UNIVERSITY OF RWANDA

COLLEGE OF SCIENCE AND TECHNOLOGY

AFRICAN CENTRE OF EXCELLENCE IN INTERNET OF THINGS

PREDICTIVE MODELLING AND ALERT SYSTEM FOR RAINFALL INDUCED LANDSLIDES

PhD. Thesis submitted in the fulfilment of requirements of award of PhD Degree in Internet of Things – Wireless Intelligent Sensor Networking.

Martin KURADUSENGE

JUNE /2022

UNIVERSITY OF RWANDA

COLLEGE OF SCIENCE AND TECHNOLOGY

AFRICAN CENTRE OF EXCELLENCE IN INTERNET OF THINGS

PREDICTIVE MODELLING AND ALERT SYSTEM FOR RAINFALL INDUCED LANDSLIDES

PhD. Thesis submitted in the fulfilment of requirements of award of PhD Degree in Internet of Things – Wireless Intelligent Sensor Networking.

Thesis Supervisor: **Prof. Santhi KUMARAN**

Thesis Co-Supervisor(s): **Dr. Zennaro Marco**

Dr. Frederic Nzanywayingoma

JUNE /2022

Martin Kuradusenge, a Ph.D. student of UR-ACEIoT student ID 218014514, successfully defended the thesis/dissertation entitled "**PREDICTIVE MODELLING AND ALERT SYSTEM FOR RAINFALL INDUCED LANDSLIDES**", which he prepared after fulfilling the requirements specified in the associated legislations, before the thesis examination members whose signatures are below.

Thesis Supervisor: **Prof Santhi Kumaran**

The Copperbelt University, Zambia

Co-Supervisor (s): **Dr. Marco Zennaro**

International Centre of Theoretical Physics, Trieste, Italy

Dr. Frederic Nzanywayingoma

University of Rwanda

Viva Voce Members:

Prof. Denis Ndanguza (chair)

University of Rwanda

Dr. Baraka Maiseli

Senior Lecturer

University of Dar es salaam

Dr. Udoinyang Godwin Inyang

Associate Professor

University of Uyo, Nigeria

Date of Submission : 16/09/2022

Date of Defense : 28/06/2022

To my spouse and children,

FOREWORD

This thesis is the final deliverable of PhD research conducted in the African Centre of Excellence in Internet of Things (ACEIoT) with the financial support of the Government of Rwanda through the African Centers of Excellence (ACEs) project.

The ACEIoT opened academic programs in 2017, just one year after the tragedies that occurred in the North and Western provinces of Rwanda in 2016, where landslides killed many people and destroyed various properties and infrastructures. Then, after being admitted into PhD program in the first cohort in 2017, we had to submit our research proposal, and I addressed my request to contribute to the landslide risk reduction through the prediction and landslide early warning system using IoT as one of the strategies to reduce the risks caused by rainfall induced landslides. The proposal was accepted by my supervisors and later presented to the doctoral committee and approved as PhD research work.

The full commitment was required to achieve the objectives of this research work. I started by analysing the landslide causal factors, then identifying the data sources. I had to meet various people in different institutions, such as the Ministry in Charge of Emergency Management, Rwanda Meteorology Agency, and local governance authorities. Secondary data have been collected from those institutions for further analysis and predictions. Various experiments have been conducted in the area of study (Gakenke district) to identify the correlation between rainfall, soil moisture and landslide occurrence. We had to spend many hours a day (in most cases, 8 hours or more) using the rainfall simulator, recording the simulated rainfall amount and observing whether the selected plot would crack or not. After that, we developed the prototype and spent a lot of time on the field testing it.

For the period of 4 years, I learnt a lot, like spending much time on the field, working hard, gaining experience from others I used to interact with during the entire process of this journey.

I would like to say thanks to the Government of Rwanda for its initiative to empower and strengthen the human capacity to promote the knowledge-based economy by delivering high quality, market-relevant postgraduate education in Rwanda.

Firstly, let me express my sincere gratitude to the first Director of the ACEIoT, Prof. Santhi Kumaran who made a successful proposal to initiate the Centre and her continued support/management of this Centre. I would also like to thank her for supervising my research for more than four years. Although she had other responsibilities as the Director of the Centre, she always tried to find a time to meet and discuss on a weekly basis, regardless of her busy agenda. Thank you for your guidance, discussions, opportunities, and other various support you provided to make this research successful.

I would like to thank my co-supervisor, Dr. Marco Zennaro, for his dedicated support and guidance during this research. Dr. Marco continuously provided technical support in different

ways, such as offering appropriate trainings. Your encouragement and always enthusiastic willingness to assist are invaluable to the achievements of this research. I also acknowledge the support of the resident co-supervisor, Dr. Nzanywayingoma Frederic. Although you came into the supervision team late, you did your best while writing this thesis.

Thanks to Professor Damien Hanyurwimfura for your guidance, inspiration, administrative support, and facilitation to accomplish different research activities. Thanks to the entire staff of the ACEIoT, especially Mr. Mugiraneza Jean Baptiste, for different support, including hosting the primary data on the server to be accessed by the reviewers and other researchers.

I would want to express my gratitude to Honorable Germaine Kamayirese, the former Minister of Emergency Management and Refugee Affairs, who took time out of her hectic schedule to meet with me to discuss my research. She gave all the essential help by assigning the appropriate ministry personnel to assist me and supplying all of the information or secondary data I required.

The objectives of this research could not be achieved without secondary data. I would therefore like to express my sincere gratitude to the leadership of the office of meteorology/Rwanda for their assistance in terms of providing all the rainfall data I needed for different analyses. Thank you, Mr. Amos Uwizeye, for your kindness and being patient by replying to numerous emails requesting data.

Thanks to Mr. Ndayambaje Godefroid and Mrs. Mukunduhirwe Benjamine, the former mayor and Vice mayor of Social Affairs of Ngororero district, respectively, for your permission to collect primary data in the district chosen as the area of study. Likewise, I would like to thank the mayor of Gakenke district, Mr. Nzamwita Deogratias, for allowing me to conduct some experiments and capture data.

Thanks to my classmate, Mr. Habiyaremye Joseph, for technical assistance. To the department of civil engineering at the University of Rwanda for the lab service of testing and analyzing the soil composition. Thanks to the other co-authors who were not part of the supervision team, Mr. Minani JB and Mr. Niyonzima Albert.

Finally, yet importantly, I am thankful to my entire family. My spouse, Ingabire Felicula, for your usual patience during this long journey and your prayers. My daughter Iza K. Emma Lise and my sons Ize K. Odilo Landry as well as Izi K. Odo Landericus without forgetting my sister Mukeshimana Francine. I apologize for being obnoxious while I was busy writing papers and couldn't allow you to converse as you desired. I'll never forget your constant encouragement, without which I would have given up on my studies long ago. You are all a part of the accomplishments that brought this lengthy journey to a close.

JUNE/2022 **Martin KURADUSENGE**

TABLE OF CONTENTS

ABBREVIATIONS

ACEIoT	African Centre of Excellence in Internet of Things
ANN	Artificial Neural Networks
AUC	Area Under the Curve
CRED	Centre of Research on the Epidemiology of Disasters
CSV	Comma-Separated Value
EDA	Exploratory Data Analysis
FAO	Food and Agriculture Organization
FNR	False Negative Rate
FPR	False Positive Rate
GIS	Geographic Information Systems
GPRS	General Packet Radio Service
GSM	Global Systems for Mobile
IDE	Integrated Development Environment
IDW	Inverse Distance Weighting
IoT	Internet of Things
LED	Light Emitting Diodes
LEWS	Landslide Early Warning System
Lo-LEWS	Local Landslide Early Warning System
LoRa	Long Range
LoWPAN	Low Power Personal Area Network
LR	Logistic Regression
LTE	Long Term Evolution
MIDIMAR	Ministry of Disasters Management and Refugees
MINEMA	Ministry in charge of Emergency
MLT	Machine Learning Technique
RAB	Rwanda Agriculture Board
RF	Random Forest

RFADT	Rotation Forest with Alternating Decision Tree
RLMUA	Rwanda Land Management and Use Authority,
ROC	Receiver Operating Characteristics
SA	Site A
SB	Site B
SC	Site C
SD	Site D
SE	Site E
SFBA	San Francisco Bay Area
SLCG	Sandy Lean Clay with Gravel
SMC	Soil Moisture Content
SMOTE	Synthetic Minority Oversampling Technique
SMS	Short Message Service
SSEWS	Site Specific Early Warning System
SVM	Support Vector Machine
SWCC	Soil Water Characteristics Curve
Te-LEWS	Territorial Landslide Early Warning System
TNR	True Negative Rate
TPR	True Positive Rate
UNDRR	United Nations Office for Disaster Risk Reduction
WMO	World Meteorological Organization
WSN	Wireless Sensors Networks.

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Natural disasters are among the leading causes of death worldwide. Statistics indicate that many people lose their lives in different disaster incidents. Damages to properties and infrastructure from this type of hazard are worth millions of dollars, and much more money is spent on disaster recovery. In many countries, including the East African region, landslides and floods are the most common natural disasters causing fatalities. In the least developed or developing countries, there are no efficient mechanisms for prediction or early warning of rainfall-induced disasters for rescuing people before the occurrence.

In Rwanda, the common strategies used to reduce the risk of landslides are to move people from high-risk zones to low-risk areas and cover the land with vegetation (forestation, grass). The first method takes a long time as it has a financial implication and citizens have to be responsible since the support from the government is limited. Today, technological solutions are available, such as wireless sensor networks. The last plays a key role in solving many problems by monitoring environmental parameters and providing alerts to the public. With the help of machine learning techniques, the prediction of disaster occurrence can be done by using historical rainfall data and landslide incidence records in the past.

This research aimed to reduce the risks of landslides by identifying and analysing the internal and external landslide-causing factors, the correlation between disasters' occurrences and the causing factors, and then designing and developing an early warning system for predicting rainfall-induced landslides. The system uses wireless sensors to collect hydrological data that is used to predict landslide incidence and alert the public before the occurrence of the incident. Rainfall, topographical, and geological data were collected, and machine learning techniques have been used to predict landslide occurrences. The wireless sensor network was designed and developed to collect real-time data, send it to the cloud where it is processed, and predict landslide incidence.

This research was conducted in 3 systematic phases: The prediction of landslide incidence using machine learning models; experimental study to determine thresholds to be used in the prototype implementation; design, development, and testing of the IoT prototype. In the first phase, Random Forest (RF) and Logistic Regression (LR) are the two machine learning models that were used for the prediction of landslides using historical data. The models' performance was evaluated using false negative rate (FNR) and the receiver operating characteristics, area under the curve (ROC-AUC).

The prediction results revealed that the antecedent rainfall has a significant impact on the occurrence of landslides. The AUC for RF was 0.995 and 0.997 for LR, whereas FNR was 4.80% and 3.84% for RF and LR, respectively. The comparative analysis showed that LR performed better than RF. The correlation between rainfall, soil moisture, and landslide

incidence was identified in the second phase of this research. The results from this phase revealed the amount of rainfall and soil moisture content inducing landslides.

The results of the experiments showed that for a particular site, the minimum time required to cause slope failure was 8h41, with an intensity of rainfall of 8 mm/hour and soil moisture levels exceeding 90% for the sensors placed more than 100 cm deep in the ground. Those thresholds were used for the early warning system prototype, and the delivery of the warning message is based on threshold values. The system prototype was successfully tested at the selected sample sites at a rate of 71.4%. The study area of this research is the Ngororero and Gakenke districts in Rwanda.

CHAPTER 1

INTRODUCTION

1.1 Background

The number of natural disasters' occurrences has been increasing over time due to unavoidable natural phenomena. Climate change is the catalyser of some hazards such as landslides, floods, droughts, storms, and winds. These disasters have a negative effect on the lives and assets that lead to the economic loss. The cost of the damage is expressed in billions of dollars every year in many countries [1,2]. The general consensus is that the extent to which properties or people are exposed to disasters increases the loss associated with them [3]. One of the most common natural disasters is landslides. The last, which is also known as "slope failure" or "debris flow", is defined as the uncontrollable downhill movement of a slope forming materials from a mountain under the influence of gravity [4].

Rainfall induced landslides occur after a certain period of precipitation (high rainfall intensity or low intensity with long duration). Landslides lead to the deaths of people or animals living at the foot of the hills or in the path of debris flow. In addition, this disaster causes damage to the properties such as infrastructure as well as the natural environment around the incident, and this leads to economic harm. For instance, a single landslide can lead to the serious damage of roads, railways, communication systems, and buildings such as schools, hospitals, etc. Hence, more funds will be allocated for reconstruction or repair of the damaged infrastructure. According to the Centre of Research on the Epidemiology of Disasters (CRED), Landslides are responsible for 17% of all fatalities worldwide [5].

1.2 Landslides in the context of Rwanda

Rwanda is a mountainous country with an average altitude of 1,700 meters [6]. The highest peak is found at Kalisimbi volcano, with an elevation of 4,507 meters [7]. The steep slopes that characterize Rwanda's topography render it vulnerable to landslides, particularly in the west and north regions (Figure 1). There are four distinct climatic seasons in the country: (i) the short rainy season (September–November); (ii) the short dry season (December–February); (ii) the long rainy season (March–May); (iv) the long dry season (June–August) (June-August).

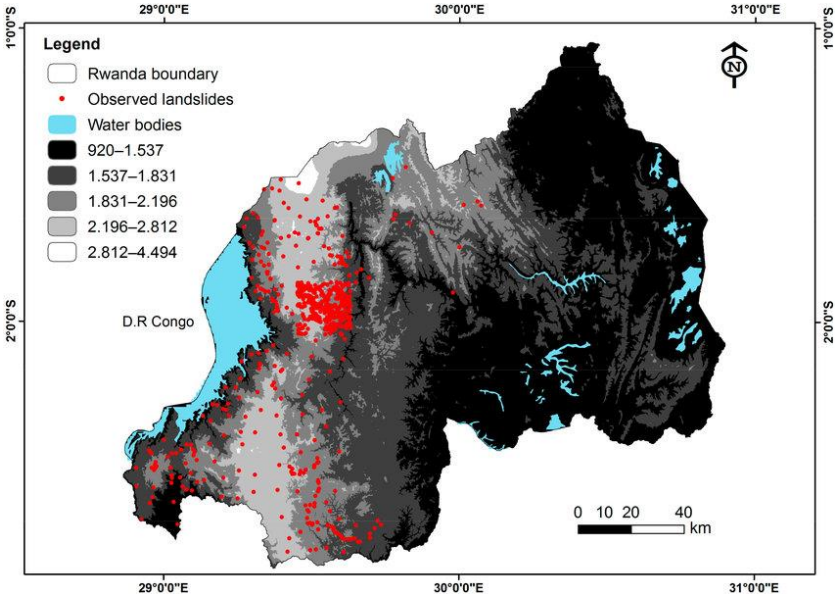


Figure 1. Landslides' inventory in Rwanda and elevation [8]

The long rainy season is characterised by abundant precipitation, especially in the western and central parts of the country (Figure 2), and the regions with steep slopes are much affected by landslides. The last led to the loss of lives, the destruction of infrastructure, and damage to citizens' properties and the environment. For instance, on May 6-7, 2022, the ministry in charge of emergency management recorded 65 deaths in the entire country, many injuries as well as livestock, hundreds of houses destroyed, and many hectares of land overlapped by landslide debris [9].

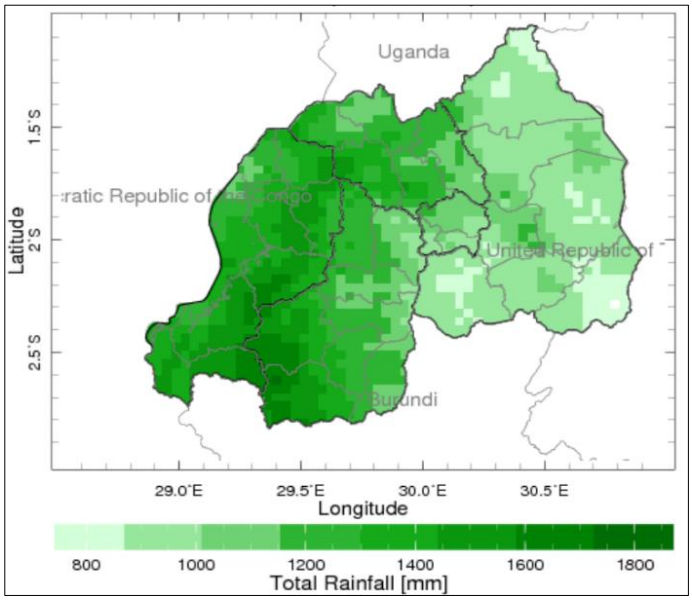


Figure 2. Mean annual rainfall in Rwanda 2011-2021 (source: Meteorology Rwanda)

1.3 Motivation

According to the World Meteorological Organization (WMO), weather related disasters have increased over the past half-century, causing many deaths and about US $202 billion of loss daily. In its report, WMO said that disasters related to climate change were amplified by a factor of five over the last 50 years, but due to the early warning systems, the deaths related to the threats were reduced almost three-fold [10]. Landslides form one of the leading natural hazards causing loss of life and damage to infrastructure worldwide [11]. Landslides strike several parts of Rwanda almost every year during the period of heavy rainfall. Some of the victims drowned in the debris flows or floodwater, while others died after houses collapsed under the heavy rainfall.

Disasters which ravaged several parts of the country and caused many deaths (Figure 3) and damage. For instance, during the year 2017, many damages were estimated at RWF 6.7 billion, according to the Ministry in Charge of Emergency Management [12]. Another example, MINEMA reported that during the rainfall events in 2017, around 5,000 houses were damaged and almost 5,111 hectares of different crops were destroyed [13]. The Northern Province was hit the worst and had the greatest fatality rate [14,15].

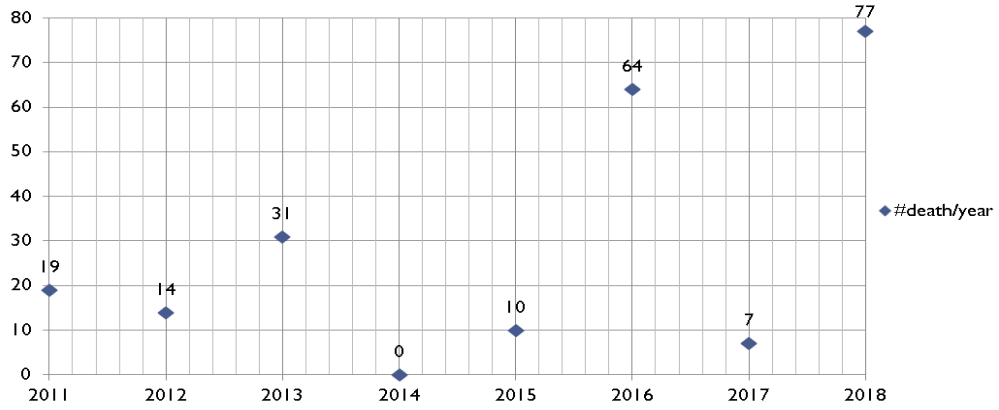


Figure 3. Annual statistics of deaths caused by landslides (2011-2018)

Due to the countless hills and steep slopes that characterize Rwanda, several areas are prone to landslides (Figure 4). The most vulnerable areas are in the northwestern districts. During the last 10 years, the most affected districts were Ngororero and Gakenke. This is due to the geo-environmental factors that characterize these districts (and the entire region). The most common factors that cause this area to be vulnerable are high slope inclinations, soil types, land use, and many others (Figure 4).

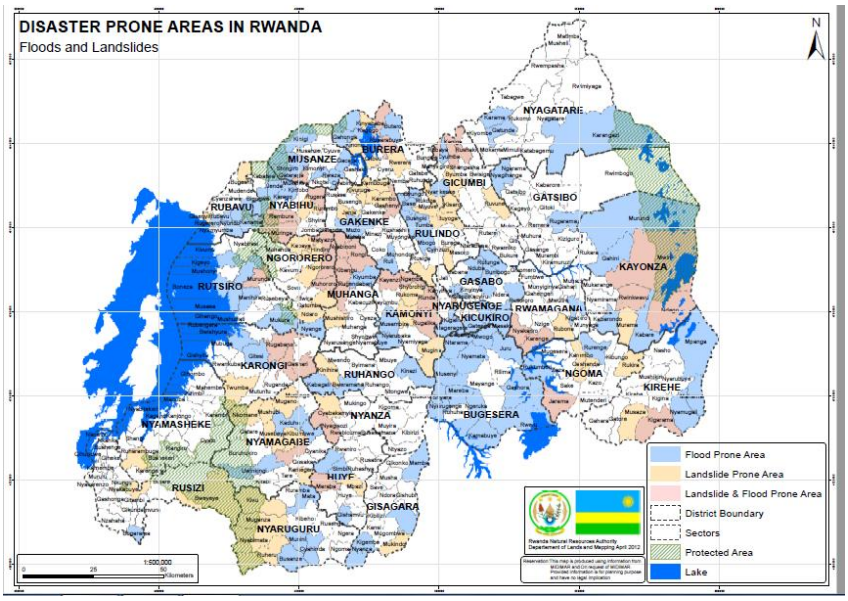


Figure 4. Disaster prone areas in Rwanda: Floods and Landslides (source: MINEMA, 2012)

Rainwater causes changes in soil properties, leading to slope failures, which are also linked to the different activities on land such as agriculture, deforestation, excavation, and other internal factors. These landslide contributing factors might provide crucial information for reducing or mitigating the hazards posed by this natural disaster.

The current strategies to cope with landslides in Rwanda are: construction and maintenance of drainages, relocation of populations in low-risk areas, measures to reduce erosion (reforestation, implementation of terraces), and crop diversification to reduce runoff. Although there are different studies about risk reduction and various strategies to solve this problem, there is still a gap in those studies because of poor coverage of the existing landslide early warning system (LEWS), since most of them are regional. The reliability of LEWS depends on the coverage area. Yet, the coverage of the regional LEWS used to be wide because they used weather stations [16–19].

The Internet of Things (IoT) is one of the emerging technologies that can be used to address the issue of landslides by remotely collecting data and delivering them into the LEWS to alert citizens of the specific area. The rationale for the use of IoT for disaster monitoring and prediction is that this technology can rapidly respond to the change of data, therefore providing accurate information about the imminent disaster prior to its occurrence. By using IoT, many parameters can be monitored, which increases the accuracy of prediction. The previous studies that used the IoT did not consider the parameters in the specific regions of the studies [20]. In this research, various parameters have been considered by conducting experiments in the respective areas of study, which improved the accuracy of prediction.

1.4 Problem statement

Many people suffer because of unpredicted disasters like landslides, floods, and droughts due to the prolonged period of abnormal high or low rainfall. If proper preparation is made against the impending landslide, the disaster's effects can be reduced. Although various studies are ongoing, and elsewhere many early warning systems are being developed, information about disasters is not provided before the occurrence in Rwanda. Citizens are communicated about disasters by means of radio broadcasts, where in most cases, the information is about weather forecasts and no warnings about imminent landslide incidences are provided before their occurrence [21]. The most efficient way to reduce the risks caused by this hazard is to deliver an early alert to the community. However, many proposed or developed systems do not provide a reliable solution to the problem due to the prediction capability caused by unconsidered parameters (features) during their design or coverage incapability.

This research focused on identifying an appropriate prediction model by using machine learning techniques and developing an early warning system using wireless sensors to collect hydrological data in a real-time pattern and provide a warning message to the people located in the prone areas.

1.5 Research aim and objectives

This research is aimed at reducing the risks caused by rainfall-induced shallow landslides by using a site-specific landslide early warning system (SSLEWS).

The specific objectives are:

1. Identifying the best prediction model for landslides and their causal factors
2. Determining the correlation between rainfall, soil moisture, and the occurrence of landslides
3. Developing and testing a prototype for the landslide early warning system

1.6 Methodology

This research was carried out in three systematic phases. It consisted of preliminary literature about the landslide risk reduction techniques and early warning systems used elsewhere. Therefore, phase I was aimed at assessing the existing strategies for landslide risk reduction in the study area. This stage also consisted of finding the best parameters to be used in the landslide predictions as well as identifying the best prediction model (machine learning technique). The data were collected based on the identified (available) parameters and machine learning techniques were used to train those data. Then, the best predicting model and parameters were identified.

The second phase consisted of determining the thresholds that could be used for the landslide early warning systems. At this stage, the field experimental study was done on various sites

characterized by the landslide causing factors identified in phase I. To conduct those experiments, we used the soil moisture sensors, a rain gauge, and a rainfall simulator. In total, eleven (11) sites were selected based on their landslides' incidences history, and the results of these experiments were the thresholds for sites that resulted in the slope failure.

In phase III, an IoT prototype for LEWS was developed by using the thresholds found in the previous phase. This prototype was successfully tested on sites characterized by similar geo-features to those used in the second phase.

1.7 Thesis outline

Chapter 1 comprises the background of natural disasters in general and landslides in the context of Rwanda. The research motivation and problem statement are clearly explained. At the end of this chapter, the research aim and specific objectives are provided.

Chapter 2 comprises the literature review on landslides, their causal factors, and the mechanisms used to reduce their risks.

Chapter 3 is about the study of rainfall induced landslides' prediction using machine learning techniques. Two machine learning models are used to predict landslide incidences using rainfall data and other features. The performance of the models was evaluated, compared, and a recommendation is provided for the landslide early warning system.

In Chapter 4, the determination of the thresholds for the soil moisture content and rainfall for landslide prediction was done through the experimental study. The chapter comprises the description of different sites with various geo-morphological and environmental characteristics. The results from the experiments are presented with the identified threshold values of soil moisture and rainfall for LEWS.

Chapter 5 describes the LEWS designed and developed to reduce the risks of landslide incidences. The components of the system are described, and the test results are presented.

Chapter 6 concludes the thesis.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This planet confronts a great loss of lives, properties, and infrastructure due to natural disasters such as rainfall-induced landslides. Most natural disasters are unpredictable and may occur in a very short period of time. The earth's movement is caused by various factors, including geological, morphological, environmental, hydrological, and anthropogenic aspects [22]. More extensively, landslide determining factors are variations in rainfall, slope, aspect, and geology [23]. The occurrence of landslides is more frequent in hilly regions with steep slopes where rainfall used to be abundant. This natural disaster has high impacts when it occurs in clusters. Their impacts touch on three main sectors: economy, lives, and environment.

The slope failure involves two main issues: (i) the internal strength of the slope material, (ii) the friction of the material of the slope. The failure occurs when the gravitational (driving) forces exceed the shear strength to pull the slope making material downhill [22]. Hydrological factors are common to induce landslides. This is because water adds substantial weight to the slope as it penetrates the ground, leading to increased gravitational force. Another role that water plays on the ground is lowering the strength of the earth's material as well as reducing friction, which makes the movement of the material downwards easier [22].

2.2 Landslides categorization

Classification of landslides can be done by considering different parameters. For instance, type of movement (flow, sliding, falling), earth material (debris, soil, rock), movement speed (slow or fast) [22,24]. They can also be determined based on the amount of water.

Considering the speed and size, landslides can be classified as shallow or deep-seated. Shallow landslides, also known as mudslides, are debris-flow slides that occur within the topsoil layer between 1.5 and 10 meters deep [25,26]. They are usually initiated by intense precipitation over several hours or a few days that causes soil saturation and loosening, which then triggers sliding. They occur as a fast moving debris flows downhills or slumps along roadways or any other man made cut slopes like those associated with the site preparation during building houses. Contrary, deep-seated landslides are slow moving debris-flow slides rooted in bedrock, usually on a large scale, devastating infrastructure, environment, and buildings. They can cover a large area and are devastating. The depth of this type of slope failure can range from 10 meters and above. Deep seated landslides are caused by the change of geologic and hydrologic processes like an earthquake or increased amount of earth water [3].

Regarding the type of movement, landslides can be classified as flows, slides, topples or rock falls. Flows are defined as a concoction of water, soil, rock, and (or) debris moving speedily

downhill [26]. Flows can also be of different types, such as earth flows, debris flows, debris avalanches, lahars, lateral spreads, soil creep, etc. Slides are defined as a downhill movement of soil or rock along a surface and can be either shallow or deep-seated. This type of slope failure can also be categorized as transitional slides, rotational slides (slumps), or brock slides [26].

2.3 Landslides' risks reductions strategies

The main constraints of landslides risk mitigations are the lack of information about landslides risks in the area, how those risks change versus the climate change, and information about land-use at the local scale [5]. Broadly, strategies used to mitigate or reduce the risks caused by landslides can be categorized into: (i) landslide susceptibility assessment, ii) structured measures that consist of slope stabilization structures, (iii) non-structured measures comprising land use planning and digital technology such as early warning systems for awareness of the imminent hazard prior to the occurrence.

2.3.1 Landslide susceptibility assessment

Although landslides are unavoidable and undetected, their risks can be reduced in different ways. Some of the techniques used to minimize the risks include hazard zonation or susceptibility assessment of slope failures in different areas, whereby the analysis comes with an indication of the areas that are highly susceptible and those with moderate to low susceptibility. The techniques used for vulnerability zonation may comprise landslide inventories, statistical models, vulnerability zonation, probabilistic methods, etc. [27].

Many countries possess building acts stating the rules and criteria for site selection of buildings such as houses, schools, hospitals, and other infrastructure. The rules also specifically state that no building should be constructed in the hazardous area. The identification of hazardous zones and their respective levels is done through the landslides' susceptibility assessment, then the map is established. Thereafter, the analyses may be used for decision making in terms of infrastructure settling or construction.

2.3.2 Landslides risks reduction by structural techniques

Structural measures for landslides' mitigations consist of all techniques used to reduce water infiltration into the ground. It comprises the construction of strong walls and draining the rainwater into constructed channels. Forestation or grass planting are the most effective and cheapest techniques to deal with landslides' risks reduction. Vegetation increases the shear strength by adding a water resistive layer on the topsoil that will limit the amount of rainwater penetrating the soil.

2.3.3 Digital Technology

Various studies and research have been conducted worldwide using the latest technologies of wireless sensor networks (WSN). Worldwide, the development and use of weather-related landslide early warning systems are of great importance to reduce the risks caused by slope failures triggered by abundant precipitation [28]. LEWSs operate at either a regional (territorial) or local scale. For instance, on a territorial scale, LEWS for Java Island (Indonesia) can be mentioned [29,30]. The system predicts slope failures based on the rainfall threshold and can detect the incidence using an extensometer [30]. The initial LEWS for Indonesia could have provide a warning some minutes or a few hours prior to the incident. The new version of Indonesian LEWS can provide warning four days before the landslide occurrence [28].

Other regional systems developed a long time ago include the LEWS for Hong Kong island, more than 40 years operational but with many revisions [31]. This system uses rainfall data from 20 rain gauges [32]. The LEWS for the San Francisco Bay Area (SFBA) that began operating in 1986 [32]. The SFBA system uses the hourly rainfall from 60 rain gauges (each rain gauge covering 300 Km^2). The values from rainfall measurements are compared with the empirical thresholds for a possible slope failure incidence in the SFBA [33]. The near real-time debris-flow warning system for Vancouver (Canada) operated in 2011 [34]. The Vancouver system consisted of four precipitation catchments covering 600 Km^2 and using 25 rainfall variables such as cumulated rainfall, rainfall duration, rainfall intensity, and various antecedent rainfall [32].

The LEWS for Rio de Janeiro (Brazil) established in 1996 uses 1, 24, and 96 hours rainfall data measured by meteorological stations (each station covers 37 km^2) compared with empirical thresholds for the landslides' prediction [32]. The LEWS for south Taiwan uses 3 hours mean rainfall and cumulated 24-h rainfall measurements from 96 rain gauge stations (each covering 77 km^2 approximately), susceptibility zonation (low, medium, and high class), and empirical thresholds [35]. Different regional LEWS are found in Italy, such as Emelia-Romagna and Piedmont [16,17] (north Italy), Umbria and Tuscany [18,19] (Central Italy), Sicily, Apulia and Sardinia [36,37] (southern Italy).

The local LEWSs use the WSN, which is one of the major technologies used for remote sensing or real-time monitoring and provides information for alerting. They are basically designed to operate in a well-identified area where they provide warnings to the local citizens [38]. In addition to the operational LEWS, several studies proposed systems using the Internet of Things. Aji F. et al. [39] proposed a system made of a potentiometer and a soil moisture sensor controlled by the Arduino (ATMega 328) microcontroller. The warning is provided using light emitting diodes (LEDs) and an alarm.

A landslide detection system developed by [40] consists of a lower layer made of wireless sensor nodes for data collection, and an upper layer to aggregate data to the sink node (gateway) at the

deployment site. This system connects to the data management center through Wi-Fi and satellite terminals for real time data transmission. Ravi et al. discussed a proto-model of a node design for landslide monitoring in heavy rainfall and hilly areas. The WSN deployment enables access to many sensors' information by using Ethernet, Wi-Fi, satellite or any other wireless protocol [41]. Yong et al. developed a landslide monitoring system based on a wireless sensor network that consists of geological sensor nodes and camera sensor nodes. The anomaly detection based on geological sensor nodes is analyzed in terms of its self-learning threshold [42].

2.3.4 Machine learning for landslide prediction

Machine learning is a cutting-edge analytics method that has been widely employed in the prevention of landslides [43]. Machine learning plays an important role in landslide risk reduction in different ways, such as detection, susceptibility assessment, and prediction using early warning systems. Machine learning for landslide detection entails training and testing of the landslide inventory dataset. The last consists of landslide inventory produced through remote sensing technology that captures images, positions, areas, and the level of demolition [43].

The likelihood of landslide occurrence in a given location can be determined by landslide susceptibility based on meteorological and topographical and topographical parameters. Based on the landslide inventory, the topography is classified into several levels of danger [44]. Machine learning models process and produce susceptibility judgments on the study region using the acquired data. The level of contribution of each causing factor is calculated so that those with an insignificant contribution can be eliminated. The most common supervised models are Logistic Regression (LR), Artificial Neural Network (ANN), Naïve Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT) [45–49,49–53].

Landslide prediction is an important component of operational early warning systems. Time-series data on landslide displacement can properly reflect slope failures [54]. Real-time data such as precipitation and groundwater data can be collected by remote sensing systems or IoT sensors and play a crucial input in predicting slope failures. SVM [55], ANN [56], RF [57], and other conventional machine learning models have been applied to predict landslide events using precipitation and groundwater data. In experiments by Xie P. et al. [58] and Yang B. et al. [59], for example, the long short-term memory (LSTM) achieved good results in landslide prediction.

2.3.5 IoT for landslide prediction

A reliable LEWS should have the following components: data collection, data transmission, modeling, warning, and response [60]. The data collection and transmission components are performed by IoT or remote sensing systems, whereas the modelling and warning are performed by machine learning. The failure or bad performance of those components affects the LEWS reliability. However, where machine learning is not used in LEWS development, a prior study should be conducted to identify the landslide triggering parameters. Then, the real-time data collected from the IoT components will be processed and compared with the thresholds to learn the potentiality of the landslide occurrence. Thresholds can be determined from the experimental study or by using machine learning methods.

As described in section 2.3.3, remote sensing systems are suited for regional or territorial LEWS. Sensor-based IoT, on the other hand, can provide a dependable solution for local LEWS. Therefore, local sensors can increase the reliability of the regional prediction system. Regional deployment of local sensors is not really possible when low-cost, low-maintenance sensors are mass-deployed [61]. According to Thomas et al. [62], a single soil moisture sensor can cover a broader region and better depict soil moisture levels than a satellite soil moisture product.

CHAPTER 3

PREDICTION OF RAINFALL INDUCED SHALLOW LANDSLIDES USING MACHINE LEARNING MODELS

Landslides fall under natural, unpredictable and most distractive disasters. Hence, early warning systems for such disasters can alert people and save lives. Some of the recent early warning models make use of the Internet of Things to monitor the environmental parameters to predict disasters. Some other models use machine learning techniques (MLT) to analyse rainfall data along with some internal parameters to predict these hazards. The prediction capabilities of the existing models and systems are limited in terms of their accuracy. In this research paper, two prediction modelling approaches, namely random forest (RF) and logistic regression (LR), are proposed. These approaches use rainfall datasets as well as various other internal and external parameters for landslide prediction and hence improve the accuracy. Moreover, the prediction performance of these approaches is further improved by using antecedent cumulative rainfall data. These models are evaluated using the receiver operating characteristics, area under the curve (ROC-AUC) and false negative rate (FNR) to measure the landslide cases that were not reported. When antecedent rainfall data was included in the prediction, both models (RF and LR) performed better, with an AUC of 0.995 and 0.997, respectively. The results proved that there is a good correlation between antecedent precipitation and landslide occurrence rather than between one-day rainfall and landslide occurrence. In terms of incorrect predictions, RF and LR improved FNR to 10.58% and 5.77% respectively. It is also noted that among the various internal factors used for prediction, slope angle has the highest impact of all the other factors. Compared to both the models, the LR model's performance is better in terms of FNR and it could be preferred for landslide prediction and early warning. The LR model's incorrect prediction rate FNR = 9.61% without including antecedent precipitation data and 3.84% including antecedent precipitation data.

3.1 Introduction

Landslides and floods are the common natural disasters that strike the northern and western provinces of Rwanda due to their topographical, geological, and climatic profile [11,63]. According to the National Risk Atlas of Rwanda report published by the Ministry in Charge of Emergency Management (MINEMA), 42% of areas are classified as moderate to very susceptible areas to landslides [12]. Every year, during the rainfall period, landslides affect many people in mountainous regions [64]. These disasters led to the loss of lives and left many homeless and without a livelihood. Since the establishment of an institution in charge of disaster management (MINEMA) in 2010, systematic records indicate that there have been 227 deaths and 160 injuries from 2011–2018 (June). Also, many houses have collapsed, and many hectares

of crops have been washed away by landslides and floods [8]. Table 1 and Figure 5 summarize the statistics of deaths caused by landslides during the period of 2011–2018 in Rwanda.

Table 1. Number of deaths caused by landslides in Rwanda (2011–2018).

Year	2011	2012	2013	2014	2015	2016	2017	2018	Total
Number of deaths	19	14	31	0	10	64	7	77	222

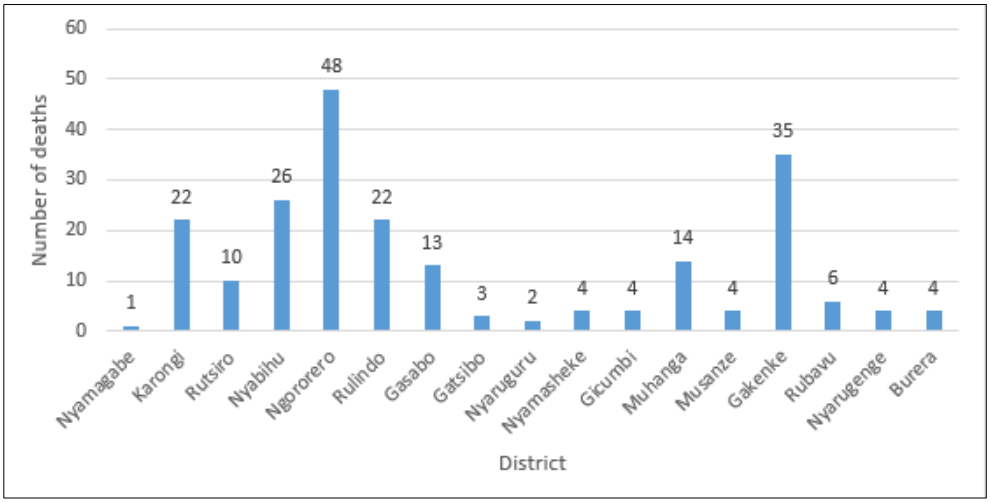


Figure 5. Number of deaths caused by landslides per district.

In most cases, landslides and floods in Rwanda occur in a cascading manner as debris is dumped into rivers, which in turn causes riverine floods [65]. One example is the case of floods and landslides that occurred in 2016 and 2018, when many people died after heavy rainfall that caused widespread flooding and landslides across parts of Rwanda. At least 222 people have died in different landslide events since 2011–2018, where victims drowned in floodwater or others died after houses collapsed due to debris movement or landslides caused by heavy rainfall. The types of floods most threatening Rwanda are riverine floods due to its dense river network and large wetlands [66].

The triggering factor of a landslide is rainfall infiltration into the soil, making the groundwater level increase, resulting in a reduction of the shear strength, which is also closely related to antecedent rainfall, cumulative rainfall, and rainfall duration [67]. The slope instability increases with high intensity or long duration rainfall but does not relate to rainfall alone. It is affected by other factors such as lithological material, type of soil and depth, the surrounding vegetation, slope inclination or aspect, curvature, altitude, land use patterns, and drainage networks [8,68,69].

Although the rainfall intensity may be the same in different regions, landslides may or may not occur according to the geological and topographical characteristics. Therefore, rainfall intensity

cannot be the only cause of landslides [69]. According to the study carried out by MINEMA, the following landslide contributing factors have been assigned weights based on the past landslide characteristics in Rwanda: rainfall (20%), slope (20%), altitude (15%), soil type (14%), lithology (10%), land cover (9%), soil depth (7%), and distance to the main roads (5%) [66]. Various machine learning models and assessment studies have been carried out on both landslide susceptibility assessments and flood predictions [69–71]. The models are built based on the assumption that landslides are more likely to occur in situations similar to those of past landslides [11].

Although there have been numerous researches undertaken on disaster susceptibility assessment and prediction models, the literature about landslides data for Rwanda is scarce [66]. Hence, Rwanda lacks efficient early warnings about disasters and therefore, the risks of vulnerability to incidences are high [72]. Therefore, there is a need for continuous research on the prediction of disasters due to landslides and floods to reduce the risks. Recently, a few studies were conducted on landslide susceptibility in Rwanda [8], but none of them were about disaster prediction for early warning and risk reduction.

Elsewhere, numerous machine learning techniques have been used to predict landslide occurrence but most of them have been focusing on disaster susceptibility mapping using internal (geological, topographical, environmental) factors without taking into consideration the triggering internal factors such as rainfall [21,73]. Those that include rainfall in their dataset do not consider antecedent rainfall or vice versa [8,74,75], while others use both daily and cumulated previous precipitations without internal conditioning parameters [76]. False negative rate was not considered to measure the performance of the models, yet it is a crucial evaluation metric in landslide predictions.

The main purpose of this research is to improve the performance of the prediction models by including the antecedent rainfall data among other parameters used in the previous studies, such as daily rainfall, hill slope angle, soil type, soil depth, and land cover. Another objective is to minimize the incorrect predictions (FNR), as this is a very important metric to be considered since it counts landslide cases that have not been seen by the prediction model and has significance in early warning systems. The preferred MLTs are random forest (RF) and logistic regression (LR). These two MLTs were selected among others because of being among the most broadly used in landslide prediction [64] and due to their ability to deal with discrete and continuous data for classification problems. Besides, LR calculates regression coefficients while RF can show the importance of different parameters. Its training speed is high and the computational cost is low [64]. This study is aimed at: 1) analyzing the correlation between historical rainfall data and other topographical and geological factors impacting landslide occurrence in Rwanda; and 2) proposing a machine learning model for the prediction of this disaster which can be used for early warning. This study is geographically limited to the Ngororero district in Rwanda, and the period of study was from 2011–2018.

3.2 Materials and Methods

3.2.1 Area of Study

The district of Ngororero is one of the seven districts of the western province. The district is situated in the northwestern region of Rwanda. The district has a surface area of around 679 km2, and is composed of 13 administrative sectors, 73 administrative cells, and 419 villages. The district shares borders with five other districts. The district has a relief characterized by high mountains with very steep slopes that flow into valleys. The altitude varies between 1460 m and 2883 m above sea level (Figure 6), the highest point being on Bweru Mountain, situated in Muhanda sector at 2883.4 m of altitude. The average annual temperature is 18 °C which varies with the altitude. The average altitude is 1500 m.

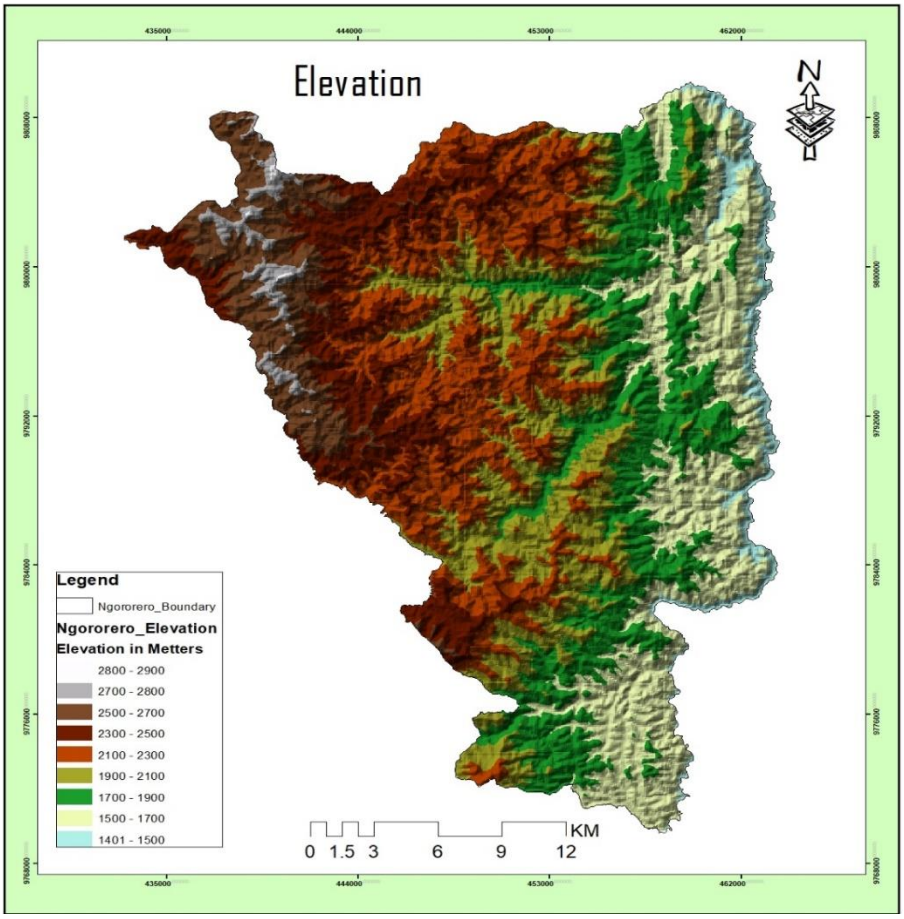


Figure 6. Study area, elevation map.

The climate of the region is more of the tropical type with four distinct seasons. (1) Short rainy season of October–December corresponding to the agriculture season A. (2) Short dry season of January–February. (3) Long rainy season of March–June corresponding to the agriculture season B. (4) Long dry season of July–September corresponding to the swamp agriculture

season C. Rainfall is regular, with a rainfall of 1527.7 mm per year, although irregularities are recorded sometimes with shortages or excess rainfall. Due to heavy rainfall and the topographic structure, which is characterized by a steep slope, the district is often affected by landslides (Figure 7) and floods almost every year during the period of heavy rainfall. In the past, some of the victims drowned in floodwater, others died after houses collapsed under heavy rainfall or landslides [77].

Figure 7. Photographs of some landslides in the study area.

3.2.2. Data Acquisition and Landslide Inventory

3.2.2.1. Data Collection

Site visits have been conducted in the study area for primary data collection where different landslide-prone areas were visited during the period of 9 months (October 2018–June 2019) to collect GPS coordinates of landslide locations and investigate contributory factors. In addition to the primary data collection, historical data about landslide incidence records have been collected from the ministry in charge of emergency management (MINEMA). Since this ministry was established in 2010, no landslide data before 2011 could be found. Therefore, the chosen period for the study is 2011–2018. Daily rainfall data were collected from Rwanda's office of Meteorology. Soil type, land cover, and slope inclination were collected from different government institutions: Rwanda Land Management and Use Authority, Rwanda Agriculture Board, and the Centre for Geographic Information Systems and Remote Sensing (CGIS) at the University of Rwanda. It has been realized that landslides occur under the same or similar conditions as those that caused them in the past [68]. Precipitation is the common internal triggering factor of landslides in Rwanda. However, depending on the selected disaster-prone area, the conditioning factors may vary.

3.2.2.2. Dataset

The internal factors most contributing to the landslide occurrence in Rwanda are slope angle, soil type/texture, soil depth, and land cover while rainfall is the external triggering factor [66]. The daily rainfall data and other parameters collected from stage 1 have been arranged in one datasheet. Geo-topographical data were extracted from Geographic Information Systems (GIS) data.

3.2.2.3. Rainfall

Daily historical rainfall data (2011–2018) for the study area were collected from the Rwanda Office of Meteorology. The dataset was made up of daily precipitation from Sovu and Muramba weather stations located in the Ngororero district. The number of rainfall data records from each station are 1893 and 2077 entries respectively. Sovu station had 978 rainfall instances while Muramba had 1305 (Figure 8). Because no landslides occurred during the dry seasons in the study area, some non-rainfall days were not included in this dataset. Therefore, 3970 instances (rainfall and non-rainfall) have been used.

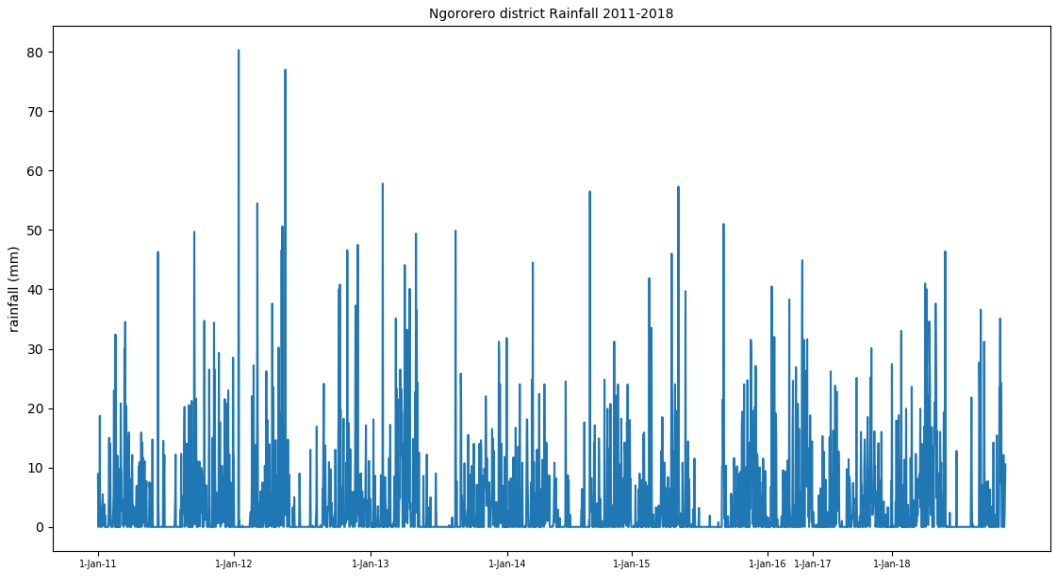


Figure 8. Rainfall in the study area.

The two rainfall gauge stations were not sufficient for rainfall data at every location in the study area. To estimate the rainfall in the areas distant from the rainfall gauge stations, the data from neighboring weather stations (Figure 9) were collected and the data interpolated using the inverse distance weighting (IDW) method.

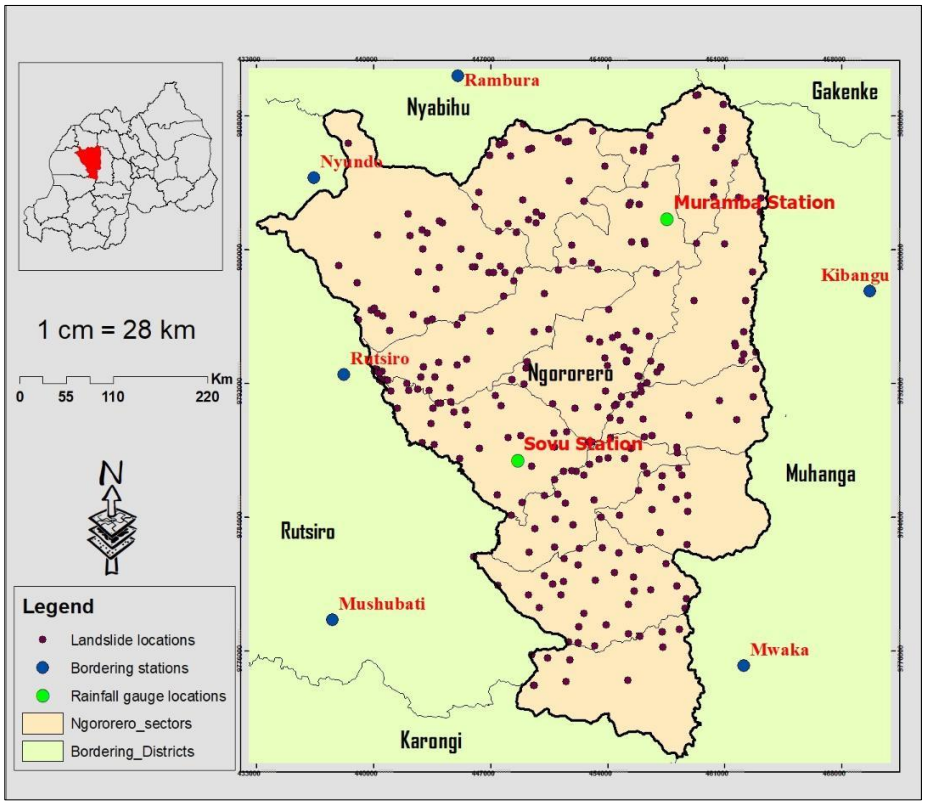


Figure 9. Landslide inventory in the study area and rain gauge stations.

The IDW was chosen because it is a commonly used technique for the estimation of missing data in hydrology and geographical sciences [78]. It is based on the functions of the inverse distances in which the weights are defined by the opposite of the distances, and normalized so that their sum equals one [79]:

$$Z(S_0) = \sum_{i=1}^{N} \lambda_i Z(S_i) \quad (1)$$

where $Z(S_0)$ represents the interpolated value at point S0, Z(Si) represents the observed value at point Si, n is the number of observations, and λi is the weight. The weights decrease as the distance increases. The weights λi can be calculated as follows:

$$\lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^{N} d_{i0}^{-p\prime}} \Rightarrow \sum_{i=1}^{N} \lambda_i = 1 \quad (2)$$

where p is a power and di0 is the distance between a target and observations. Rainfall data from the two weather stations in the study area and those in three neighboring districts (Figure 9) were used as input into the IDW model, and a GIS software tool was used for implementation.

3.2.2.4. Antecedent Rainfall

Analysis indicates that landslides can be triggered by 1-day of prolonged rainfall or by many consecutive days. From a physical viewpoint, the antecedent rainfall determines the initial water content and matric suction of the soil, but the effects of the antecedent rainfall last for a certain period of time [69]. From the rainfall dataset, 5-days antecedent rainfall data has been computed and used as an additional parameter. This parameter has the same size (number of records) as that of daily rainfall.

3.2.2.5. Slope

One of the most important landslides' causal factors is slope [80]. An increase in the slope decreases its stability if the soil depth is sufficient [81]. According to the "National risk atlas of Rwanda" by the ministry in charge of emergency management, the slope is classified into ranges from 0 to 10 according to the steepness angle [66].

3.2.2.6. Soil Type

Based on the grain-size distribution analysis and prevailing range of particle sizes, soils in the study area are classified into three categories: sand, silt, and clay [66].

3.2.2.7. Soil Depth

This was considered as one of the landslide's causal factors. Three classes of soil depth were used in the dataset according to MINEMA soil depth classification [66].

3.2.2.8. Land Cover

Various past studies have pointed out that permanently covered lands are more protected against landslides than non-covered ones. According to their potential influences, there are six main types of vegetation in Rwanda [66], but three are dominant in the area of study, as shown in Table 2.

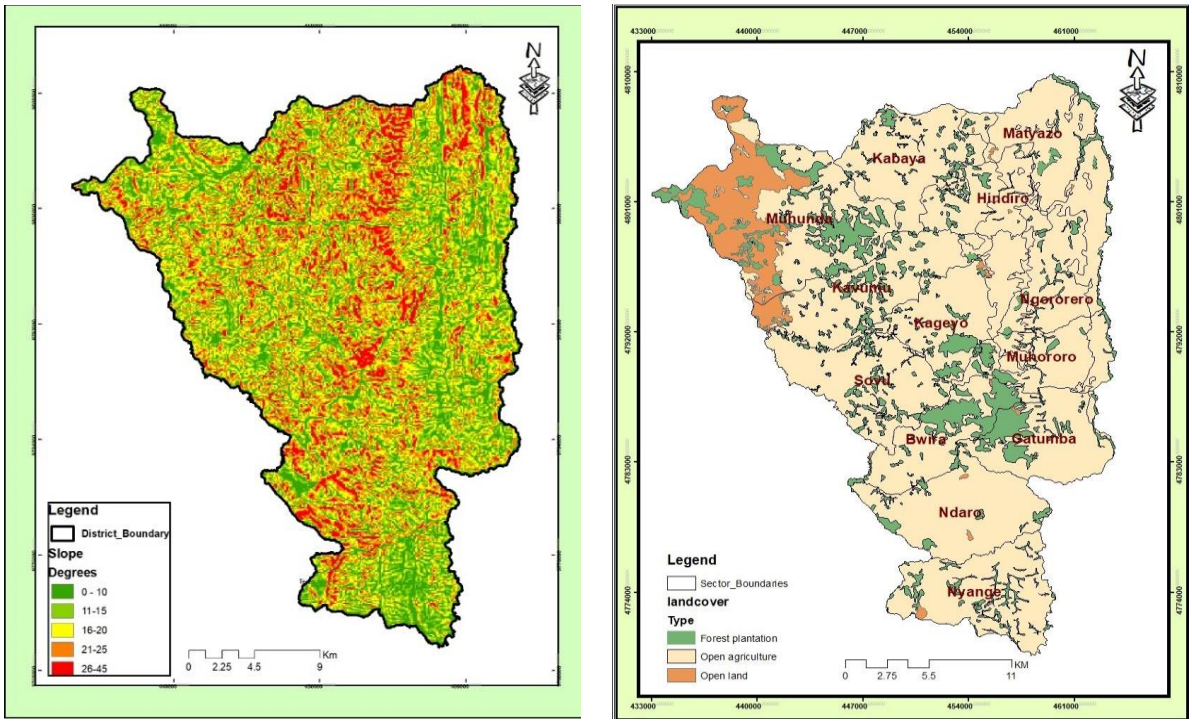
Table 2. Summary of classification of key parameters in the study area and their standardized scores used in the model.

Slope		**Soil**				**Land**	
Slope Angle (Degree)	**Score**	**Soil Type**	**Score**	**Soil Depth (cm)**	**Score**	**Land Cover**	**Score**
0–10	0	Clay	0	<50	1	Forest plantation	3
>10–15	1	Sand	4	>50–100	4	Agriculture	7

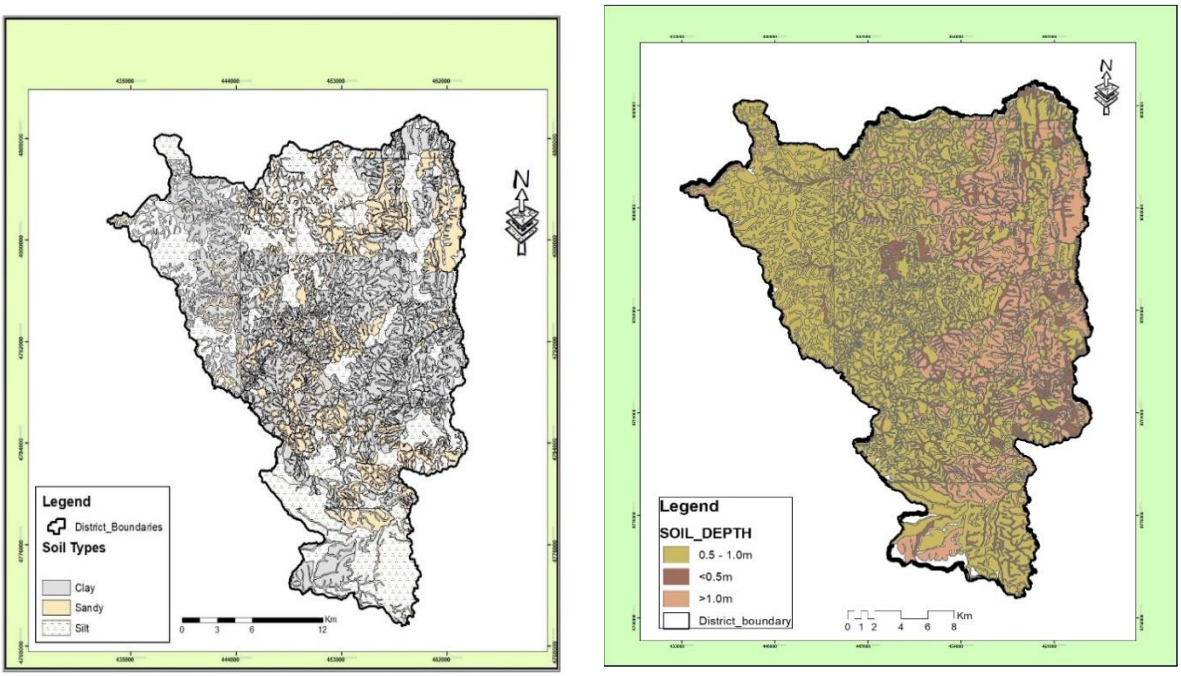
>15–20	4	Silt	6	>100	10	Open land	10
>20–25	6						
>25–45	10						

All the data in Table 2 have been collected as secondary data. Slope and soil depth are continuous, while on the other hand, the soil type and land cover are categorical. Both datasets were obtained from the source (government institutions) with the assigned scores (claimed to be FAO scores) for the purpose of getting qualitative data as defined in some official documents [66]. The standard scores were used as categorical data during the model training process. Figure 10 shows the geographical characteristics (as defined in the area of study).

The complete dataset was made up of 3970 daily rainfall, 5 slope categories, 3 soil type categories, 3 soil depth categories, 3 land cover categories. Then, the entire dataset was made of 535,950 records.



(**a**) (**b**)



(**c**) (**d**)

Figure 10. Geographical characteristics in the study area: (a) slope, (b) land cover, (c) soil type, (d) soil depth.

3.2.2.9. Landslide Incidences

Data on landslide events, including their dates of occurrence, were obtained from the ministry in charge of emergency management (MINEMA) and coded as zero (0) for "no landside event" or one (1) for "landslide event". The datasheet was converted to the comma-separated value (CSV) format for the purpose of being used in the machine learning algorithm. At this stage, the missing values were handled using the fill in missing values with the median technique, and labels were removed from the features.

3.2.2.10. Splitting Dataset into a Training and a Test Dataset

The dataset was divided into a training set and a test set at different proportions.

3.2.2.11. Training and Testing the Models

The training dataset was used to train two machine learning models, and predictions were made using the test dataset. Test labels were used for the accuracy calculation.

3.2.2.12. Results Analysis

Based on the performance, the best model will be recommended to be implemented for an early warning system.

3.2.3. Machine Learning Models

This study aims to apply two machine learning techniques and choose the best performing one for predicting landslide incidences in Rwanda, which can be used for early warning to reduce risks. Two classification models that have been selected to accomplish this research area are random forest (RF) and logistic regression (LR). The simplicity or complexity of each MLT differs from one to another. Hyper-parameter tuning has been done to optimize the performance of each. These classification models were chosen because this research deals with the binary classification problem. The complete research methodology is indicated in Figure 11.

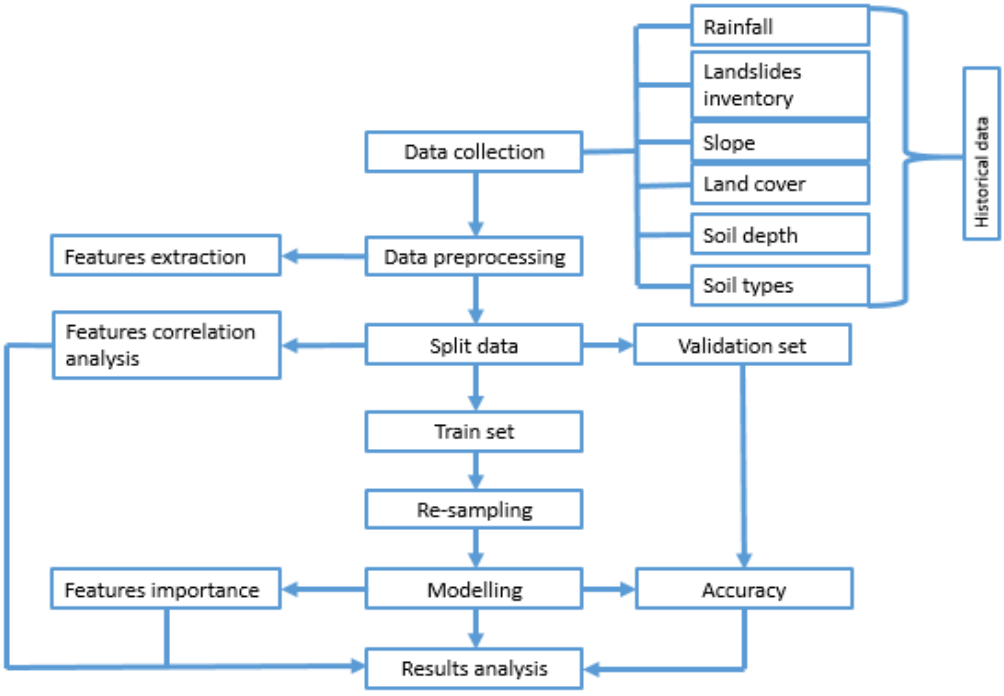


Figure 11. The proposed landslide prediction flow chart.

3.2.3.1. Random Forest

Random forest (RF) was first introduced by Breiman [82] as an ensemble learning algorithm. It is a commonly used machine learning technique for data prediction for both classification and regression problems [83–85]. It is called random forest because of the randomness in the selection of features on each decision tree and also in the training samples from the dataset [86]. It operates by building up several decision trees from random subsets of the dataset during the training time to make up a relationship between features and labels [87]. A decision or classification tree (Figure 12) is a tree-shaped diagram which is a basic structural block of RF used to decide a route of action. RF consists of the construction of classification rules (tree) from bootstrap samples of the dataset. The ultimate classification decision is made using the majority vote of various classification trees.

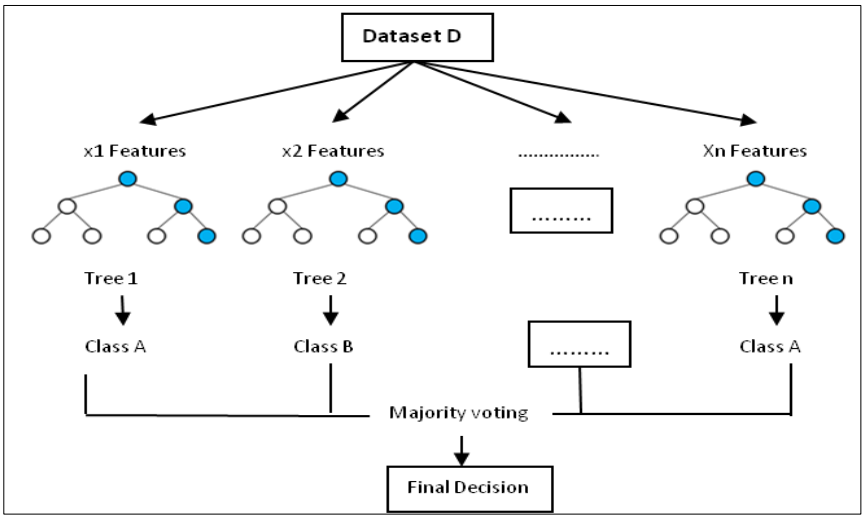


Figure 12. Random Forest decision tree.

The larger number of trees in the model plays a role in improving the accuracy and predictive capability [88] and requires defining the number of trees (T) and the number of predictive variables (m). RF can work with both categorical and numerical data [80]. RF was chosen because of its power to deal with mixed variables (categorical and numerical), over-fitting risks reduced, efficiently working on large databases with high accuracy, robustness with less training time, estimation of missing data, and feature importance indication [89].

3.2.3.2. Logistic Regression

Logistic regression (LR) is a machine learning algorithm that was initially used in the biological sciences (during the early 20th century) and later in other areas. It is used in solving classification problems by measuring the relationship between independent variables and categorical dependent variables. The technique has become one of the most used in landslide susceptibility assessment [80]. It is in three categories: Binary (dichotomous or has two possible classes coded as 0/1: absence/present), Multinomial (more than two categories without ordering), and ordinal LR (more than 2 categories with ordering). In its simplest form, the LR model can be expressed using a sigmoid function to makeover regression values in a range of $-\infty$ to $+\infty$ producing a probability between 0 and 1 for non-landslide and landslide respectively with 0.5 being the threshold [90]:

$$P(Y = 1) = \frac{1}{1 + e^{-y}} \quad (3)$$

where P is the probability of event occurrence and y is the dependent variable (landslide or no-landslide) and is defined as:

$$y = \beta 0 + \beta 1X1 + \beta X2 + \cdots \beta nXn \quad (4)$$

where y is dependent variable and X1, X2, … and Xn are explanatory variables. Then:

$$P(Y=1) = \frac{1}{1+e^{-(\beta 0+\beta 1X1+\beta 2X2+\cdots \beta nXn)}} = \frac{e^{\beta 0+\beta 1X1+\beta 2X2+\cdots \beta nXn}}{1+e^{\beta 0+\beta 1X1+\beta 2X2+\cdots \beta nXn}} \quad (5)$$

where $\beta 0$ is an intercept of the model and β1, β2,….βn are regression coefficients that measure the contribution of x1,x2, …xn to the prediction model. With LR, mathematical equations can be generated from regression coefficients and intercept, hence the probability of landslide incidence can be computed.

3.2.4. Re-Sampling

One of the challenges of various machine learning models is that their performance is affected by imbalanced data in the training dataset. To deal with this problem, three re-sampling techniques have been used: oversampling, under-sampling, and the synthetic minority oversampling technique (SMOTE), and the later was found to be the best. SMOTE was initially proposed in 2001 by Nitesh [90].

3.2.4 Synthetic Minority Oversampling Technique

SMOTE is an oversampling technique that uses the nearest neighbor's algorithm to create new instances of minority categories by forming a convex combination of neighboring instances. SMOTE works by picking examples that are near in the feature space, sketching a line between the examples in the feature space, and drawing a new sample at a point along that line. With this approach, "the minority category is oversampled by taking each minor category sample and introducing synthetic examples along the line segment joining any/all of the k minority class nearest neighbors" [90]. Synthetic data is generated in the following two main steps: (i) The difference between the feature vector (sample) under consideration is taken and its nearest neighbor. (ii) Multiplying the difference by a random number between 0 and 1, and add it to the feature under consideration. This makes the selection of an arbitrary point along the line segment between two specific features [90].

3.2.5. Preliminary Analysis Using Exploratory Data Analysis (EDA)

Exploratory data analysis (EDA) was used to get patterns and the relationship between landslide incidences and various parameters [90]. The scatter plot matrix tool was used as a method of EDA and helped to identify the relationship among variables.

3.2.6. Models Evaluation

Proper evaluation of the performance of the model is a crucial aspect of predictive modelling. A wide ranges of performance metrics are available for classification and regression models. A common metric for binary classification is the area under the receiver operating curve (ROC). The most widely used evaluation metric is the confusion metric, which distinguishes measures between errors and measures overall accuracy or percent correct classification. In this study, the

statistical index-based method and ROC have been used to evaluate the performance of the models. The following evaluation indices were calculated using the table of summary of classification errors (confusion matrix):

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (6)$$

$$FPR = \frac{FP}{FP + TN} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

where Accuracy is the overall performance of the classifier, false positive rate (FPR) is the rate of incorrect predictions. TP or Recall denotes the true positive (correctly predicted incidences), TN represents true negative (correctly predicted as negative) [91]. FP is the false positive (predicted occurrences but did not occur), FN is the false negative (predicted negative but in reality, it was positive). These performance metrics were selected based on their high impact on predicted landslides. Moreover, the metrics are most relevant for the proposed approach as the model will be combined with a wireless sensor network (WSN) that will be used as an early warning system to alert people about landslide occurrences and reduce risks.

The receiver operating curve is one of the most widely used models' performance tools for classification problems. It summarizes all the confusion matrices that each threshold produces by plotting FPR (1 - sensitivity) on the X-axis vs. Sensitivity or Recall (TPR) on the Y-axis [92]. The ROC curve makes it easy to identify the best threshold to make a decision. The area under the curve (AUC) is also used to compare two or more ROC curves. AUC is used to decide which classification method is better. The following equation is used to calculate AUC [93]:

$$AUC = \frac{(\sum TP + \sum TN)}{P + N} \quad (9)$$

where P is the total number of positives (landslides), N is the total number of negatives (No landslides).

3.3. Results

3.3.1. Preliminary Analysis: Correlation Among Features Used in the Models

As shown in Figure 13, there is no specific threshold for 1-day rainfall inducing landslides. Depending on how much rain fell in the previous days, a rainfall of 0 mm could trigger a lands lide. As indicated by the graphs, very few landslide hazards occurred for antecedent rainfall of less than 50 mm. In most cases, landslides occurred when previous cumulative precipitation was close to 100 mm.

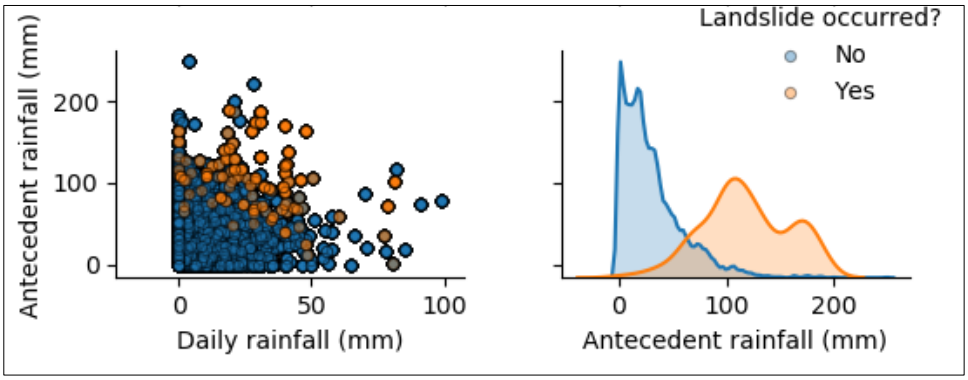


Figure 13. Correlation between landslide and rainfalls.

Also, as indicated in Figure 14, many landslides occurred in hilly areas where the slope inclination was 25–45° (categories 6 and 10) and very few landslides occurred in area where the slope was less than 15° (categories 1 and 4). The land covered by the forest resists landslide hazards, but also the low number of landslide incidences in that area can be justified by the low dominance of that category. Agricultural land is the most affected by slope failure, while the forest land is less affected.

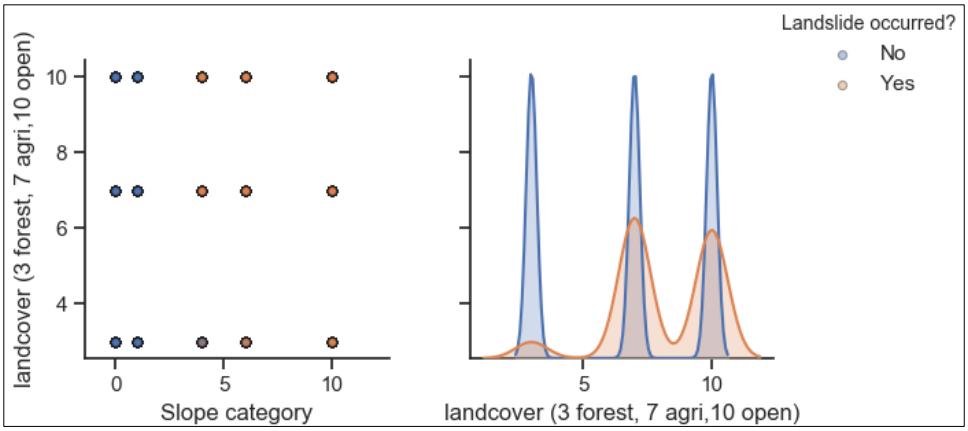


Figure 14. Correlation analysis between landslides, slope and land cover.

The soil under the category of silt was much affected by landslides, while areas of sand had few incidences, not just because of the physical properties of the soil particles but also because not much of this type of soil is found in the district. Figure 10 points out that few landslides' incidences occurred in the areas where the soil depth was thin (less than 50 cm), while more occurred in the regions with thick soil (depth greater than 50).

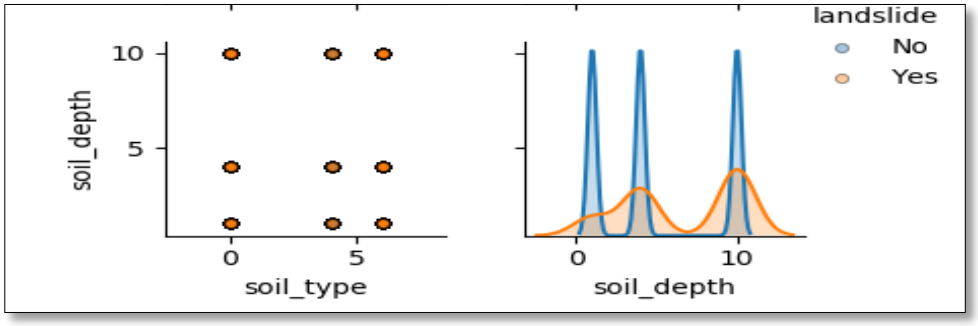


Figure 15. Correlation analysis between landslide and soil.

3.3.2. Models Results

The performance evaluation of the classifiers was done by using five different metrics, including Recall, accuracy, FPR, FNR, and the ROC-AUC. Four internal parameters and one external (triggering factor) have been used by the models. Initially, the models were trained by using 1-day rainfall, followed by training models with the inclusion of the cumulated 5-days antecedent precipitation as a new parameter into the dataset.

Initially, the models have been trained using 5-fold cross validation and later tested using different train/test ratios (0.8 and 0.2, 0.75 and 0.25, 0.7 and 0.3, 0.65 and 0.35, 0.6 and 0.4, 0.55 and 0.45) to determine which provides better results. By applying 5-fold cross validation, the variance is very low, as indicated by the insignificant standard deviation for both models, which is 0.0003 and 0.0004 for RF and LR, respectively (Table 3). The overall prediction accuracy was 95.30% (RF) and 94.63% (LR) by using cross validation, while 0.6 and 0.4 ratio provided better results with an accuracy of 95.35% for RF, and 0.55 and 0.45 was the best ratio for LR. On the other hand, if the antecedent rainfall is included as the new feature in the models, the overall accuracy was 98.74 and 98.79% (RF, LR respectively), the best ratio for RF cross validation was 0.65 and 0.35 with an accuracy of 97.67%, while 0.60 and 0.40 ratio looks to be better in LR model with 98.40% accuracy. It has been observed that there is no significant difference in applying different train/test ratios as indicated by the standard deviation (0.0001). However, the ratio of 65% of the dataset was adopted as training, while 35% was used as a test sample. This ratio was preferred because LR 0.65 and 0.35 is the best ratio for incorrect predictions (FNR) and for maintaining the same ratios for both models.

Table 3. Models' overall accuracy.

1-Day Rainfall (Without Antecedent Rainfall)						5-Days Antecedent Rainfall Included				
Fold	Accuracy (cross-validation)		Accuracy (train/test ratios)			Accuracy (cross-validation)		Accuracy (train/test ratios)		
	RF	LR	Ratio	RF	LR	RF	LR	Ratio	RF	LR
1	0.9530	0.9466	0.80, 0.20	0.9414	0.9322	0.9875	0.9877	0.80, 0.20	0.9742	0.9836
2	0.9536	0.9466	0.75, 0.25	0.9398	0.9330	0.9871	0.9882	0.75, 0.25	0.9759	0.9835
3	0.9531	0.9459	0.70. 0.30	0.9399	0.9352	0.9876	0.9878	0.70, 0.30	0.9744	0.9838
4	0.952	0.9457	0.65, 0.35	0.9394	0.9392	0.9870	0.9879	0.65, 0.35	0.9767	0.9838
5	0.9530	0.9467	0.60, 0.40	0.9469	0.9413	0.9876	0.9879	0.60, 0.40	0.9740	0.9840
Av.	0.9530	0.9463	0.55, 0.45	0.9409	0.9415	0.9874	0.9879	0.55, 0.45	0.9744	0.9837
Std	0.0003	0.0004		0.0028	0.0041	0.0002	0.0001		0.001	0.0001

The overall accuracy could not be considered as the only metric to judge the performance of the models because of the imbalance of dependent values in the dataset. Therefore, to assess the performance of the models, other metric indexes have been used. FN and FP were used to evaluate the models in terms of incorrect predictions such as landslide cases which have been considered as no landslides, yet landslides occurred. In addition to the prediction results presented in Table 4 RF generated features of importance indicating how much each conditioning factor contributes to the landslide incidence. As the triggering factor, rainfall has the highest contribution rate on landslide occurrence. If 1-day rainfall data are used (without antecedent precipitation), rainfall contributes 40.49%, slope 32.74%, soil type 16.48%, land cover (land use) 8.46%, and soil depth 1.80% (Figure 16a).

Table 4. Performance results.

Performance Metric	1-Day Rainfall, Antecedent Rainfall Excluded (%)	5-Days Antecedent Rainfall Included (%)

	RF	LR	RF	LR
Recall (TPR)	84.61	90.38	95.19	96.15
Specificity (TNR)	93.91	93.92	97.64	98.38
False Positive Rate (FPR)	6.08	6.07	2.35	1.61
False Negative Rate (FNR)	15.38	9.61	4.80	3.84

Upon including antecedent cumulated rainfall, the most contributing factor was antecedent rainfall at the rate of 40.64%, followed by the rainfall on the day of landslide occurrence contributed 22.91%, slope with 22.50%, soil type 7.87%, land cover 5.21%, and soil depth 0.84% (Figure 16b).

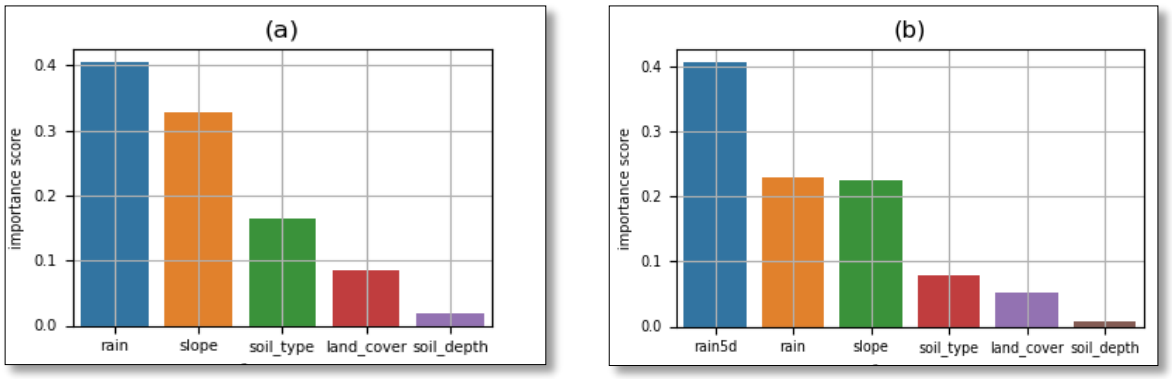


Figure 16. RF features importance (a) without antecedent rainfall (b) with antecedent rainfall.

Also, LR model provided the intercept and coefficients of variables from which the mathematical equation establishing the relationship between labels (dependent variables) and features (independent variables) was derived. Variables' coefficients reveal which parameters have a high or low impact on landslide occurrence as indicated by positive or negative values in Table 5.

Table 5. LR intercept and coefficients of parameters.

Intercept	Daily Rainfall	Antecedent Rainfall (5-Days)	Slope	Soil Type	Soil Depth	Land Cover
−7.06	1.32	2.47	−9.34	−2.90	−4.10	−4.37
			−9.29	−5.39	−1.85	−1.12
			1.20	1.23	−1.10	−1.56
			3.87			

6.49

By using equations (3) and (4) and the values of intercept and coefficients in Table 5, the probability of landslide occurrence can be expressed by the following equation:

$$P(Y=1) = \frac{1}{1+e^{-y}} = \frac{e^{y}}{1+e^{y}} \tag{10}$$

where P is the probability of landslide occurrence and:

$$\begin{aligned} Y = {} & -7.06 + 1.32rf + 2.47Ar - 9.34Sl_{0-10} - 9.29Sl_{11-15} + 1.2Sl_{16-20} + 3.87Sl_{21-25} \\ & + 6.49Sl_{26-45} - 2.9St_{clay} - 5.39St_{sand} + 1.23St_{silt} - 4.10Sd_{<0.5} \\ & - 1.85Sd_{0.6-1.0} - 1.10Sd_{>1.0} - 4.37Lc_{forest} - 1.12Lc_{agri} \\ & - 1.56Lc_{open} \end{aligned} \tag{11}$$

where −7.09 is the intercept, Rf is one day rainfall, Ar is antecedent rainfall, Sl is a slope in degrees (0–10, 11–15, 16–20, 21–25, 26–45) respectively, St is the type of soil (clay, sand, and silt), Sd is soil with (>0.5 meter, 0.5–1.0 meter, more than 1.0 meter) depth, and Lc is land cover (forest, agriculture, and open land).

3.4. Discussion

Prediction of landslide occurrence and early warning systems are the primary keys to awareness and preparedness for risk reduction from this type of disaster. Different prediction models and warning systems (both local and regional) are available for this purpose [38]. The most popular machine learning models that have recently been used for landslide prediction are namely random forest [85,94,95], artificial neural network [23,96], support vector machine [8,85], logistic regression [23,95,97], etc. The results obtained from these different studies were good depending on the variables used and the metrics used to evaluate their performance. In some studies, only internal factors were used to determine landslide susceptibility. Some others used rainfall, while a few of them included antecedent precipitation to determine its impact on the disaster occurrence. We chose RF and LR for this study for the reasons stated earlier, and more specifically to be used in Rwanda because, to our knowledge, no similar study has been conducted on this territory.

The use of both internal (geological and morphological) factors together with external (triggering) factors (rainfall: antecedent & current) led us to the better performance of the models compared to when either one of these factors is not considered. The results from a comparative analysis of one-day rainfall and 5-days antecedent cumulative rainfall before landslide occurrence revealed that antecedent cumulative rainfall has more impact on landslide occurrence than one-day rainfall (Table 6). This is because the shear strength of the soil is degraded by rainfall and continuous if the interval between two or more rainfall events is short (a few hours) but also depends on the hydraulic conductivity of the soil [98].

Table 6. Models' performance summary (various metrics.

Performance Metric	Random Forest			Logistic Regression		
	Correct Predictions (%)					
	Without Antecedent Rainfall	With Antecedent Rainfall	Improvement (%)	Without Antecedent Rainfall	With Antecedent Rainfall	Improvement (%)
Recall (TPR)	84.61	95.19	10.58	90.38	96.15	5.77
Specificity (TNR)	93.91	97.64	3.73	93.92	98.38	4.46
	Incorrect Predictions (%)					
False Positives	6.08	2.35	3.73	6.07	1.61	4.46
False Negatives	15.38	4.80	10.58	9.61	3.84	5.77

All performance metrics improve when the antecedent rainfall (Ar) is used for both models. For correct predictions, RF improves 10.58 and 3.73% for recall and specificity respectively, while LR improves 5.77 and 4.46% on the same metrics. Incorrect predictions improve in the same manner as the correct one since one is the opposite of the other. FN is a very important parameter to consider in the case of landslide prediction, as its low value means few incidence cases were not predicted. In this case, LR performs better than RF as Table 6 shows that only 3.84% cases were not predicted.

The receiver operating curve was used to identify the performance of the models at different classification thresholds and the AUC was used to compare two models by including or not including antecedent cumulated rainfall. Considering ROC-AUC, RF performs with an AUC = 0.973 (Figure 17a) if only daily rainfall is considered as a landslide triggering factor and with an AUC = 0.995 by including antecedent rainfall as an additional landslide triggering factor (Figure 17b). This means that antecedent cumulated rainfall implicates better performance of the model to the rate of 2.2%. Likewise, Ar improved the LR model as before the inclusion of previous cumulative precipitations, AUC was 0.979 (Figure 17a). However, after including the Ar, the AUC was equal to 0.997 (Figure 17b), indicating that this parameter contributed 1.8% to the model prediction.

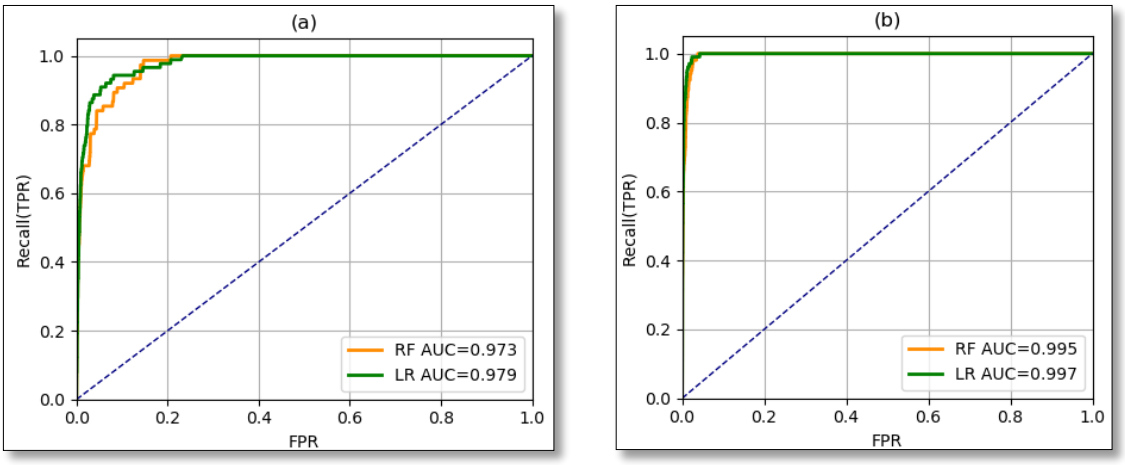


Figure 17. ROC-AUC (a) with no antecedent rainfall, (b) 5-days antecedent rainfall taken into consideration.

Both models indicated that five 5-days antecedent precipitation has a high impact on the occurrence of landslides in the study area. Table 5 shows that the features most triggering landslide incidences are rainfall (one-day, 5-days antecedent) as indicated by their coefficients of 2.47 and 1.32, respectively. Among the internal factors, the slope is the one most affecting the occurrence of the disaster. The magnitude of the impact increases as the slope increases or vice versa. For low slopes less than 15 degrees coefficients are negative, while the probability of landslide incidence is very high on slopes of 25–45 degrees, with a coefficient of 6.49. This means that high slope zones are prone to landslides, and people should not reside in such areas to reduce risks. The silt soil, the soil more than 1m deep, and agricultural land are the most vulnerable to the landslide (Table 5).

Table 6 shows that both models performed well in terms of different performance indicators such as accuracy, errors, and AUC. The use of antecedent precipitation made the models predict better. We also compared the performance of these prediction models with other models in recent studies using the same machine learning algorithms [99–105] and those using different models such as artificial neural networks (ANNs) [86], support vector machine (SVM) [100,102,103,105], Boosted Regression Tree (BRT) [100], Multivariate Adaptive Regression Spline (MARS), Quadratic Discriminant Analysis (QDA), Linear Discriminant Analysis (LDA), and Naive Bayes (NB) [102]. The comparison of the models was done using AUC, which seems to be widely used by researchers, and FNR, which has been used in a few studies [106–108].

The RF proposed by Yacine et al. [100] had almost the same performance as one of the RF and LR models proposed in this study before the use of antecedent rainfall data, but after including the antecedent rainfall data in the dataset, both RF and LR models predict better with good AUC results compared to other models. Thus, adding antecedent rainfall data has made the RF and LR models proposed in the current study achieve good results (Figure 18).

A comparative analysis has also been done with other models on incorrect predictions. Figure 19 indicates that the model proposed by Utomo et al. [106] is the best to minimize false negative errors up to 1.60%, while the LR in our study can minimize FNR up to 3.84%. On the other hand, our proposed LR can minimize FP errors up to 1.61% against 2.01% of Utomo et al. [106].

When comparing the two proposed models RF and LR of the current study among themselves, there is a slight difference in the prediction parameters. For instance, if ROC-AUC is considered, both models are good, as indicated in Figure 17, but in terms of prediction errors, LR is considered the best because its FNR is 3.84% against 4.80% of RF. This evaluation parameter (FNR) should be highly considered because it indicates how many landslide cases were predicted as NO landslides, which can be a dangerous outcome. Minimizing the ratio of false negatives is crucial for better performance of the system and disaster risk reduction. In this regard, logistic regression is the best model to be used due to its performance in terms of error reduction (false negatives). Besides error reduction, LR performs faster than RF, which is another important aspect that LR should be considered in the early warning system as the delay is reduced. Though no model can guarantee 100 percent accuracy, the objective is to reduce false negatives (warnings that there will be no landslides, but landslides do occur). This is because they have a greater detrimental impact on lives than false alarms (warning that there will be a landslide while it is false).

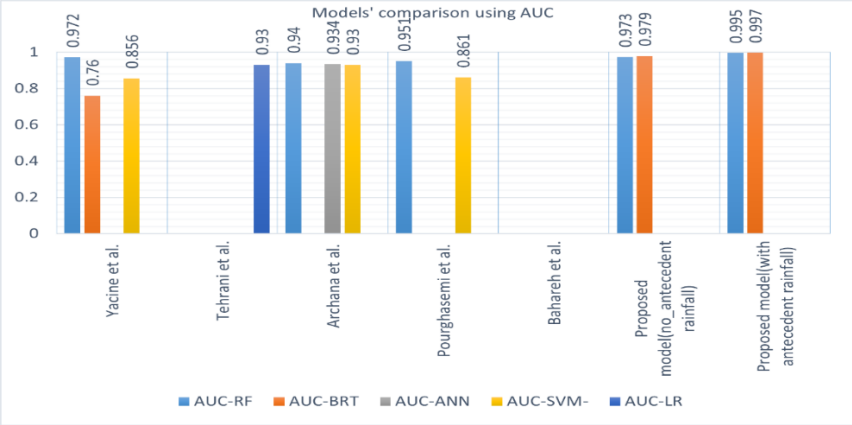


Figure 18. Comparison of the models results (using AUC) vs. other studies.

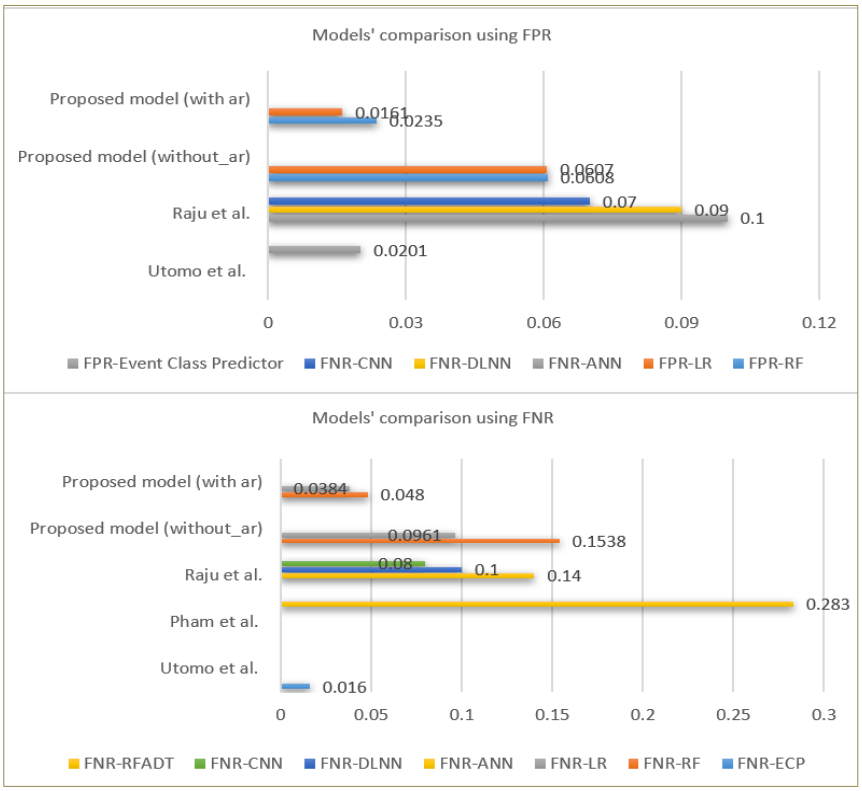


Figure 19. Comparison of the models results (false predictions) vs. other studies.

3.5. Conclusions

In this research, two approaches, random forest (RF) and logistic regression (LR), were applied to analyze the rainfall data along with other external and internal factors to develop a prediction model for landslide incidences for an early warning system. Performance parameters such as ROC-AUC, error rate (TP and FN) have been used to evaluate the best prediction models. The results prove that the prediction performance of these two models is better than those established in other research studies.

Results from this study revealed that landslides are triggered due to too much daily (or low intensity but prolonged) rainfall, but in most cases, they occur after a few consecutive (like 2–5) days of precipitation. This was proved by the impact of antecedent rainfall on disaster occurrences as marked by prediction models used in this study. Both models indicated that five days antecedent precipitation has a high impact on the occurrence of landslides in the study area. In addition to the rainfall data, other parameters have been utilized to assess their impact on the disaster. The slope of hills is the most internal parameter affecting disaster occurrence after rainfall (external factor), meaning that the areas of high slope angle are more susceptible to landslides than the regions where the terrain is almost a plateau. Agricultural land or non-protected land is the most susceptible to landslide occurrences while land covered by forest has few incidences. As previously discussed, landslides are triggered by current rainfall but very

highly correlated with consecutive precipitation before the day of incidence. This is because the soil strength reduces with water content depending on the type of soil material and gets strong again during the evapotranspiration process or the sunny season. Based on the results, logistic regression proves to be the best approach to be used for landslide prediction and early warning. The LR model's incorrect prediction rate FNR is 9.61% without including antecedent precipitation data and is 3.84% after including the antecedent precipitation data. Therefore, the LR model can be used for the early warning system.

Instead of antecedent cumulated rainfall, in our future work, we will study the correlation between landslide occurrences, rainfall and soil moisture level on different types of the soil, which will be gathered by sensors and through WSN, landslides will be predicted and citizens warned earlier. The prediction capability combined with the Internet of Things (IoT) where rainfall gauges can be used to capture rainfall intensity in real-time and soil moisture sensors used to measure the water content in the soil will be used for rescuing citizens from landslide disasters.

CHAPTER 4

EXPERIMENTAL STUDY OF SITE SPECIFIC SOIL WATER CONTENT AND RAINFALL INDUCING SHALLOW LANDSLIDES

Shallow landslides are among the natural threats causing death and damage. They are mostly triggered by rainfall in mountainous areas where precipitation used to be abundant. The amount of rainfall inducing this natural threat differs from one site to another based on the geographical characteristics of that area. In addition to the rainfall depth, the determination of soil water content in a specific zone has a major contribution to the landslide prediction and early warning systems. Rwanda, being a country with hilly terrain, and some areas are susceptible to both rainfall and soil water content inducing landslides. But an analytical study of the physical threshold determination of both rainfall and soil water content inducing landslides is lacking. Therefore, this experimental study was conducted to determine the rainfall and soil water content thresholds that can be fed into the landslide early warning system (LEWS) for alert messages using Internet of Things (IoT) technology. Various experiments have been conducted for the real-time monitoring of slope failure using a toolset composed of a rain gauge, soil moisture sensors, and a rainfall simulating tool. The results obtained show that the threshold for landslide occurrence does not solely correlate with the total rainfall amount (or intensity) or soil moisture, but it is also influenced by internal (geological, morphological) and environmental factors. Among the sampled sites, the sites covered by forest indicated no sign of slope failure, whereas sites with crops could slip. The experiments revealed that for a specific site, the minimum duration to induce slope failure was 8 hours and 41 minutes with a rainfall intensity of 8 mm/hour and the soil moisture was above 90% for deeper sensors. These values are used as thresholds for LEWS for that specific site to improve predictions.

4.1 Introduction

Rainfall induced shallow landslides are among the most common natural disasters that cause deaths and substantial economic losses by damaging infrastructure or plants in different mountainous regions around the world [67,109,110]. In Rwanda, about 1,000 landslide cases have been identified during the past decade [8], affecting a significant number of citizens, agriculture land, livelihoods, and infrastructure that are valued in billions of dollars. For instance, almost 200 people died by landslide incidences during 2016-2018 [111,112]. The most susceptible areas in Rwanda are the northern and western provinces, which are characterized by mountains and steep slopes [12,66].

Rainfall induced landslides are mainly caused by intrinsic factors like geological and geomorphological parameters and extrinsic factors such as hydrological conditions, climatic conditions, earthquakes, and volcanic eruptions [113–115]. Hydrological factors such as rainfall

and ground water table location influence the slope stability. The process of this type of natural disaster is intricate. The water content in the shallow soil is almost absent during the dry season. When it rains, the rainwater starts penetrating into the earth, characterized by different permeability properties, and then the ground becomes moist. It keeps on penetrating until it reaches the layer of low hydraulic conductivity, where it accumulates up to the complete saturation level [116] and forces the grains of the soil to separate. The matric suction on the topsoil decreases, and once the ground is fully water-logged, the soil matric suction finally vanishes completely [117], making the ground unstable, and that is a critical condition for landslide risk [118]. Thus, depending on the other geo-physical characteristics of the area, a landslide incidence may follow [118]. Chiorean in [119] describes the process of a rainfall-induced landslide as follows: (i) rainfall permeation results in a reduction of the matric suction of the slope soil, (ii) the diminution in soil matric suction decreases the soil shear strength, and (iii) the decrease in soil shear strength afterwards causes the slope to become unbalanced and finally fail. The matric suction (ψ_m) plays a crucial role in landslide occurrence and is defined as the difference between pore air pressure (u_a), and pore water pressure (u_w) [119–122].

$$\psi_m = u_a - u_w \tag{12}$$

ψ_m is influenced by the movement of water (most of the time rainwater) within soil pores [123]. At the initial state, when the soil is free from water (dry), the soil has the peak value of matric suction, and this attains its lowest value (zero) when the soil is fully saturated [124]. It is associated with the degree of soil saturation over the soil water characteristics curve (SWCC) [125]. The SWCC is broadly defined as a variation of water within a soil with respect to suction and is graphically represented by matric suction on the x-axis versus degree of saturation (*S*), gravimetric water content (*w)*, or volumetric water content (θ) on the y-axis [126]. The degree of soil saturation is expressed in the following mathematical equation [126]:

$$S_r = \frac{V_w}{V_v} \tag{13}$$

where Sr is the degree of saturation, V_w is the volume of water, and V_v is the volume of voids in a soil representative elementary volume. Gravimetric water content (w), which is the commonly used term in geotechnical engineering, is calculated using equation (3):

$$w = \frac{M_w}{M_s} \tag{14}$$

where *w* is the gravimetric water content, M_w is the mass of water, and M_s is the mass of soil solids. The volumetric water content (θ), which is mostly used in agriculture field, is the amount of water content in the soil over the total volume of the soil. Its formula is expressed as follows:

$$\theta = \frac{V_w}{V_v + V_s} \quad (15)$$

where θ is the volumetric water content, V_w is the volume of water, V_v is the volume of voids, and V_s is the volume of solids. Landslides are closely linked to the rainfall infiltration that causes dissipation of matric suction [127]. Therefore, the SWCC is essential for modelling rainfall infiltration in the soil and is useful for landslide prediction.

Various mitigation methods have been extensively utilized to reduce the risk of landslide incidence. They include both structured and non-structured warning systems. Structural methods require high construction costs and may take many years to complete [128]. Therefore, high precision LEWS is more efficient and should be emphasized.

Numerous studies have been undertaken broadly to reduce the impact of landslides on human lives and economic loss. They include various prediction models, susceptibility models, and landslide early warning systems (LEWS) [110,114]. Some of the models establish the relationship between landslide occurrence and rainfall intensity through laboratory field tests as well as numerical analysis [129,130]. Others estimate the rainfall intensity and duration that could cause landslide incidences [114,131–135]. The parameters used in each study are different and depend on the author's selection and the availability of data. The most common parameters used in different studies are rainfall (an external triggering factor) and internal factors such as slope (inclination, aspect), soil type, lithology, and land cover [107,136–142].

Past studies have used the laboratory flume test to identify the correlation between soil water content and landslide occurrence where rainfall, soil type (particle size), and slope inclination have been taken into consideration [120,124,143–146]. Others used empirical field tests [146,147] or numerical modelling [110,148], and various studies conducted on LEWS have generally used historical rainfall data together with landslide occurrence records to determine the rainfall threshold that induces landslides [140,149–155]. The studies conducted in [129,142,156–162] have shown that they have used antecedent rainfall as a key parameter that has a great impact on landslides' occurrence. In this research paper, the parameters considered are rainfall intensity together with soil moisture content (SMC), which can help to achieve a better landslide prediction modelling tool that might be used for LEWS. Though, the preceding studies demonstrated the role that rainfall plays in triggering slope failure and came up with good results, but the optimization is needed because of the complexity of conditions causing landslide incidence. Therefore, there is a need to derive the correlation between the present rainfall, antecedent rainfall, and the content of water in the soil at the time of slope failure to determine the site specific threshold values of these parameters.

In relation to the above, the gaps found in the past studies are the exploration of how rainfall has different effects on slope stability according to geophysical and environmental characteristics. For example, some experiments used the embankment soil in the slope flume [130,163,164] that has no original strength (cohesion) and could not provide a reliable threshold like the one conducted on the terrain. Therefore, the field experiment to determine the SMC and rainfall depth is of great importance for the LEWS. The present study was carried out on site and aimed at: 1) analyzing how gradually the rainwater penetrates into the soil during or after rainfall events until the landslide occurrence; 2) identifying the rainfall amount and threshold values for the SMC as a step prior to feeding in to the LEWS using IoT technologies. Both quantitative and qualitative techniques were used on different environmental covariates, namely, slope, soil type, vegetation (or land coverage), and rainfall intensity as an external parameter. The soil water content inducing landslides will be determined for each selected site. The geographical scope of the study is the district of Gakenke while the period of study is October 2019 to June 2021 (21 months).

4.2 Materials and Methods

4.2.1 Study Area

This study was conducted in five sectors (Figure 20c) of Gakenke district (Figure 20b), located in the northern province of Rwanda (Figure 20a). The district shares borders with Rulindo, Burera, Musanze, Nyabihu, Kamonyi and Muhanga Districts. The district comprises 19 administrative sectors divided into 97 cells, and 617 villages. The District has an area of 704.06 km^2 [165].

. The population density is 473 residents/km^2. The climate in this district is commonly a type of humid climate, with the average annual temperature ranging between 16^{o}C and 29^{o}C. The rainfall is quite plentiful, with a scale between 1,100 and 1,500 mm/yr. This district has four main seasons: the small dry season from January to February, a high rain season that spans from March to May, marked by plentiful rainfall and landslide incidences, the long dry season extending from June to August, and finally the short rain season from September to December. The high hills separated by rivers and swamplands characterize this district. The highest altitude attains 2,647 meters (Mount Kabuye), whereas the lower altitude is 1,362 meters [165].

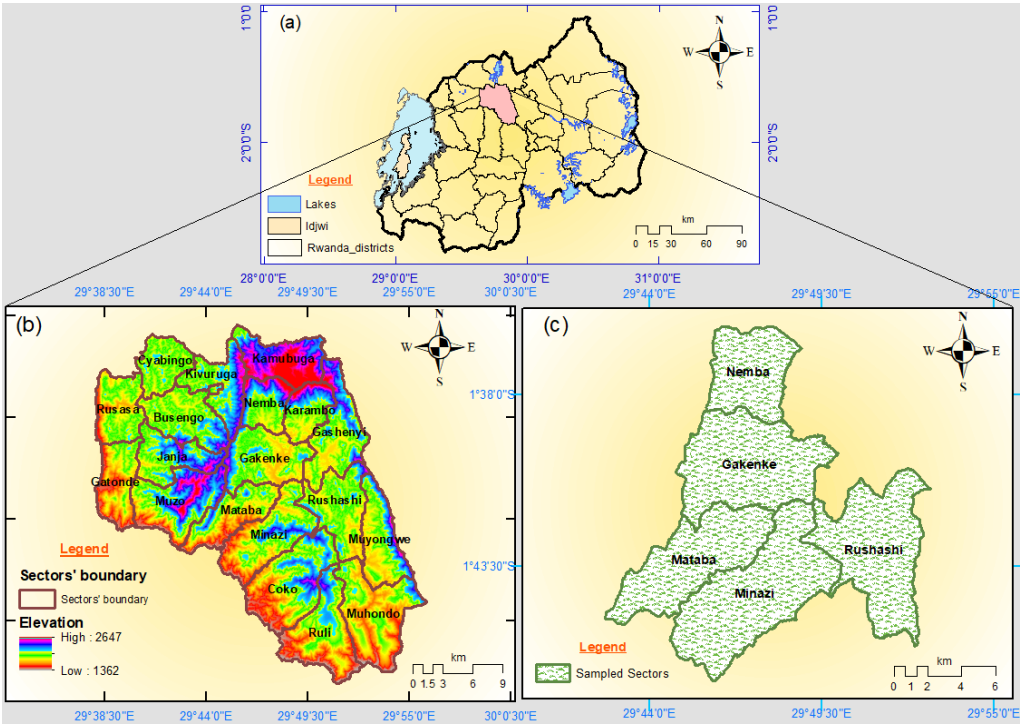


Figure 20. Study area: (a) Rwanda, (b) district elevation map and sectors' boundaries, (c) district's sectors of experiments.

Due to its climatology and geographical characteristics such as topography, geology, the district is characterized by numerous landslides that cause few to many deaths and property damage in different heavy rainfall events [166–168]. In addition to the frequent and abundant rainfall, high slopes, land cover, and soil texture contribute to the slope failure in this region [169,170]. The landslides in this region can be classified into two categories: (i) landslides related to natural slopes that initiate from anywhere on a hill slope (Figure 21a) and (ii) human-made slopes which are cut slopes related to house plots (Figure 21b), roads, or excavation activities.

Figure 21. Pictures of landslides in the study area during 2016 (a) and 2020 (b) events.

4.2.2 Methods and Data

The early warning system is one of the methods that can be used to reduce the risks related to landslides by providing incident information to the citizens prior to the occurrence. Two main

methodological approaches can be used for landslide mitigation techniques. The first method uses physically based models considering the infinite soil mechanism, the second counts on experimental studies to determine the rainfall intensity and duration threshold [114,116]. The challenge associated with the first approach is that the landslide can be detected but not predicted. The second approach can be appropriate, but the threshold for each influencing factor (such as hydrological, land use, lithological, and soil characteristics) should be determined [171]. In the current study, the rainfall intensity causing shallow landslides is estimated considering the soil-forming factors and environmental covariates. Various experiments and numerical analysis were carried out to estimate the rainfall amount and the soil water content level that may lead to slope failure.

4.2.2.1 Daily Rainfall Data

The slope failure in the study area and in the entire country is solely dependent on rainwater. Therefore, the rainfall data were necessary for this study. Primary and secondary data have been used for both quantitative studies and qualitative assessment. Firstly, historical rainfall and soil moisture data were collected from Rwanda Meteorology Agency for analysis of their correlation. Data from three rain gauge stations (Figure 23d) were used for primary analysis (Figure 22). This analysis was envisaged to identify rainfall amount induced slope failures and landslide locations in the neighborhood of rain gauge stations.

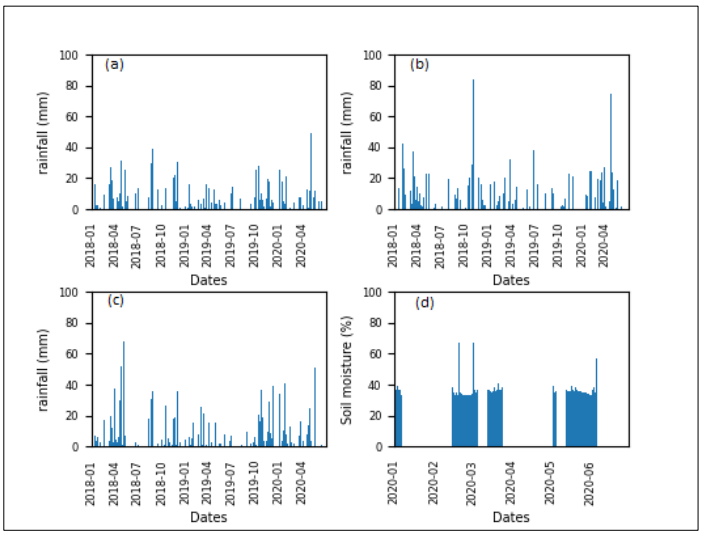


Figure 22. Past daily rainfall in the district of Gakenke from three rain gauge stations and one soil moisture station: (a)Janja, (b)Minazi, (c)Nemba, (d)Rushashi soil moisture station.

Secondly, field work was conducted in the surrounding areas that have been characterized by at least two landslide incidences in the past five years. In this study, rain gauge and soil moisture sensors were used for real-time data collection.

4.2.2.2 Soil Moisture

In addition to the rainfall data, historical soil water content data were collected from the office in charge of meteorology and were analyzed (Figure 22d) on the basis that the SMC has a correlation with precipitation [172–175]. Moreover, studies revealed that the actual soil wetness can be achieved through in situ measurement [113,176]. The soil moisture sensors were used to collect the soil water content on different sites (Figure 23d).

4.2.2.3 Slope

Various studies show that the slope has a high impact on landslide occurrence. Steep slopes in several regions characterize Gakenke district, where the slope angle can be more than 45 degrees. Other studies such as [142,169] used 5 slope classes, whereas in this study, we grouped the first two classes because landslide cases are very few in areas with slopes of less than 15%. According to the data source (GIS centre at the University of Rwanda), the slopes are categorized in four classes, as indicated by Figure 23c.

4.2.2.4 Soil Types

Geotechnical properties have an important impact on slope stability [177]. At each test site, soil samples were collected and taken to the soil mechanics lab at the University of Rwanda to be tested for soil classification and other analysis. The soil classes found on the sampled sites are: Silty Sand, Sandy Silt, Sandy Lean Clay with Gravel, Lean Clay, and Elastic Silt.

4.2.2.5 Land Cover

Land use is another factor contributing to the landslide occurrence. There are five different types of land cover in the study area: forest, cropland, grassland, built area, and water, but only the first three could be used in this study.

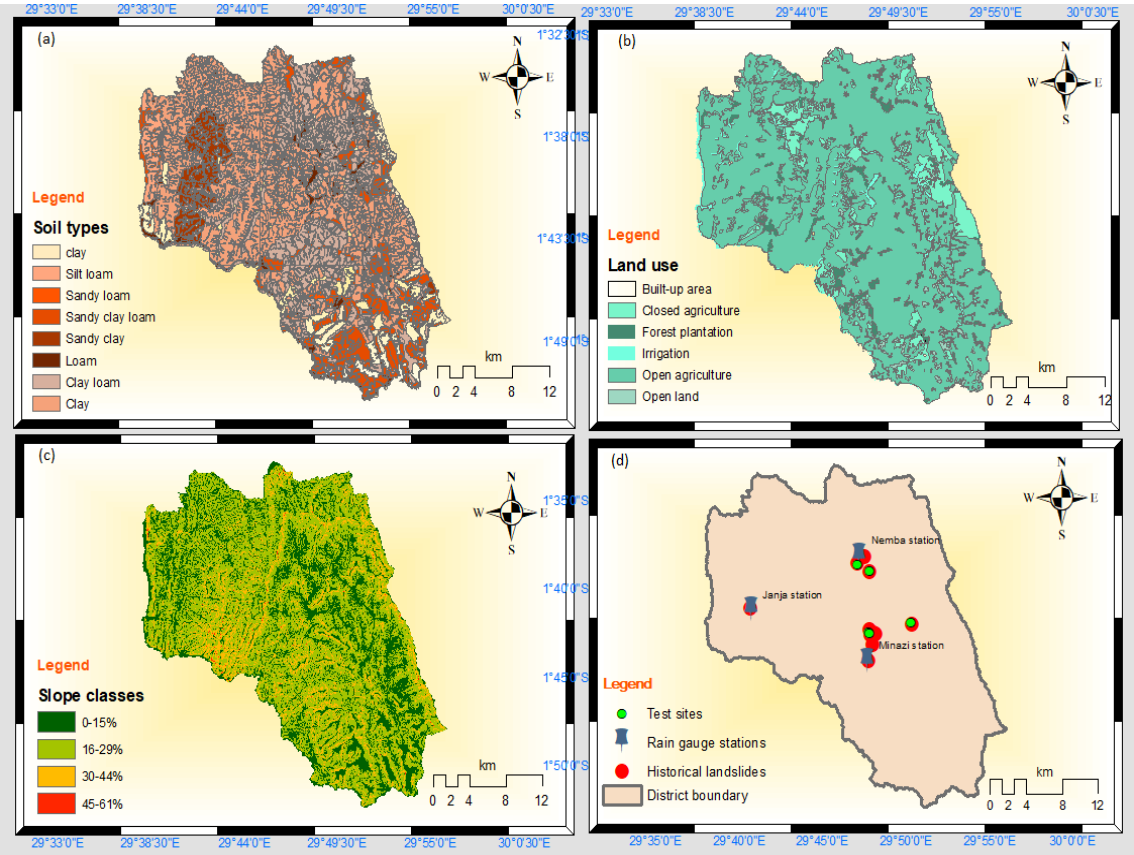


Figure 23. Maps: (a) soil types, (b) land use classification, (c) slope classification, (d) rain gauge & test site locations.

4.2.2.6 Experimenting Tools and Setup

To identify the rainfall influence on slope instability and its correlation with soil moisture, we used the rainfall simulator (Figure 24) for the direct in situ measurement of the soil humidity at various time durations until the slope starts indicating the sign of sliding, such as a horizontal crack on the ground above the slope, or sliding of cut slopes. The rainfall amount was recorded along with soil moisture at an interval of 49 seconds. The sensor-based monitoring tool was made of: (i) the sensor node comprising three analog capacitive soil moisture sensors manufactured by Paialu, whose operating voltage is 3.3 ~ 3.5V to capture soil water content in various ground depths; and the Arduino Uno Microcontroller ATmega328P. (ii)a weather station consisting of the transmitter (MISOL model: WH40) and the receiver (model: WN5360). The communication between the transmitter and receiver was via wireless, with a transmission frequency of 433 mhz and a maximum distance of 100 meters. (iii) A laptop with the Python code data logger to record readings from 3 sensors and convert into CSV file (Figure 25)

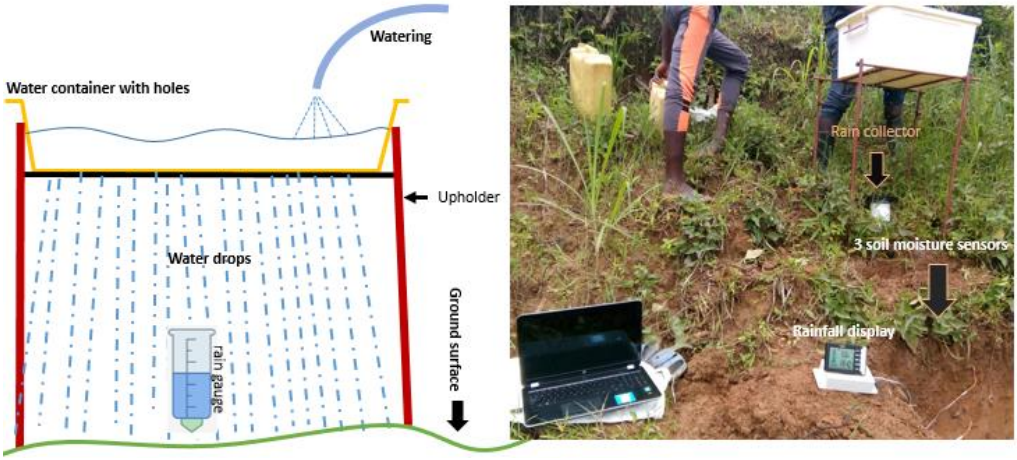


Figure 24. Rainfall simulator.

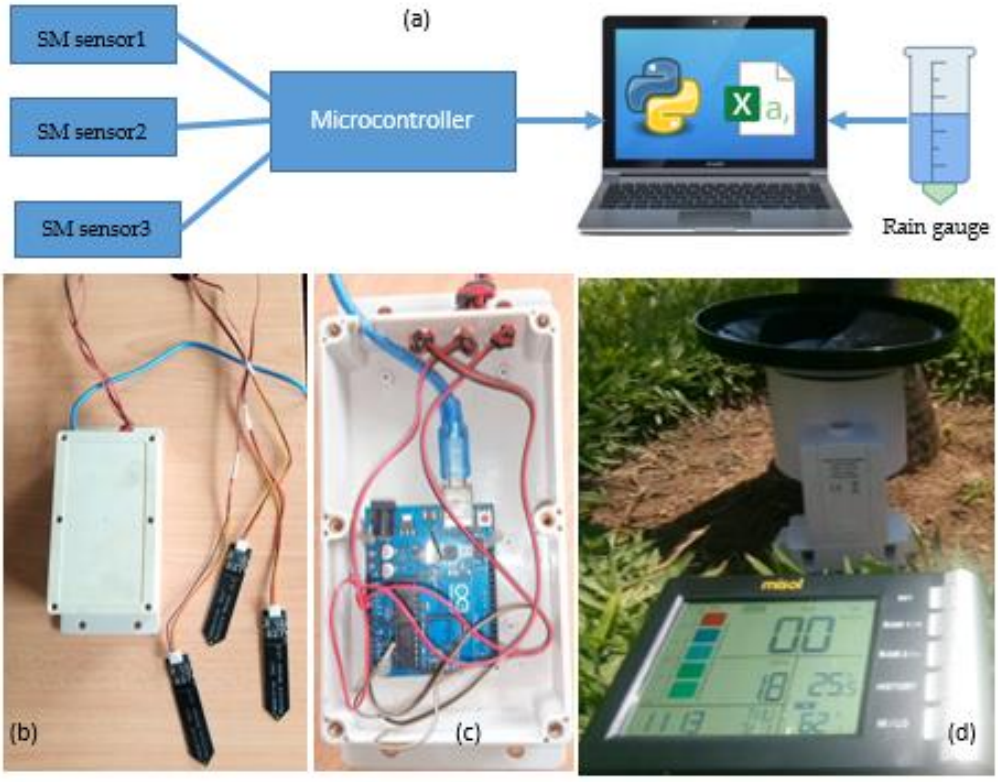


Figure 25. The block diagram & tool set of monitoring equipment: (a) block diagram, (b) sensor node box, (c) inside box, (d) rain gauge with a digital display.

The soil moisture sensor used is a resistive sensor that outputs the voltage variation as water penetrates the ground, i.e., the increase of water in the soil lowers the ground resistance that will then drop the voltage. By default, the sensor readings change from high to low when moisture is detected, and their values may differ from one sensor to another. Therefore, the output values from the sensors were calibrated in the Arduino IDE (Integrated Development Environment) by

putting the sensors into oven-dried soil and recording the readings, and then putting the sensors into the fully wetted container of the soil (fully saturated) and again recording readings. In the Arduino IDE, the maximum value (reading in dried soil) was mapped to zero (0) and the minimum (reading in wet soil) to 100. Three sensors were placed at different ground depths as indicated in Table 7.

Table 7. In-ground depth placement of sensors

Sensor	**Site plot**										
	Depth placement (m)										
	SA1	SA2	SB	SC1	SC2	SC3	SD1	SD2	SD3	SE1	SE2
Sensor1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Sensor2	0.6	0.6	0.5	0.6	0.7	0.6	0.6	0.6	0.6	0.5	0.6
Sensor3	1.2	1.1	0.9	1.2	1.3	1.2	1.1	1.2	1.2	1	1.2

4.2.2.7 Experimental Sites

All field experiments were conducted in the surrounding zones that have the historical background of landslides (Figure 23d). As shown in Table 8, several typical sites were chosen based on geo-topographical and environmental factors such as slope, soil types, and land use.

Table 8. Geo-topographical characteristics of representative sites in the study area

#	Site	Location (Sector, Cell, Village)	Plot	Slope (^{o})	Land cover	**Number of experiments**
1	SA	Minazi, Raba, Ndegamire	SA1	26.4	Crops	3
2			SA2	35.3	Grass	2
3	SB	Mataba, Buyange, Gabiro	SB	41.8	Forest	2
4	SC	Rushashi, Mbogo, Gisanze	SC1	47.2	Crops	3
5			SC2	28.9	Crops	3
6			SC3	39.7	Grass	3
7	SD	Gekenke, Rusagara, Museke	SD1	49.5	Forest	2
8			SD2	31.3	Crops	3

9			SD3	29.8	Grass	3
10	SE	Nemba,Gisozi,Karukara	SE1	28.6	Forest	2
11			SE2	24.4	Crops	3

At least two experimental tests were done at each site plot to test the reoccurrence of the slope failure. Following the first two tests, a third test was undertaken on those who identified at least one landslide to ensure that the results were reliable.

4.3. Results and Discussion

4.3.1 Soil Classification Test Results

Five main sites were chosen as sample sites (SA, SB, SC, SD, and SE). Soil samples were taken for the laboratory test of soil particle size distribution analysis because this plays an important role in different hydrologic features such as water retention characteristics and slope failure as well. The relative composition of the soil in the sampled sites is summarized in the table below (Table 9).

Table 9. Soil particle size and classification

Sample	Gravel (%)	Sand (%)	Fines (%)	LL (%)	PL (%)	PI (%)	Group symbol	Group name
SA	1	52	48	37	25	12	SM	Silty Sand & Sandy Silt
SB	24	24	53	41	23	18	CL	Sandy Lean Clay with Gravel (SLCG)
SC	0	10	90	42	25	17	CL	Lean Clay
SD	0	3	97	66	37	29	MH	Elastic Silt
SE	0	10	90	42	23	16	CL	Elastic Silt & Lean Clay

As shown in Table 8, from each site, two or three plots were sampled (except site SB) to identify the soil-related effects on the infiltration process. Some of the plot samples were found to have different soil textures. For instance, the soil types on site SA were Silty Sand and Sandy Silt for SA1 and SA2, respectively. Likewise, site SE plots had Elastic Silt and Lean Clay for SE1 and SE2, respectively.

4.3.2 Simulation Results

The rainfall simulation was applied to the selected sites while recording the rainfall amount and soil moisture content (using a wireless rain gauge and three sensors, respectively) until the slope sliding (or crack) is observed or not. The activity lasted a different amount of time depending on the site. The limit of rainfall simulation and duration was based on the rainfall events that induced landslides recently (Figure 22) or in the past [142]. Out of twenty-nine (29) experimental tests carried out on 11 plots, sixteen of them (55.5%) resulted in slope failure (or crack) as shown in Table 10.

Table 10. Observed simulation results

Site	Soil type	Intensity (mm/h)/duration (hours, minutes)			Slope failure observed (Test 1, 2, 3)?
		Test 1	Test 2	Test 3	
SA1	Silty Sand	8.7/10h15	8.7/12h31	8.7/11h45	Y, N, N
SA2	Sandy Silt	8.7/11h03	8.7/11h26	NA	N, N, -
SB	SLC/G*	8.4/11h33	8.4/12h24	NA	N, N, -
SC1	Lean Clay	9.4/10h59	7.8/10h27	7.8/10h23	N, Y, Y
SC2	Lean Clay	7.8/9h04	7.8/8h48	7.8/9h40	Y, Y, Y
SC3	Lean Clay	9.4/9h54	7.8/9h32	7.8/10h37	Y, Y, N
SD1	Elastic Silt	11.1/10h37	7.2/11h04	NA	N, N, -
SD2	Elastic Silt	8.0/8h41	7.2/11h10	7.2/11h21	Y, Y, N
SD3	Elastic Silt	7.2/9h36	7.2/11h58	7.2/9h37	Y, Y, Y
SE1	Elastic Silt	10.8/11h14	7.7/10h14	NA	N, N, -
SE2	Lean Clay	7.7/9h11	7.7/8h53	7.7/9h28	Y, Y, Y

As shown in Table 8, most sites are characterized by steep and very steep slopes, while the laboratory results indicate that lean clay and elastic silt are the most dominant soil classes in the sampled sites. The results in Table 10 show that landslides are possible in all categories of soil, except in sandy lean clay with gravel, which is less represented among the selected sites because the sampled sites are those that had landslides in the past.

4.3.3 Correlation between Rainfall and Soil Moisture Content

The shortest time used to simulate rainfall and provoke slope failure was 8h41 at a rainfall rate (intensity) of 8 mm/h. This occurred on site SD2, where the land is covered by crops at a slope of 31.3º. The soil moisture at the point of failure was 82%, 92%, and 95% (Figure 26a). In addition to other factors, the short duration of time to simulate slope failure is linked to the antecedent rainfall indicated by high initial soil moisture.

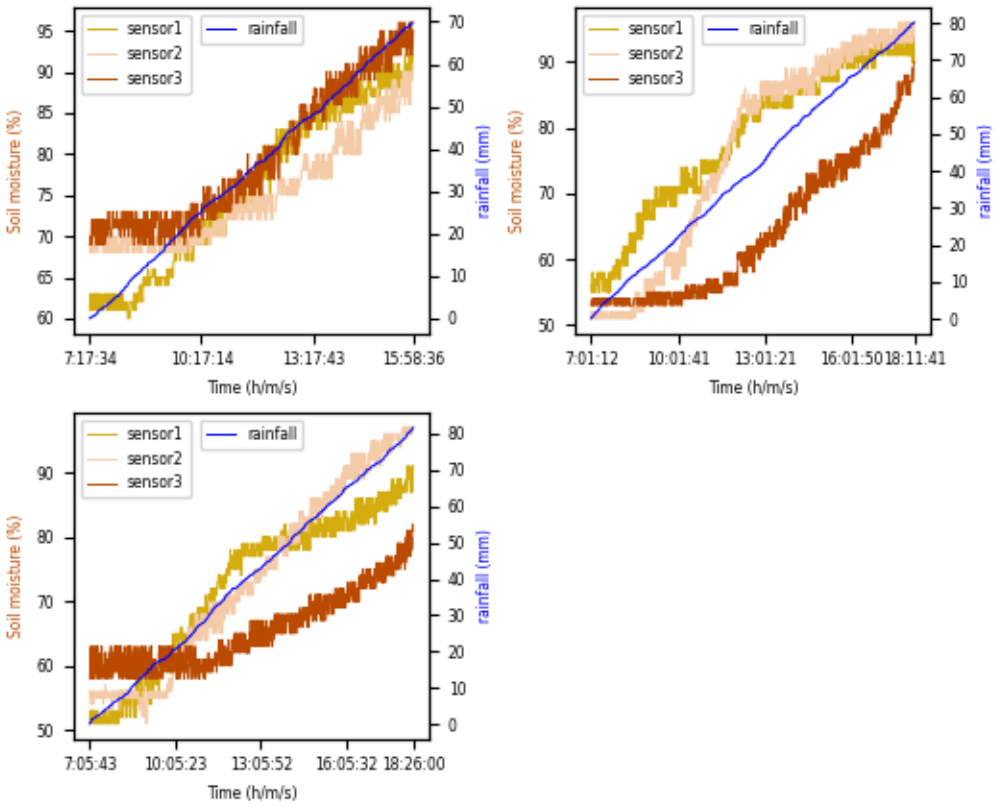


Figure 26. Variation of soil moisture content versus cumulative rainfall for three experimental tests on site SD2. The slope failure was observed for the first two tests (a, b) and did not occur during the last test (c).

The nearby site SD3 was selected with a different land cover (grass) and a slope inclination of 29.8º. The slope failure was observed for all three tests on this site after 9h36, 11h58, and 9h37 with a rainfall intensity of 7.2 mm/h. It took a different duration of rainfall simulation to induce a landslide, but the soil moisture content recorded by the three sensors was above 80% for the top sensor, while the deeper sensors recorded more than 90%, as shown in Figure 27 a-c. Furthermore, three tests conducted on site SC2 resulted in slope failure. This site has a slope inclination of 28.9º and the land used for agriculture (covered by crops). The maximum soil moisture content attained by the three sensors was 98%, while the least value was 86%, recorded by the top sensor (sensor1) during the first test (Figure 27). As shown in Table 10, the duration of each test was different due to factors like the initial soil moisture content (prior to the experiment) or other internal geological factors.

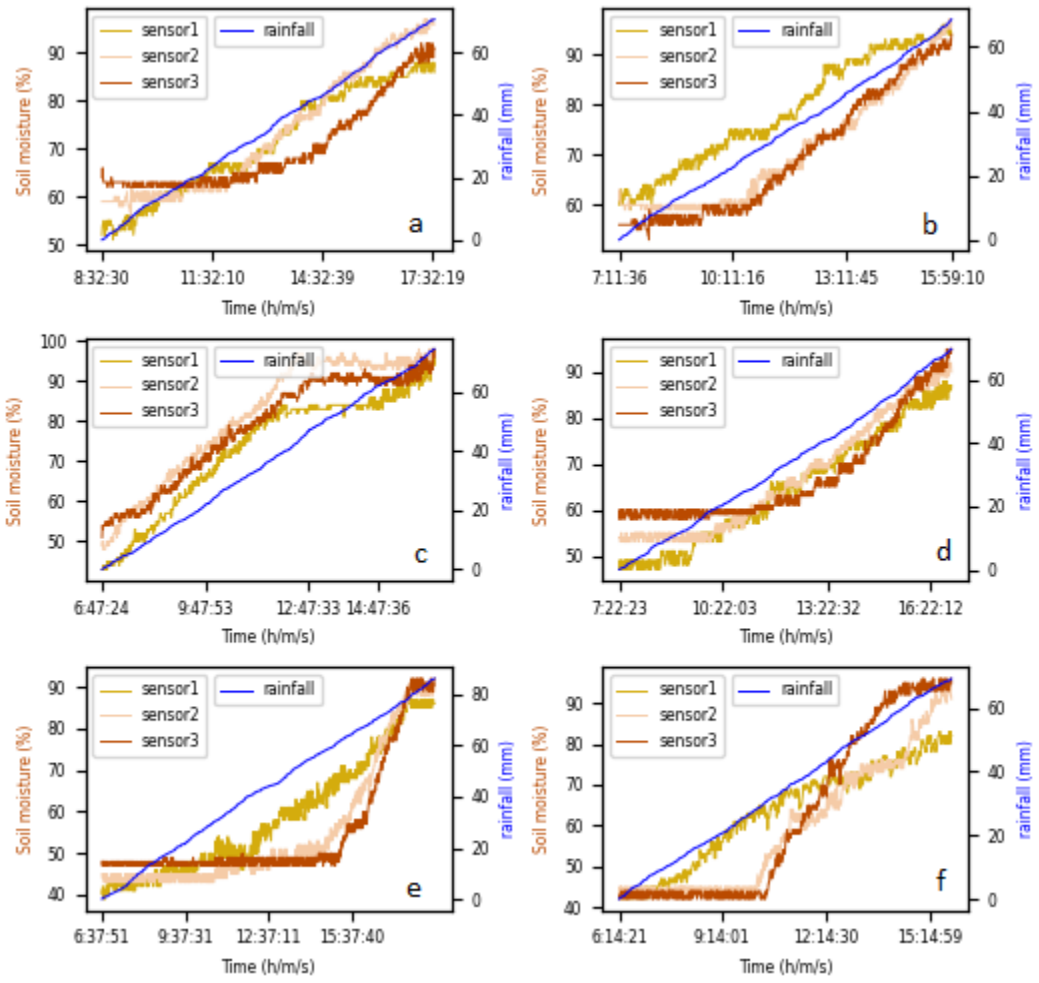


Figure 27. Variation of soil moisture content versus rainfall for three tests on site SC2 (a-c), and Site SD3 (d-f).

It was also observed that the two sites were characterized by lean clay. It indicates that this soil type is more susceptible to landslides because even on sites SC1 and SC3, two tests out of three resulted in slope failure. On site SC1, the rainfall simulation of 10h59 and intensity of 9.4 mm/h did not result in slope failure, while the last two tests at an intensity of 7.8 mm/h resulted in slope failure after 10h27 and 10h23, respectively. On the other hand, the first two tests on site SC3 resulted in slope failure; however, the last did not even show any sign of sliding. Figure 28 shows the correlation between rainfall and soil moisture content for the two sites.

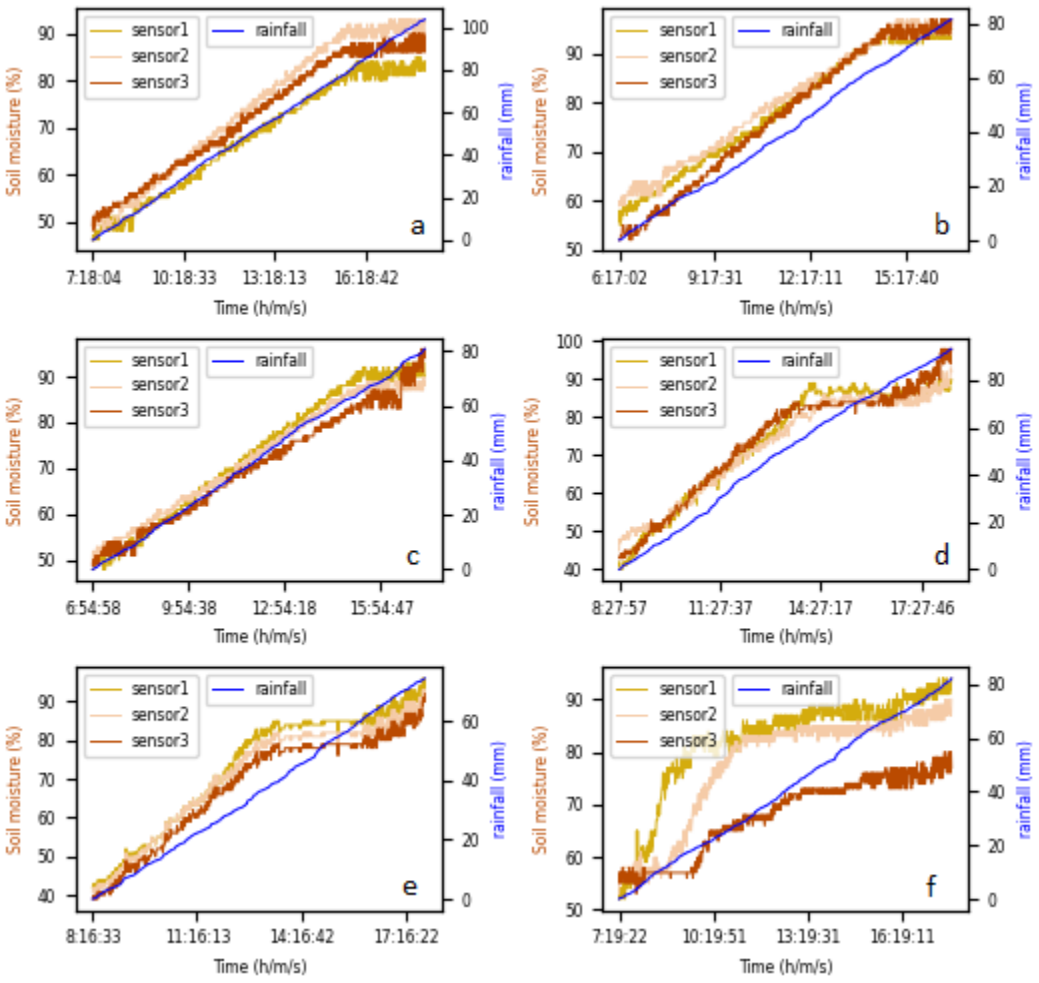


Figure 28. Variation of soil moisture content versus rainfall for three tests on site SC1 (a-c), and site SC3 (d-f).

The land covered by the Eragrostis spectabilis type of grass has shown to be resistant to rainwater infiltration in the soil. This is the only site among the grass-covered plots (Figure 29 a&b) that did not show any sign of slope failure for the first two tests. However, out of the three tests carried out on site SA1, only one slope failure was observed as shown in Table 10. Generally, the simulation took a long time on sites that did not show any sign of slope failure compared to those that manifested landslides. The reason was to test the effect of daily total cumulative rainfall and duration on landslide occurrence. Figure 29 (a-f) shows that the simulated cumulative rainfall was about 100 mm while the duration was more than 11 hours (Table 10).

The land cover is common to the sites that showed no signs of slope failure, as they are covered by forest (Figure 29 c-f), whilst the other is covered by grass (Figure 29 a-b). Figure 29 (a-f) also indicates that, even though it required a long duration of simulation and the high amount of total rainfall, in some cases, the saturation level was less than 90% for the three sensors (Figure

29 a, c-f)). This means that it requires a long time to make the soil fully saturated with the slopes covered by forest or protected by Eragrostis spectabilis.

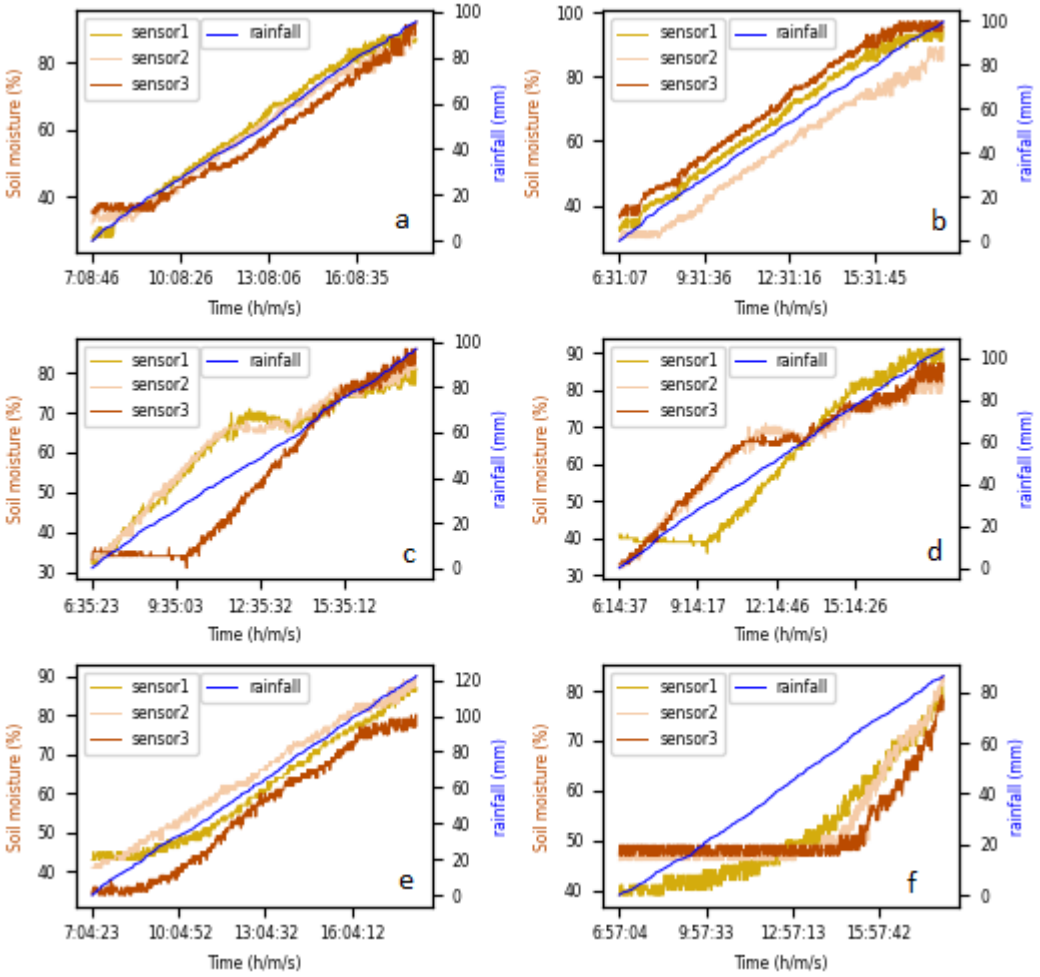


Figure 29. Variation of soil moisture content versus rainfall for sites SA2 (a, b), SB (c, d) and SE1 (e, f).

From the figures above, the linear regression of both rainfall and soil moisture (Figure 30) is observed due to the continuous simulation of rainfall. In practice, this is not the case because rainfall is characterized by discontinuous events (with inter-event periods) of different durations. Even though one-day rainfall can reach the values simulated in this study, the duration either expands to many hours or else can be related to the previous

rainfall events (antecedents). Hence, the use of soil moisture sensors for LEWS has a crucial importance because their data are informative to the antecedent rainfall.

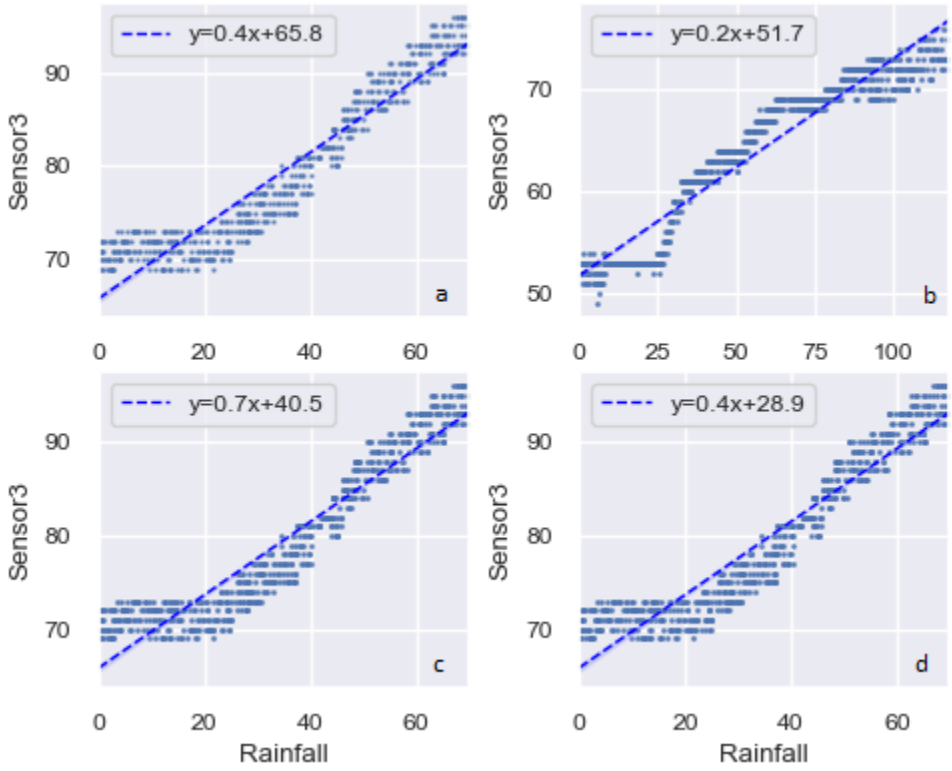


Figure 30. Rainfall vs. soil moisture: (a) SD2, (b) SD1, (c) SE2, (d) SE1.

The first two figures (Figure 30a & b) compare cropland and forestland, respectively. As shown by the slope lines of the best fit, the slope of 0.4 indicates that the infiltration rate was faster for crop land than that of forest land and that it got saturated before forest land (slope = 0.2). Likewise, Figure 30c & d, slope = 0.7 for cropland and 0.4 for forestland.

4.3.4 Slope Failure,Total Rainfall and Intensity

The numerical analysis of rainfall intensity and duration (Table 10) shows that each site required a different duration and rainfall intensity to initiate the slope failure. This shows that the thresholds for the two parameters are not identical, and in order for them to be identical, all other factors should be identical, which is almost impossible for different sites. As shown in Figure 31, there are many landslide cases that occurred when the rainfall intensity was lower than that of the highest intensity.

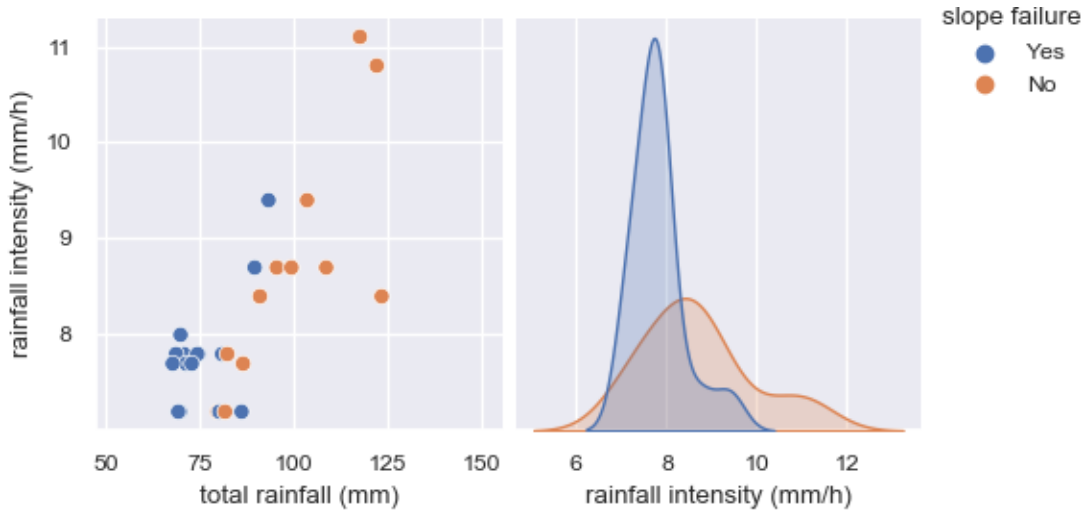


Figure 31. Slope failure vs. total rainfall and intensity.

4.3.5 Slope Failure and Geo-environmental Factors

It is not forthright to explain the correlation between hydrological and mechanical processes that occurred before and during the slope failure. Although the triggering factors may be clearly known, the process itself is complex. Rainfall has been discussed in the literature as an external landslide's causing factor but cannot be considered alone without the physical characteristic changes of the soil at the near stage of sliding. The rainfall intensity and duration that are the basis of the hazard prediction cannot be determined because their values cannot be the same in all susceptible areas characterized by different environmental factors. Therefore, the field experimental analysis was crucial in this study to identify the threshold of both rainfall and soil water content, leading to water-induced shallow landslides in different susceptible locations.

4.3.5.1 Landslides, Slope, and Land Cover

Although steep slopes are associated with landslide occurrence, in this study, it has been realized that very steep slopes (45^{o} and higher) are not more prone to soil failure compared to slopes less than 45^{o} and greater than 25^{o} (Figure 32). Two reasons that may justify this statement are: (i) the most very steep slopes are characterized by sturdy rocks that make the slopes to be more stable, (ii) due to the high inclination, much rainfall water runs off instead of seeping into the soil compared to the moderate slopes. Apart from the slope inclination, the rainwater runoff also depends on the land cover and soil texture. There was no landslide indication observed on all sites covered by forest, although simulating rainfall for a long duration has been used for simulating rainfall (more than 11 hours). This is an indication of the role of forest cover in slope stability.

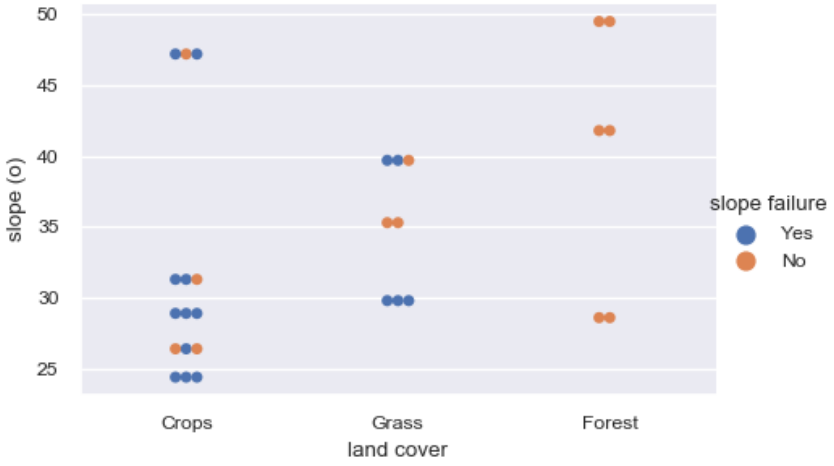


Figure 32. Slope failure vs. slope angle & land cover.

In the study area, most of the forest areas are at the same time covered by Eragrostis spectabilis, which is a natural grass type found in the high mountains in the study area. Site SA2 was covered by this type of grass and did not show any signs of slope failure.

4.3.5.2 Landslides, Slope and Soil Types

Table 10 shows that the soil types in most of the sampled sites are lean clay and elastic silt. The results do not really indicate which soil type is more susceptible to landslides. But the sites were selected based on the historical background of landslide events, and the laboratory tests revealed the two types of soil that are most dominant among the sampled sites. It was noted that the sites (lean clay or elastic silt) that did not indicate any sign of slope failure were those that were protected by forest or had a high slope angle. Therefore, lean clay and elastic silt are the most affected by landslides compared to the other types of soil in the sampled sites (Figure 33).

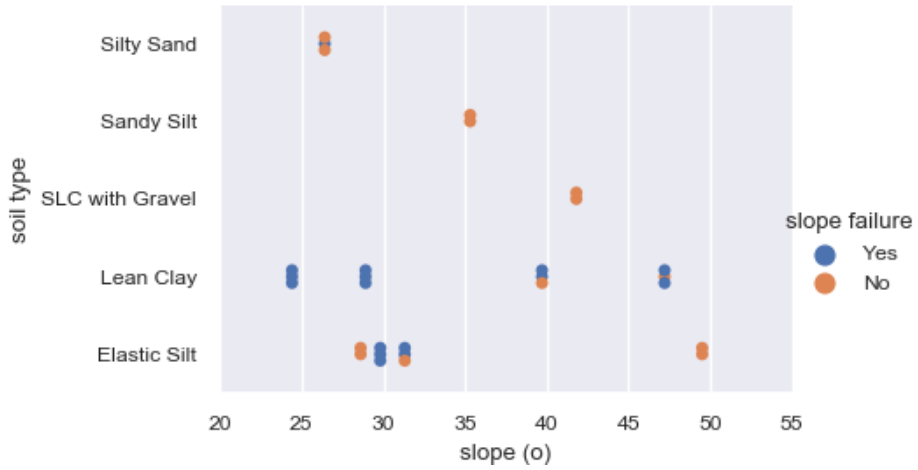


Figure 33. Slope failure vs. soil type.

4.3.6. Rainfall and Soil Moisture Thresholds

The maximum records of rainfall and soil moisture content from the experiments conducted in this study help us establish thresholds for both parameters in the specific sites (area of study). As stated earlier, three sensors were placed at various depths underground to determine which one could better predict landslide incidence. The red dots in Figure 34 are more clustered in the right lower corner of the figure, indicating that the slope failure was observed when sensor1 soil moisture was greater than 80% and 90% for sensors placed deeper (sensor2 and sensor3). The minimum rainfall inducing slope failure, as indicated by the same figure is around 70 mm. Therefore, we can conclude that these values can be used for local LEWS. On the other hand, rainfall can be used for regional LEWS as it is not possible to identify the soil moisture content at every site. Even though, a rainfall of more than 100 mm did not cause slope failure according to this study, such a daily rainfall depth is also dangerous as the increased rainwater runoff may cause floods, which can also depend on different factors (which are out of the scope of this study). It was also noted that the sites that did not experience any slope failure were those lands covered by forest or types of grass that reinforce the shear strength of the topsoil. But, under normal circumstances, a one-day rainfall of more than 70 mm should be taken into consideration for early warning systems by considering the antecedent rainfall and other geo-factors.

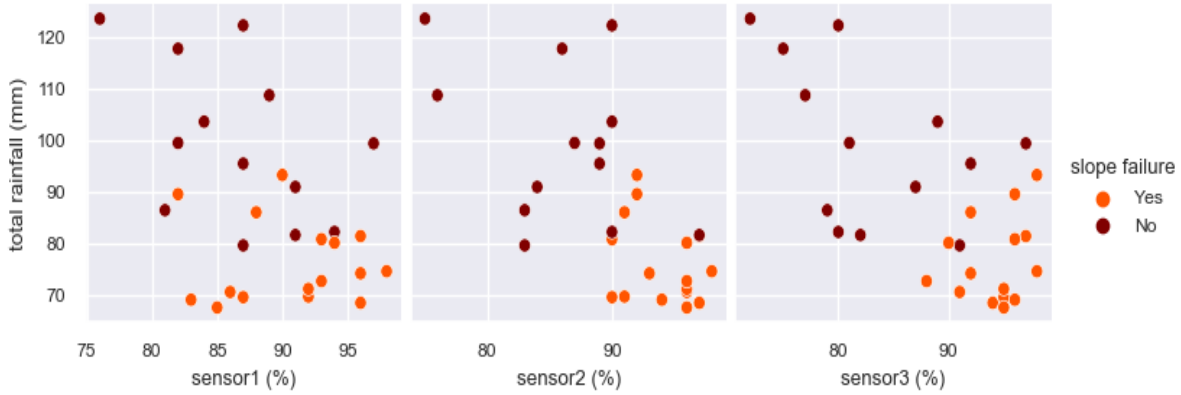


Figure 34. Rainfall and soil moisture thresholds.

As mentioned before, most of the non-landslide cases in this study are related to the forestland cover that persists to rainwater penetration and requires more rainfall duration. This explains the reason for having fewer incidences of high total rainfall as shown in Figure 34. It was also explained that rainwater runoff was much more on very steep slopes than on steep or medium slopes. Hence, the probability of slope failure increases with soil permeability, which is dependent on the soil texture, rainfall intensity, and duration.

4.4. Conclusions

In this study, field experiments were conducted at various sites selected in the surrounding areas that had landslide events in the past. The study consisted of the measurement of rainfall and soil water content using a rain gauge and soil moisture sensors. The site sample profile considered different parameters such as slopes, angles, soil texture, soil depths, and land coverage.

The experimental results show that rainfall triggers slope failure and that the total rainfall amount inducing this hazard depends on various other parameters. Rainfall alone cannot be considered as a parameter to predict slope failure and it has an impact on the hydrological properties of the soil. In addition, the level of soil water content at the near stage of slope failure differs from one site to another depending on the internal and external features. In general, steep slopes are more susceptible to shallow landslide incidence compared to very steep slopes. Furthermore, we noticed that land coverage plays an important role in the stability of the slope due to the more time required for saturation of land covered by natural grass or forest than that covered by plants. This is because the vegetation adjusts the hydrological equilibrium of the involved location through the evapotranspiration process, whereas roots add some reinforcement by increasing soil shear strength [178–180] and the degree of slope stabilization varies according to the vegetation [181–183].

According to the experimental results in this study, the following major insights are taken: (i) For LEWS, no common rainfall intensity or soil water content threshold could be determined. Instead, location-specific thresholds must be determined using empirical models. Then, the identified threshold can specifically be used to predict slope failure in areas with identical (or almost identical) geomorphological features. (ii) Depending on other geo-factors in specific locations, daily rainfall of more than 70 mm and soil water content of more than 90% can cause landslides. (iii) The thresholds found in this study can be used in designing local LEWS for areas having almost similar environmental covariates or soil forming factors, especially for cut slopes (manmade slopes) such as in-house plots or roads. (iv) This study proposes similar experiments to be conducted at various sites to derive site-specific thresholds to feed in to local IoT-based LEWS as the parameters differ from one zone to another.

CHAPTER 5

DESIGNING A SITE SPECIFIC LANDSLIDE EARLY WARNING SYSTEM (SSLEWS) FOR RISKS REDUCTION

This chapter describes a Site Specific Early Warning System (SSEWS) for rainfall-induced landslides using the soil moisture sensors deployed in the regions prone to landslides' incidences to minimize the risks caused by this natural threat. An experimental study was conducted to estimate the soil moisture level that induces a slope failure. The estimated thresholds were used in the development of the SSEWS prototype. The system on site consists of a sensor node that gathers the soil moisture data from remote sites (landslide prone areas). The collected data are transmitted using a GSM/GPRS (Global Systems for Mobile/General Packet Radio Service) module over the cellular network to the database server and analyzed for issuing alerts and visualized on a dashboard. The system was tested on different sites and shown to be 71.4% successful.

5.1 Introduction

Rainfall induced landslides have been challenging the world due to the loss they cause to lives, the natural environment, and infrastructure. Universally, different strategies and techniques have been used to mitigate or reduce the risks caused by this natural threat. Landslide mitigation strategies can be categorized into structural and non-structural strategies. Structural strategies are like drainage, walls, planting, directing, etc. Their implementation may be challenging in developing countries due to economic constraints. Non-structural measures are consequently reducing measures, which may include land use planning, public willingness, emergency management, early warning etc. [184].

Digital technologies appear to be the most effective landslide risks mitigation techniques used in different countries. The use of the digital technologies for landslide risk reduction comprises of digital mapping, landslides' prediction and early warning systems [184]. It has been proved that landslide early warning systems (LEWS) have a great importance in providing information about the imminent disaster so that precautions can be taken before the incidence and hence reduce the risks they would cause [185]. One of the digital technologies used nowadays is the Internet of Things (IoT). IoT supports quick data collection from the remote site, transmission, analysis, and feed them to the early warning system for alerting the people under risks. L. Piciullo et al. have discussed a landslide early warning systems (LEWS) that has proved efficient in providing information about the imminent disaster so that precautions can be taken before the incidence occurs and hence reduce the risks that would be caused [155].

Globally, the schematic representation of an early warning system proposed by the United Nations Office for Disaster Risk Reduction (UNDRR) and the World Meteorological

Organization (WMO) comprises four components: (i) disaster risk information, (ii) monitoring and analysis of threats, (iii) communication and warnings, (iv) readiness and response abilities [186,187]. Generally, monitoring and warning can be considered the main components of LEWS [188].

⊕ The monitoring process involves all operations related to the in situ installation of the instrumentation, data acquisition, processing, and analysis [189]. In situ installation involves on-field placement of measuring instruments such as meteorological stations, GPS systems, webcams, sensor nodes (inclinometer, piezometer, soil moisture, etc.) [190] to gather targeted data to be later processed for LEWS.

⊕ The data acquisition procedures involve setting the measurement frequency (cycles). The data are periodically captured by the monitoring instrument according to the preset cycle and transform them in the digital format. This process is performed locally and is terminated by the local export, which involves the transmission of data from the measuring node to the nearest gateway, and the central database (server).

⊕ The data transmission involves short range such as Bluetooth, Low Power Personal Area Network (6LoWPAN), etc. [190]; the long range data transmission uses technologies like LoRa (Long Range), SigFox, and cellular network technologies such as GPRS, 3G, and LTE (Long Term Evolution) [191].

⊕ The warning procedure involves the process of comparing the actual data to the preset threshold derived from the historical data collected through various experimentation.

5.2 Related works

Organizations in different countries have implemented LEWS using different techniques, whereas the research and studies are still ongoing to find a reliable solution for notification of landslide incidence prior to the occurrence. LEWS can be categorized as global, territorial, or local systems, operating at a global, regional, or local area scale, respectively [184]. Regional (also known as territorial) LEWS predict the rainfall-landslide occurrence over a large region (the area of concern covers a single slope) by monitoring meteorological parameters and providing an alert to the population, authorities, or other personnel in charge of disaster management [155,192]. Examples of regional LEWS are the Hong Kong Landslip Warning System, LEWS for the San Francisco Bay area (USA), and the LEWS for Java (Indonesia) as described in [32]. Instead, the local LEWS monitors variables inducing landslide occurrence at a small well-identified slope and provide warnings to the previously stated people [184], The Åknes rock slope in Norway is a typical example of this category [193]. Whether local, regional or worldwide, important components constituting LEWS are described similarly.

The most common components described by different authors for both local and territorial LEWS include monitoring, transmission, analysis, warning, and response [193]. T.F Fathani et al. proposed a seven sublayers schematic representation for LEWS by the extraction of four

components of UNDRR. They proposed a new layer of building the commitment of local authorities and communities while 3 sublayers were extracted from the preparedness and response capabilities layer: (i) establishment of a disaster preparedness and response team; (ii) evacuation map; (iii) and standard operating procedures for evacuation [188]. On the other hand, the structure proposed by [191] consists of three phases: (i) data acquisition procedures (which concern installation of measuring devices, onsite data collection, preprocessing, transmission, storage, and organization), (ii) early warning application (which defines the warning/alarm threshold), and (iii) monitoring results representation and dissemination (consisting of data representation for easy analysis and understanding).

The transmitted data is uploaded into the database where they will be processed and analyzed. This stage comprises various procedures, mainly data preprocessing, validation, storage, presentation for visualization, and analysis before issuing an alert if necessary.

The warning procedure involves the process of comparing the actual data to the preset threshold according to what has happened in the past. Threshold determination is done based on historical data analysis and landslide causing factor study in the specific area. Different thresholds should be used according to the warning level. Three levels are the most common: (i) Normal, (ii) Attention, (iii) Alarm [188]; Caution, Warning, Evaluate [194]. A model of different number of levels (more than 3) can also be used like those described by [189]: (i) normal situation, (ii) awareness, (iii) increased awareness, (iv) high hazard, (v) critical situation.

Regional LEWS use weather stations, radars, and other technologies that are expensive, so that non-developed countries cannot afford them. Besides, the coverage areas of LEWS can be a challenge for reliability. Local LEWS can overcome different challenges of territorial LEWS. The purpose of this study is to design and develop a tailored, cost-effective prototype for local LEWS. Many systems use rainfall data while few of them use soil moisture and wetness information [32].

The purpose of this study is to design and develop a tailored cost-effective prototype for Site-specific LEWS (SSLEWS). We used the soil moisture data, as recent studies showed that the hydrological threshold can improve the prediction capability [32]. Thresholds for rainfall and soil moisture have been determined through field experiments that found threshold values differ from one site to another depending on the geo-environmental characteristics such as slope, soil types, and land cover.

5.3 Data and System Architecture

5.3.1 Rainfall and Soil Moisture

Regional LEWS considers different rainfall characteristics such as current and antecedent rainfall, intensity, and duration. The local LEWS monitors landslide events using some more

parameters based on personal experience (or expert judgement) [193]. Soil moisture has been proved to be one of the most useful parameters [113,185,193,195–197].

To overcome this problem, we conducted an experimental study to determine rainfall thresholds for different areas. Different sites have been chosen according to the various geo-factors, including the slope inclination, soil types, and land cover/use. We used the rainfall simulator to provoke slope failure, and a rain gauge was used to capture the rainfall amount in a time series.

Although the hydrological thresholds are not much used in designing the LEWS, recent studies have shown that the in situ measurement of soil moisture has a crucial importance in the improvement of LEWS [32,185] because this parameter has specific information on local landslide activity [196]. The soil moisture thresholds on specific sites were identified concurrently with that of rainfall, as explained in the previous paragraph. Three sensors were placed underground at various depths (20 cm, 70 cm, and 120 cm). All data collected during field experiments are available online [198].

5.3.2 System Design

The site specific landslide early warning system (SSLEWS) is made through a wireless sensor network (Internet of Things) that uses sensor nodes, which are spatially distributed to collect soil moisture data, transmit them to the nearest base station of the cellular network, and then to the cloud for landslide prediction processing, storage, visualization, analysis, and alerting. The system is made up of three main layers: perception, network, and application, as shown in Figure 35.

The perception layer: consists of sensor nodes. The network layer consists of a gateway (base station) and other interconnecting devices of a cellular network. The application layer comprises data acquisition, storage, visual visualization, analysis, and alerting.

Sensor node: Each sensor node comprises the power supply, three soil moisture sensors, a microcontroller, and a GSM/GPRS module (Figure 36a). The power supply unit consists of an off-grid system comprising a solar panel, charge controller, battery, inverter, and enclosing accessories. The sensing part involves the capacitive soil moisture sensor v1.2 for measuring the ground water content. The specifications of the sensors are as follows: The device dimension is 98mm x 23mm, supports Gravity 3-pin interface, analog output, operating voltage is 3.3 ~ 5.5 VDC, and the output voltage is 0 ~ 3.0VDC. It works by measuring the variations in capacitance caused by the water infiltration into the soil and the analog output is the charge and discharge timing.

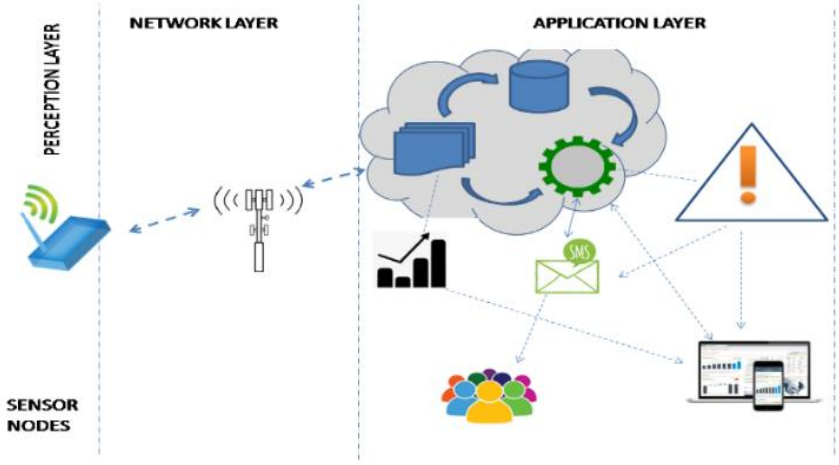


Figure 35. Proposed SSLEWS Network Architecture.

As shown by Figure 36 (b), the calibration was done in three steps: (i) measuring and recording the water content of dry soil; (ii) making them fully saturated by pouring water in the soil containers and recording the output value; and (iii) mapping the values of dry soil to 0 and those of fully saturated to 100 (the calibrated output ranges between 0-100). The soil moisture sensors were interfaced with a microcontroller (Arduino UNO) for central processing. In addition, the GSM/GPRS (SIM800L) module was connected to the microcontroller for connecting the sensor node to the server via a cellular network. The power unit and GSM module are interfaced by a DC-to-DC converter for adapting 12V from the power unit to the 4V needed for the GSM/GPRS module (Figure 36 (a)).

The sensor probes were buried underground at 0.5m, 1m, and 1.5m depths to allow the sensor node to measure the soil water content at various levels of the earth.

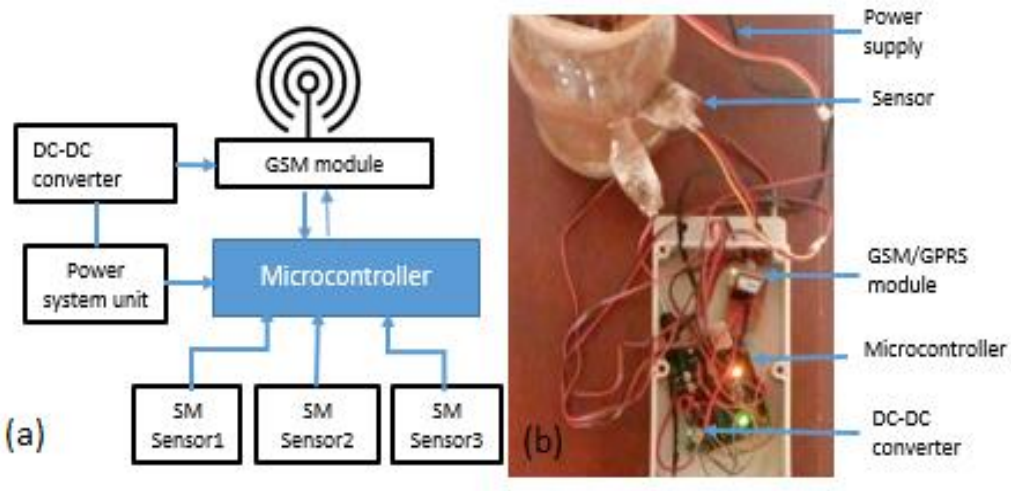


Figure 36. The block diagram of the sensor node, (b) inside the box and calibration.

Application: The application layer consists of data storage, processing, analysis, visualization, and alerting. The new data received from the network layer are stored in the database where they can also be visualized in real time. The incoming data are then compared to the threshold to classify the warning level (Table 11).

Table 11. SSLEWS Warning levels

Warning level	Description	Action
Level 1: Normal	No threshold attained	No action should be taken. Authorities/responsible agents can visualize the data.
Level 2: caution	One of the thresholds is attained. There is a possibility of a landslide occurrence.	Alarms should be sent through short messages on the mobile phones of local citizens. Data and warnings can be visualized by authorities or agents in-charge.
Level 3: Critical	Both thresholds are attained. There is a high possibility of a landslide occurrence,	An alarm message should be sent through short messages on the mobile phones of local citizens. Data and warnings can be visualized. People should be evacuated.

Thresholds for the soil moisture content that have been identified in our previous experimental study [20] were used to predict the disaster based on the new data and, via short messages, alerts are sent to the people located in the area of the disaster. The real time information together with the predicted disaster can also be accessed on the system dashboard so that further actions may be taken even before the alert is issued.

5.4 Experimental Results and Discussion

The prototype setup was preceded by an experimental study to identify the correlation between rainfall and soil moisture content. The study was conducted at various sites characterized by different factors such as soil types, slopes, and land cover. This activity is carried out by using a rainfall simulator on different sites while recording the rainfall amount from the rain gauge and soil moisture by capacitive soil moisture sensors. It took about 8 to 11 hours with a rainfall intensity of 7-11mm per hour to induce landslides on the selected sites. The result shows that there is a linear regression between rainfall and soil moisture content, as shown in Figure 37.

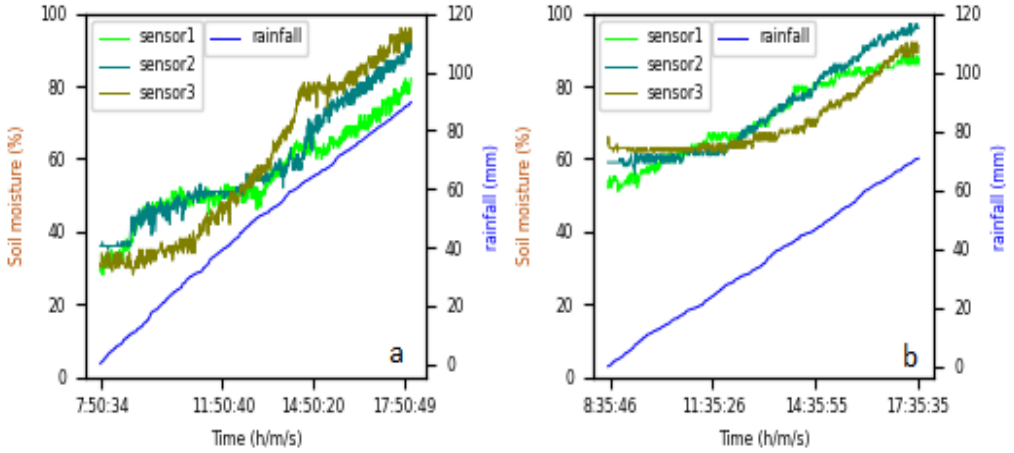


Figure 37. Correlation between rainfall and soil moisture. (a) Features characterizing those sites were Silty Sand as soil type, the slope inclination of 26%, land covered by crops. (b) site features were Lean Clay, slope inclination of 29%, land covered by crop.

An experimental study was conducted at 11 sites. Depending on the site specifications, some sites indicated the sign of slope failure, others did not. Six sites showed the signs of land sliding. Slope failures that occurred at different levels of rainfall and soil moisture are shown by Figure 38.

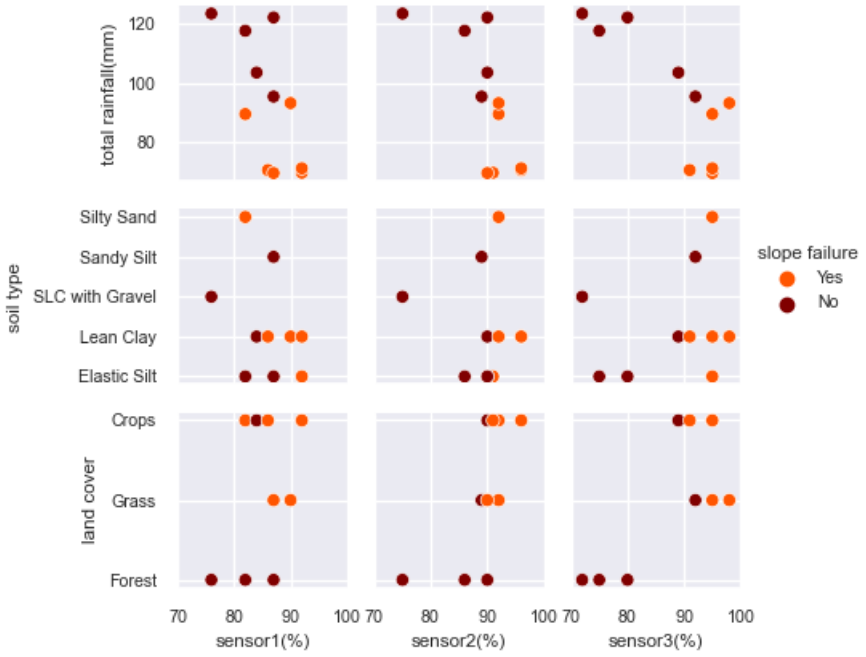


Figure 38. Slope failure vs sites' parameters.

The first row of Figure 38 plots rainfall vs. soil moisture; the second row plots soil types vs. soil moisture; and the third row plots land cover vs. soil moisture. The estimation of rainfall and soil moisture inducing landslides could be drawn from the plots in the above figure (Figure 38). The experiments revealed that slope failure occurred when the soil moisture content was at least 90% for the two sensors placed between 50–120 cm depth, whereas the top sensor (20 cm) recorded a minimum of 80%. It was also observed that Lean Clay and Elastic Silt are prone to slope failure as well as land covered by crops or grass. The prototype was tested on the land covered by crops, soil type of Lean Clay, and the following thresholds were used to provide various warning levels as described in Table 12.

Table 12. Warning levels and Soil moisture thresholds.

Warning level	**Soil moisture thresholds**		
	Sensor1(20 cm)	Sensor2 (70cm)	Sensor3 (120 cm)
1	<=60	<=70	<=70
2	>=60	>=70	>=70
3	>=80	>=90	>=90

Seven sites were selected to test the prototype. The test was done by wetting the soil above the cut slope above the house until a portion of the wet soil falls (or indicates the sign falling). Out of seven sites, five slopes showed the sign of failure (71.4%), while two sites could not fall even if the soil water content was above 92% for all sensors. The soil moisture contents recorded at each site are shown in Table 13.

Table 13. Soil moisture content at slope failure on the tested sites.

Sensor	**Site1**	**Site2**	**Site3**	**Site4**	**Site5**	**Site6**	**Site7**
1	88%	96%	93%	89%	93%	95%	83%
2	97%	99%	97%	88%	96%	92%	89%
3	94%	99%	95%	93%	99%	99%	92%
Slope failure?	Yes	No	Yes	Yes	No	Yes	Yes

The prototype of SSEWS was placed above the cut slopes of the homes with the historical background of slope failures. The sensor nodes collect the soil moisture content and transmit it to the remote database. In real time, data can be visualized and analysed over the system dashboard, as indicated by Figure 39.

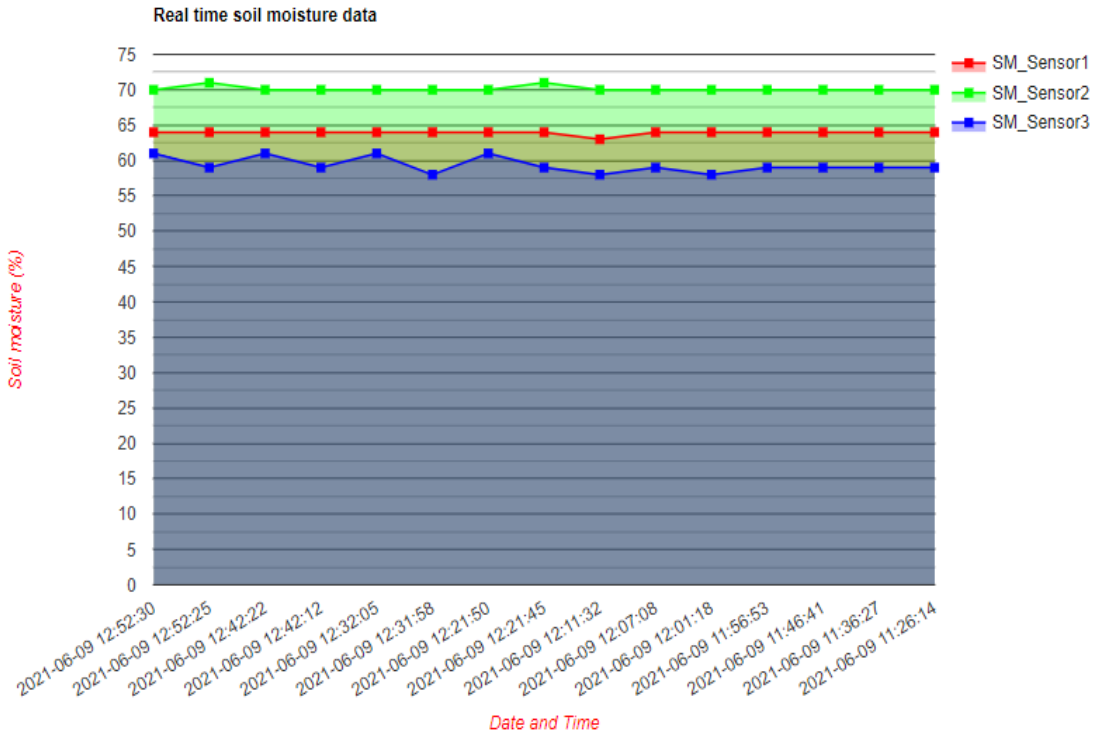


Figure 39. The real-time soil moisture data from the system dashboard.

In addition, an alert can be issued to the people in the area through the short message service (SMS), and only warning level 2 (Figure 40) and 3 (Figure 41) can be received via SMS whereas level 1 doesn't need any caution.

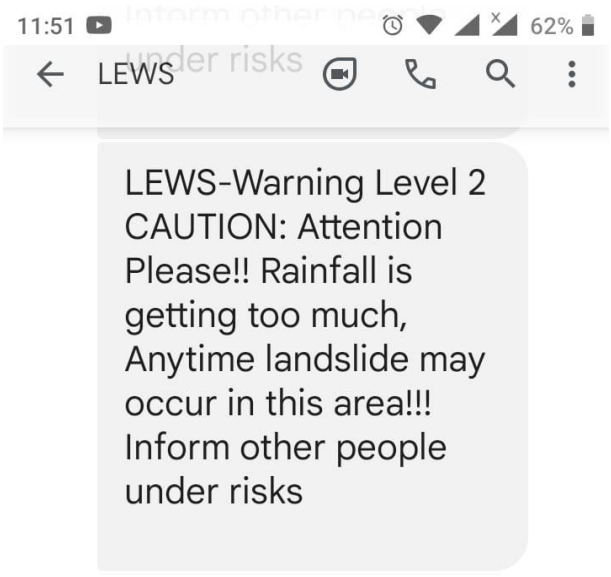


Figure 40. LEWS warning message (level 2)

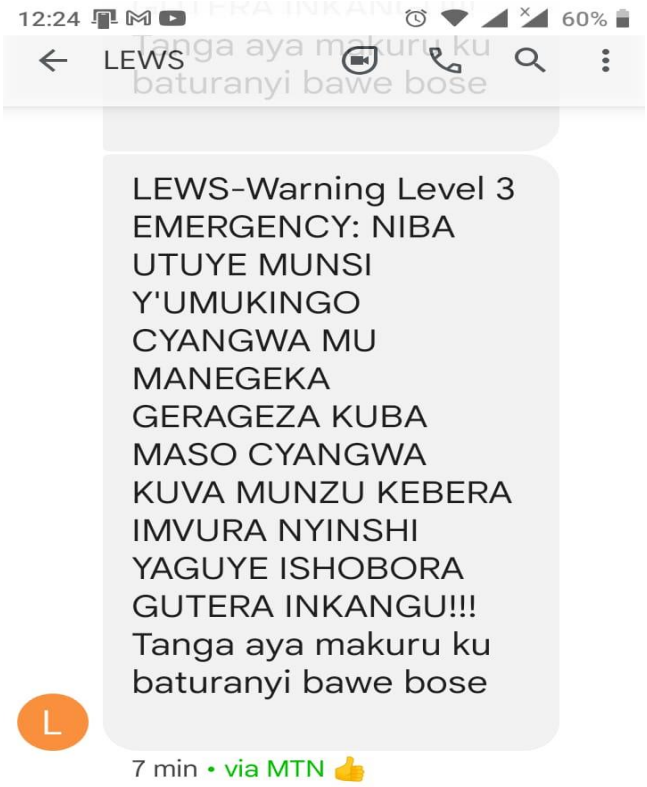


Figure 41. LEWS warning message (level 3)

Every ten minutes, the new data entry from the sensor node is recorded and stored in the database for later use and analysis. The system dashboard can be visualized online.

5.5 Conclusion

The main objective of this research is to design a low-cost SSLEWS for risks reduction in the landslide prone areas of Rwanda. The system prototype was developed, tested, and deployed. The data are collected from remote sites and pushed to a web database hosted on a remote server where they are stored, analyzed, visualized in a dashboard, Early warning messages are sent to the citizens around that area via SMS when the pre-set threshold values of the parameters inducing landslides are attained. In our future work, we will integrate rainfall data and machine learning into SSEWS.

CHAPTER 6

CONCLUSIONS AND RECOMMENDANTIONS

6.1 Conclusions

The aim of this synopsis was to present the outcomes of PhD research on the prediction model and development of an early warning system for rainfall induced landslide risks reduction. This research was conducted by analysing historical rainfall and landslide incidences data along with other external and internal factors to develop a prediction model for landslide incidences for an early warning system. The prediction of landslide occurrence was done by using Random Forest (RF) and logistic regression (LR) machine learning techniques to identify the best model that can be used for LEWS. Different parameters were used to train the models, and the results revealed that landslides are triggered due to too much daily rainfall or low intensity over a prolonged period, but in most cases, landslides occur after a few consecutive (like 2–5) days of precipitation. In addition to the rainfall data, other parameters have been utilized to assess their impact on the disaster. The areas of high slope angle are more susceptible to landslides than the regions where the terrain is almost a plateau. Upon the results of the models' evaluation, LR was considered the best model recommended for the early warning system.

After the identification of the best prediction model, an experimental study was conducted to determine the rainfall and soil water content thresholds that can be fed into the landslide early warning system (LEWS) using the IoT technology. Various experiments have been conducted for the real-time monitoring of slope failure using a toolset composed of a rain gauge, soil moisture sensors, and a rainfall simulating tool. The findings reveal that the landslide occurrence threshold is controlled not just by total rainfall amount (or intensity) or soil moisture but also by internal (geological, morphological) and environmental factors. The experiments revealed that for a specific site, the soil moisture inducing slope failure was above 90% for sensor nodes placed at 50 cm or deeper. These values were used as thresholds for LEWS for that specific site to improve predictions.

As presented in section 3.3, the site specific LEWS was designed, developed, and tested. The estimated thresholds determined in section 3.2 were used in the development of the SSEWS prototype. The system on the site consisted of a sensor node that gathers the soil moisture data from remote sites (test sites). The collected data are transmitted using a GSM/GPRS (Global Systems for Mobile/General Packet Radio Service) module over the cellular network to the database server and analysed for issuing alerts and visualized on a dashboard. The system was tested on different sites and shown to be 71.4% successful.

6.2 Recommendations

The following recommendations are drawn from the above conclusions:

- ⊕ Researchers should explore more machine learning models to find out which ones can provide improved prediction capabilities, especially the FNR.
- ⊕ The most efficient model has been discovered to be logistic regression, which is suggested for LEWS development.
- ⊕ A similar experimental study as those conducted in chapter four [20] is recommended for the specific site before implementing LLEWS. The purpose of such an experiment is to determine the threshold for two landslide predictors, which are rainfall and soil moisture content.
- ⊕ More research is recommended to determine the daily and antecedent rainfall that can be used for the RLEWS. However, any daily rainfall that is above 100 mm should be considered as dangerous precipitation and a warning to the inhabitants should be issued.
- ⊕ Local LEWS, such as SSLEWSs [199], that use both rainfall and soil moisture data are advised in landslide-prone areas because regional LEWS only use rainfall data and are hampered by poor weather station dispersion, which in most cases is insufficient to cover the full area being monitored.

References

[1] A. Smith, U.S. Billion-dollar Weather and Climate Disasters, 1980 - present, Natl. Centers Environ. Inf. (2020) 1–15. https://doi.org/10.25921/STKW-7W73.

[2] S. Hsiang, R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D.J. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer, K. Larsen, T. Houser, Estimating economic damage from climate change in the United States, Science (80-.). 356 (2017) 1362–1369. https://doi.org/10.1126/SCIENCE.AAL4369.).

[3] H.L. Koh, S.Y. Teh, Disaster Risk Reduction and Resilience Through Partnership and Collaboration, Springer Nature Singapore Pte Ltd., Singapore, 2019. https://doi.org/10.1007/978-3-319-71067-9_49-1.

[4] M.J. Crozier, Landslides: Causes, Consequences \& Environment, Croom Helm, 1986. https://books.google.rw/books?id=0Rs-AAAAIAAJ.

[5] D. Makajić-Nikolić, Disaster Risk Reduction, Springer-Verlag, Berlin Heidelberg, 2020. https://doi.org/10.1007/978-3-319-95885-9_65.

[6] REMA, Rwanda State of Environment and Outlook, Minist. Nat. Resour. 1 (2009) 98. https://www.rema.gov.rw/soe/background.php (accessed May 6, 2022).

[7] Food and agriculture organization of the United Nations, Int. Organ. 1 (1947) 350–353. https://doi.org/10.1017/S0020818300006160.

[8] J.B. Nsengiyumva, G. Luo, L. Nahayo, X. Huang, P. Cai, Landslide susceptibility assessment using spatial multi-criteria evaluation model in Rwanda, Int. J. Environ. Res. Public Health. 15 (2018). https://doi.org/10.3390/ijerph15020243.

[9] Rwanda: Flooding and landslides kill at least 65 nationwide May 7 /update 2 | Crisis24, (2020). https://crisis24.garda.com/alerts/2020/05/rwanda-flooding-and-landslides-kill-at-least-65-nationwide-may-7-update-2?origin=fr_riskalert (accessed May 6, 2022).

[10] World Meteorological Organization (WMO), Weather-related disasters increase over past 50 years, causing more damage but fewer deaths | World Meteorological Organization, World Meteorol. Organ. (2021). https://public.wmo.int/en/media/press-release/weather-related-disasters-increase-over-past-50-years-causing-more-damage-fewer (accessed December 10, 2021).

[11] L. Nahayo, C. Mupenzi, A. Kayiranga, F. Karamage, F. Ndayisaba, E.M. Nyesheja, L. Li, Early alert and community involvement: approach for disaster risk reduction in Rwanda, Nat. Hazards. 86 (2017) 505–517. https://doi.org/10.1007/s11069-016-2702-5.

[12] MIDIMAR, Republic of Rwanda Refugee Affairs National Contingency Plan for Floods and Landslides, Kigali, 2014.

[13] F. Byumvuhore, 2017 disasters cost Rwanda Rwf6.7 billion - MIDIMAR | The New Times | Rwanda, New Times. (2018). https://www.newtimes.co.rw/section/read/227141 (accessed May 21, 2022).

[14] Deadly Rwanda floods leave thousands homeless | Climate News | Al Jazeera, (2016). https://www.aljazeera.com/news/2016/5/9/deadly-rwanda-floods-leave-thousands-homeless (accessed August 20, 2021).

[15] W.S. Damaged, G. Reaction, Rwanda – At Least 49 Killed in Floods and Landslides , 500 Homes Destroyed, (2016) 1–5. https://floodlist.com/africa/rwanda-floods-landslides-gakenke-muhanga (accessed August 20, 2021).

[16] P.F. et Al., An Operative Early Warning System for Landslides Forecasting Based on Rainfall Thresholds and Soil Moisture, in: Margottini C., Canuti P., Sassa K. Landslide Sci. Pract., Springer, Berlin, Heidelber, 2013. https://doi.org/https://doi.org/10.1007/978-3-642-31445-2_82.

[17] D. Tiranti, D. Rabuffetti, Estimation of rainfall thresholds triggering shallow landslides for an operational warning system implementation, Landslides. 7 (2010) 471–481. https://doi.org/10.1007/s10346-010-0198-8.

[18] G. Martelloni, S. Segoni, R. Fanti, F. Catani, Rainfall thresholds for the forecasting of landslide occurrence at regional scale, Landslides. 9 (2012) 485–495. https://doi.org/10.1007/s10346-011-0308-2.

[19] S. Segoni, A. Battistini, G. Rossi, A. Rosi, D. Lagomarsino, F. Catani, S. Moretti, N. Casagli, Technical Note: An operational landslide early warning system at regional scale based on space-time-variable rainfall thresholds, Nat. Hazards Earth Syst. Sci. 15 (2015) 853–861. https://doi.org/10.5194/nhess-15-853-2015.

[20] M. Kuradusenge, S. Kumaran, M. Zennaro, A. Niyonzima, Experimental Study of Site-Specific Soil Water Content and Rainfall Inducing Shallow Landslides: Case of Gakenke District, Rwanda, Geofluids. 2021 (2021) 7194988. https://doi.org/10.1155/2021/7194988.

[21] L. Nahayo, C. Mupenzi, A. Kayiranga, F. Karamage, F. Ndayisaba, E.M. Nyesheja, L. Li, Early alert and community involvement: approach for disaster risk reduction in Rwanda, Nat. Hazards. 86 (2017) 505–517. https://doi.org/10.1007/s11069-016-2702-5.

[22] T.N. Singh, S.P. Pradhan, V. Vishal, Landslides: Theory, Practice and Modelling, 2019. http://link.springer.com/10.1007/978-3-319-77377-3.

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

[24] O. Hungr, S.G. Evans, M.J. Bovis, J.N. Hutchinson, A review of the classification of landslides of the flow type, Environ. Eng. Geosci. 7 (2001) 221–238. https://doi.org/10.2113/gseegeosci.7.3.221.

[25] J. Dou, U. Paudel, T. Oguchi, S. Uchiyama, Y.S. Hayakawa, Shallow and deep-seated landslide differentiation using support vector machines: A case study of the chuetsu area, Japan, Terr. Atmos. Ocean. Sci. 29 (2015) 227–239. https://doi.org/10.3319/TAO.2014.12.02.07(EOSI).

[26] Deep-seated and shallow-rapid landslides: know the difference, Washingt. For. Prot. Assoc. (2017). https://www.wfpa.org/news-resources/blog/deep-seated-landslides-shallow-landslides-washington/ (accessed January 20, 2022).

[27] L. Shano, T.K. Raghuvanshi, M. Meten, Landslide susceptibility evaluation and hazard zonation techniques – a review, Geoenvironmental Disasters. 7 (2020). https://doi.org/10.1186/s40677-020-00152-0.

[28] R. Hidayat, S.J. Sutanto, A. Hidayah, B. Ridwan, A. Mulyana, Development of a landslide early warning system in Indonesia, Geosci. 9 (2019) 1–17. https://doi.org/10.3390/geosciences9100451.

[29] Z. Liao, Y. Hong, J. Wang, H. Fukuoka, K. Sassa, D. Karnawati, F. Fathani, Prototyping an experimental early warning system for rainfall-induced landslides in Indonesia using satellite remote sensing and geospatial datasets, Landslides. 7 (2010) 317–324. https://doi.org/10.1007/s10346-010-0219-7.

[30] D. Karnawati, T.F. Fathani, S. Ignatius, B. Andayani, D. Legono, P.W. Burton, Landslide hazard and community-based risk reduction effort in Karanganyar and the surrounding area, central Java, Indonesia, J. Mt. Sci. 8 (2011) 149–153. https://doi.org/10.1007/s11629-011-2107-6.

[31] E.W. Brand, J. Premchitt, H.B. Phillipson, Relationship between rainfall and landslides in Hong Kong, in: Proc. 4th Int. Symp. Landslides, Canadian Geotechnical Society Toronto, 1984: pp. 276–284.

[32] F. Guzzetti, S.L. Gariano, S. Peruccacci, M.T. Brunetti, I. Marchesini, M. Rossi, M. Melillo, Geographical landslide early warning systems, Earth-Science Rev. 200 (2020) 102973. https://doi.org/10.1016/j.earscirev.2019.102973.

[33] R.S.Z. D.K. Keefer, R.C. Wilson, R.K. Mark, E.E. Brabb, W.M. Brown, S.D. Ellen, E.L. Harp, G.F. Wieczorek, C.S. Alger, Real-Time Landslide Warning During Heavy Rainfall, Science (80-.). 238 (1987) 921 LP – 925. https://doi.org/10.1126/science.238.4829.921.

[34] M. Jakob, T. Owen, T. Simpson, A regional real-time debris-flow warning system for the District of North Vancouver, Canada, Landslides. 9 (2012) 165–178.

https://doi.org/10.1007/s10346-011-0282-8.

[35] L.W. Wei, C.M. Huang, H. Chen, C.T. Lee, C.C. Chi, C.L. Chiu, Adopting the I3-R24 rainfall index and landslide susceptibility for the establishment of an early warning model for rainfall-induced shallow landslides, Nat. Hazards Earth Syst. Sci. 18 (2018) 1717–1733. https://doi.org/10.5194/nhess-18-1717-2018.

[36] G. Brigandì, G.T. Aronica, B. Bonaccorso, R. Gueli, G. Basile, Flood and landslide warning based on rainfall thresholds and soil moisture indexes: the HEWS (Hydrohazards Early Warning System) for Sicily, Adv. Geosci. 44 (2017) 79–88. https://doi.org/10.5194/adgeo-44-79-2017.

[37] S. Peruccacci, M.T. Brunetti, S.L. Gariano, M. Melillo, M. Rossi, F. Guzzetti, Rainfall thresholds for possible landslide occurrence in Italy, Geomorphology. 290 (2017) 39–57. https://doi.org/10.1016/j.geomorph.2017.03.031.

[38] L. Piciullo, M. Calvello, J.M. Cepeda, Territorial early warning systems for rainfall-induced landslides, Earth-Science Rev. 179 (2018) 228–247. https://doi.org/10.1016/j.earscirev.2018.02.013.

[39] F. Aji Purnomo, N. Maulana Yoeseph, G. Wijang Abisatya, Landslide early warning system based on arduino with soil movement and humidity sensors, J. Phys. Conf. Ser. 1153 (2019). https://doi.org/10.1088/1742-6596/1153/1/012034.

[40] M. V. Ramesh, Real-time wireless sensor network for landslide detection, Proc. - 2009 3rd Int. Conf. Sens. Technol. Appl. SENSORCOMM 2009. (2009) 405–409. https://doi.org/10.1109/SENSORCOMM.2009.67.

[41] G.N.L.R. Teja, V.K.R. Harish, D.N.M. Khan, R.B. Krishna, R. Singh, S. Chaudhary, Land Slide detection and monitoring system using wireless sensor networks (WSN), in: 2014 IEEE Int. Adv. Comput. Conf., 2014: pp. 149–154. https://doi.org/10.1109/IAdCC.2014.6779310.

[42] Y. Wang, Z. Liu, D. Wang, Y. Li, J. Yan, Anomaly Detection and Visual Perception for Landslide Monitoring Based on a Heterogeneous Sensor Network, IEEE Sens. J. 17 (2017) 4248–4257. https://doi.org/10.1109/JSEN.2017.2704584.

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

[44] H. Hong, Y. Miao, J. Liu, A.-X. Zhu, Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping, CATENA. 176 (2019) 45–64. https://doi.org/https://doi.org/10.1016/j.catena.2018.12.035.

[45] H. Wang, L. Zhang, K. Yin, H. Luo, J. Li, Landslide identification using machine learning, Geosci. Front. 12 (2021) 351–364. https://doi.org/10.1016/j.gsf.2020.02.012.

[46] M. Azarafza, M. Azarafza, H. Akgün, P.M. Atkinson, R. Derakhshani, Deep learning-based landslide susceptibility mapping, Sci. Rep. 11 (2021) 1–16. https://doi.org/10.1038/s41598-021-03585-1.

[47] P. Tsangaratos, I. Ilia, Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece, Landslides. 13 (2016) 305–320. https://doi.org/10.1007/s10346-015-0565-6.

[48] P. Tsangaratos, I. Ilia, Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size, CATENA. 145 (2016) 164–179. https://doi.org/https://doi.org/10.1016/j.catena.2016.06.004.

[49] A. Akgun, A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at İzmir, Turkey, Landslides. 9 (2012) 93–106. https://doi.org/10.1007/s10346-011-0283-7.

[50] H. Saito, D. Nakayama, H. Matsuyama, Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan, Geomorphology. 109 (2009) 108–121. https://doi.org/https://doi.org/10.1016/j.geomorph.2009.02.026.

[51] T. Kavzoglu, E.K. Sahin, I. Colkesen, Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression, Landslides. 11 (2014) 425–439. https://doi.org/10.1007/s10346-013-0391-7.

[52] B. Kalantar, B. Pradhan, S.A. Naghibi, A. Motevalli, S. Mansor, Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN), Geomatics, Nat. Hazards Risk. 9 (2018) 49–69. https://doi.org/10.1080/19475705.2017.1407368.

[53] B.T. Pham, D. Tien Bui, I. Prakash, M.B. Dholakia, Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS, CATENA. 149 (2017) 52–63. https://doi.org/https://doi.org/10.1016/j.catena.2016.09.007.

[54] Y. Li, R. Sun, K. Yin, Y. Xu, B. Chai, L. Xiao, Forecasting of landslide displacements using a chaos theory based wavelet analysis-Volterra filter model, Sci. Rep. 9 (2019) 19853. https://doi.org/10.1038/s41598-019-56405-y.

[55] C.H. Zhu, G.D. Hu, Time Series Prediction of Landslide Displacement Using SVM Model: Application to Baishuihe Landslide in Three Gorges Reservoir Area, China,

Appl. Mech. Mater. 239–240 (2013) 1413–1420. https://doi.org/10.4028/www.scientific.net/AMM.239-240.1413.

[56] J. Du, K. Yin, S. Lacasse, Displacement prediction in colluvial landslides, Three Gorges Reservoir, China, Landslides. 10 (2013) 203–218. https://doi.org/10.1007/s10346-012-0326-8.

[57] M. Krkač, D. Špoljarić, S. Bernat, S.M. Arbanas, Method for prediction of landslide movements based on random forests, Landslides. 14 (2017) 947–960. https://doi.org/10.1007/s10346-016-0761-z.

[58] P. Xie, A. Zhou, B. Chai, The Application of Long Short-Term Memory(LSTM) Method on Displacement Prediction of Multifactor-Induced Landslides, IEEE Access. 7 (2019) 54305–54311. https://doi.org/10.1109/ACCESS.2019.2912419.

[59] B. Yang, K. Yin, S. Lacasse, Z. Liu, Time series analysis and long short-term memory neural network to predict landslide displacement, Landslides. 16 (2019) 677–694. https://doi.org/10.1007/s10346-018-01127-x.

[60] H. Thirugnanam, M.V. Ramesh, V.P. Rangan, Enhancing the reliability of landslide early warning systems by machine learning, Landslides. 17 (2020) 2231–2246. https://doi.org/10.1007/s10346-020-01453-z.

[61] A.L. van Natijne, R.C. Lindenbergh, T.A. Bogaard, Machine learning: New potential for local and regional deep-seated landslide nowcasting, Sensors (Switzerland). 20 (2020) 1–18. https://doi.org/10.3390/s20051425.

[62] M.A. Thomas, B.D. Collins, B.B. Mirus, Assessing the Feasibility of Satellite-Based Thresholds for Hydrologically Driven Landsliding, Water Resour. Res. 55 (2019) 9006–9023. https://doi.org/10.1029/2019WR025577.

[63] I. Douglas, K. Alam, M. Maghenda, Y. Mcdonnell, L. Mclean, J. Campbell, Unjust waters: Climate change, flooding and the urban poor in Africa, Environ. Urban. 20 (2008) 187–205. https://doi.org/10.1177/0956247808089156.

[64] Y. Huang, L. Zhao, Review on landslide susceptibility mapping using support vector machines, Catena. 165 (2018) 520–529. https://doi.org/10.1016/j.catena.2018.03.003.

[65] K. Zhang, X. Xue, Y. Hong, J.J. Gourley, N. Lu, Z. Wan, Z. Hong, R. Wooten, iCRESTRIGRS: A coupled modeling system for cascading flood-landslide disaster forecasting, Hydrol. Earth Syst. Sci. Discuss. (2016) 1–23. https://doi.org/10.5194/hess-2016-143.

[66] Republic of Rwanda Ministry in Charge of Emergency Management (MINEMA), The National Risk Atlas of Rwanda, Kigali, 2015. https://reliefweb.int/report/rwanda/national-risk-atlas-rwanda.

[67] S. Jeong, K. Lee, J. Kim, Y. Kim, Analysis of rainfall-induced landslide on unsaturated soil slopes, Sustain. 9 (2017) 1–20. https://doi.org/10.3390/su9071280.

[68] A.M. Youssef, H.R. Pourghasemi, Z.S. Pourtaghi, M.M. Al-Katheeri, Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia, Landslides. 13 (2016) 839–856. https://doi.org/10.1007/s10346-015-0614-1.

[69] M. Hong, J. Kim, S. Jeong, Rainfall intensity-duration thresholds for landslide prediction in South Korea by considering the effects of antecedent rainfall, Landslides. 15 (2018) 523–534. https://doi.org/10.1007/s10346-017-0892-x.

[70] K. Zhang, X. Xue, Y. Hong, J.J. Gourley, N. Lu, Z. Wan, Z. Hong, R. Wooten, ICRESTRIGRS: A coupled modeling system for cascading flood-landslide disaster forecasting, Hydrol. Earth Syst. Sci. 20 (2016) 5035–5048. https://doi.org/10.5194/hess-20-5035-2016.

[71] M. Diakakis, Rainfall thresholds for flood triggering. The case of Marathonas in Greece, Nat. Hazards. 60 (2012) 789–800. https://doi.org/10.1007/s11069-011-9904-7.

[72] L. Nahayo, C. Mupenzi, A. Kayiranga, F. Karamage, F. Ndayisaba, E.M. Nyesheja, L. Li, Early alert and community involvement: approach for disaster risk reduction in Rwanda, Nat. Hazards. 86 (2017) 505–517. https://doi.org/10.1007/s11069-016-2702-5.

[73] S. Lee, J.H. Ryu, J.S. Won, H.J. Park, Determination and application of the weights for landslide susceptibility mapping using an artificial neural network, Eng. Geol. 71 (2004) 289–302. https://doi.org/10.1016/S0013-7952(03)00142-X.

[74] D.M. Staley, J.W. Kean, S.H. Cannon, K.M. Schmidt, J.L. Laber, Objective definition of rainfall intensity-duration thresholds for the initiation of post-fire debris flows in southern California, Landslides. 10 (2013) 547–562. https://doi.org/10.1007/s10346-012-0341-9.

[75] M. Marjanović, M. Kovačević, B. Bajat, V. Voženílek, Landslide susceptibility assessment using SVM machine learning algorithm, Eng. Geol. 123 (2011) 225–234. https://doi.org/10.1016/j.enggeo.2011.09.006.

[76] M.-J. Lee, Rainfall and Landslide Correlation Analysis and Prediction of Future Rainfall Base on Climate Change, in: Geohazards Caused by Hum. Act., InTech, 2016. https://doi.org/10.5772/64694.

[77] Republic Of Rwanda Western Province Ngororero District, 2013. www.ngororero.gov.rw.

[78] H. Modeling, Comparison of Spatial Interpolation Schemes for Rainfall Data and

Application in, (2017) 1–18. https://doi.org/10.3390/w9050342.

[79] S. Ly, C. Charles, A. Degré, Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale . A review, 17 (2013) 392–406.

[80] A. Trigila, C. Iadanza, C. Esposito, G. Scarascia-Mugnozza, Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy), Geomorphology. 249 (2015) 119–136. https://doi.org/10.1016/j.geomorph.2015.06.001.

[81] P. Vorpahl, H. Elsenbeer, M. Märker, B. Schröder, How can statistical models help to determine driving factors of landslides?, Ecol. Modell. 239 (2012) 27–39. https://doi.org/10.1016/j.ecolmodel.2011.12.007.

[82] L. Breiman, Random Forests, Mach. Learn. 45 (2001) 5–32. https://doi.org/10.1023/A:1010933404324.

[83] Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie, B. Qiang, RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2020: pp. 503–515. https://doi.org/10.1007/978-3-030-62008-0_35.

[84] H.R. Pourghasemi, N. Kerle, Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran, Environ. Earth Sci. 75 (2016) 1–17. https://doi.org/10.1007/s12665-015-4950-1.

[85] H.R. Pourghasemi, O. Rahmati, Prediction of the landslide susceptibility: Which algorithm, which precision?, Catena. 162 (2018) 177–192. https://doi.org/10.1016/j.catena.2017.11.022.

[86] Y. Wang, X. Wu, Z. Chen, F. Ren, L. Feng, Q. Du, Optimizing the predictive ability of machine learning methods for landslide susceptibility mapping using smote for Lishui city in Zhejiang province, China, Int. J. Environ. Res. Public Health. 16 (2019). https://doi.org/10.3390/ijerph16030368.

[87] J.C. Kim, S. Lee, H.S. Jung, S. Lee, Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea, Geocarto Int. 33 (2018) 1000–1015. https://doi.org/10.1080/10106049.2017.1323964.

[88] W. Lin, Z. Wu, L. Lin, A. Wen, J. Li, An ensemble random forest algorithm for insurance big data analysis, IEEE Access. 5 (2017) 16568–16575. https://doi.org/10.1109/ACCESS.2017.2738069.

[89] G. Louppe, Understanding Random Forests: From Theory to Practice, (2014). http://arxiv.org/abs/1407.7502.

[90] P. V. Gorsevski, P.E. Gessler, R.B. Foltz, W.J. Elliot, Spatial prediction of landslide hazard using logistic regression and ROC analysis, Trans. GIS. 10 (2006) 395–415. https://doi.org/10.1111/j.1467-9671.2006.01004.x.

[91] J. Mathew, V.K. Jha, G.S. Rawat, Landslide susceptibility zonation mapping and its validation in part of Garhwal Lesser Himalaya, India, using binary logistic regression analysis and receiver operating characteristic curve method, Landslides. 6 (2009) 17–26. https://doi.org/10.1007/s10346-008-0138-z.

[92] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (2006) 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

[93] P. Frattini, G. Crosta, A. Carrara, Techniques for evaluating the performance of landslide susceptibility models, Eng. Geol. 111 (2010) 62–72. https://doi.org/10.1016/j.enggeo.2009.12.004.

[94] A. Trigila, C. Iadanza, C. Esposito, G. Scarascia-Mugnozza, Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy), Geomorphology. 249 (2015) 119–136. https://doi.org/10.1016/j.geomorph.2015.06.001.

[95] H.A.H. Al-Najjar, B. Kalantar, B. Pradhan, V. Saeidi, Conditioning factor determination for mapping and prediction of landslide susceptibility using machine learning algorithms, in: SPIE-Intl Soc Optical Eng, 2019: p. 19. https://doi.org/10.1117/12.2532687.

[96] B.T. Pham, L.H. Son, T.A. Hoang, D.M. Nguyen, D. Tien Bui, Prediction of shear strength of soft soil using machine learning methods, Catena. 166 (2018) 181–191. https://doi.org/10.1016/j.catena.2018.04.004.

[97] L. Lombardo, P.M. Mai, Presenting logistic regression-based landslide susceptibility results, Eng. Geol. 244 (2018) 14–24. https://doi.org/10.1016/j.enggeo.2018.07.019.

[98] G. Capparelli, P. Versace, FLaIR and SUSHI: Two mathematical models for early warning of landslides induced by rainfall, Landslides. 8 (2011) 67–79. https://doi.org/10.1007/s10346-010-0228-6.

[99] C.N. Madawala, B.T.G.S. Kumara, L. Indrathilaka, Novel machine learning ensemble approach for landslide prediction, Proc. - IEEE Int. Res. Conf. Smart Comput. Syst. Eng. SCSE 2019. (2019) 78–84. https://doi.org/10.23919/SCSE.2019.8842762.

[100] Y. Achour, H.R. Pourghasemi, How do machine learning techniques help in increasing accuracy of landslide susceptibility maps?, Geosci. Front. 11 (2020) 871–883. https://doi.org/10.1016/j.gsf.2019.10.001.

[101] F.S. Tehrani, G. Santinelli, M. Herrera, A framework for predicting rainfall-induced

landslides using machine learning methods, 17th Eur. Conf. Soil Mech. Geotech. Eng. ECSMGE 2019 - Proc. 2019-Septe (2019). https://doi.org/10.32075/17ECSMGE-2019-0521.

[102] A.M. Youssef, H.R. Pourghasemi, Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia, Geosci. Front. 12 (2021) 639–655. https://doi.org/10.1016/j.gsf.2020.05.010.

[103] R. Archana Reddy, R. Gobinath, C.S. Khanna, G. Shyamala, Machine Learning based Landslide Prediction System for Hilly Areas, IOP Conf. Ser. Mater. Sci. Eng. 981 (2020). https://doi.org/10.1088/1757-899X/981/3/032084.

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

[105] B. Ghasemian, H. Shahabi, A. Shirzadi, N. Al-Ansari, A. Jaafari, V.R. Kress, M. Geertsema, S. Renoud, A. Ahmad, A Robust Deep-Learning Model for Landslide Susceptibility Mapping: A Case Study of Kurdistan Province, Iran, Sensors. 22 (2022) 1–28. https://doi.org/10.3390/s22041573.

[106] D. Utomo, S.F. Chen, P.A. Hsiung, Landslide prediction with model switching, Appl. Sci. 9 (2019). https://doi.org/10.3390/app9091839.

[107] B.T. Pham, B. Pradhan, D. Tien Bui, I. Prakash, M.B. Dholakia, A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India), Environ. Model. Softw. 84 (2016) 240–250. https://doi.org/10.1016/j.envsoft.2016.07.005.

[108] S. Saha, R. Sarkar, J. Roy, T.K. Hembram, S. Acharya, G. Thapa, D. Drukpa, Measuring landslide vulnerability status of Chukha, Bhutan using deep learning algorithms, Sci. Rep. 11 (2021) 1–23. https://doi.org/10.1038/s41598-021-95978-5.

[109] H. An, T. The, G. Lee, Y. Kim, M. Kim, S. Noh, J. Noh, Environmental Modelling & Software Development of time-variant landslide-prediction software considering three-dimensional subsurface unsaturated fl ow, Environ. Model. Softw. 85 (2016) 172–183. https://doi.org/10.1016/j.envsoft.2016.08.009.

[110] A. Hidalgo, M.P. Obando, M.P. Obando, Effect of of the the Rainfall Rainfall Infiltration Infiltration Processes Processes on on the the Effect Landslide Hazard Hazard Assessment Assessment of of Unsaturated Unsaturated Soils Soils in in Landslide Tropical Mountainous Mountainous Regions Regi, (2018). https://doi.org/10.5772/intechopen.70821.

[111] MINEMA, Minstry of Disaster Management and Refugee Affairs,Annual report 2016

desaster response and recovery unity, Kigali, 2016. http://minema.gov.rw/fileadmin/user_upload/ANNUAL_DISASTER_EFFECTS_REPORT_2016.pdf.

[112] MINEMA, Republic of Rwanda Ministry in Charge of Emergency Management, Kigali, 2018. http://minema.gov.rw/fileadmin/user_upload/Annual_Disaster_effects_Report_2018.pdf.

[113] A.J. Posner, K.P. Georgakakos, Soil moisture and precipitation thresholds for real-time landslide prediction in El Salvador, Landslides. 12 (2015) 1179–1196. https://doi.org/10.1007/s10346-015-0618-x.

[114] M.B. Kenanoğlu, Effect of Unsaturated Soil Properties on the Intensity- Duration Threshold for Rainfall Triggered Landslides *, (2019) 9009–9028.

[115] H. Framework, H. Yeh, Analyzing the Effect of Soil Hydraulic Conductivity Anisotropy on Slope Stability Using a Coupled, (2018). https://doi.org/10.3390/w10070905.

[116] M.G. Persichillo, M. Bordoni, M. Cavalli, S. Crema, C. Meisina, Catena The role of human activities on sediment connectivity of shallow landslides, Catena. 160 (2018) 261–274. https://doi.org/10.1016/j.catena.2017.09.025.

[117] A.G. Li, Z.Q. Yue, L.G. Tham, C.F. Lee, K.T. Law, Field-monitored variations of soil moisture and matric suction in a saprolite slope, 26 (2005) 13–26. https://doi.org/10.1139/T04-069.

[118] S. Alsubal1, N. bin Sapari, H. Mohammed, Indra S.H., A. Al-Bared,Mohammed, A review on mechanism of rainwater in triggering landslide, in: IOP Conf. Ser. Mater. Sci. Eng. 513012009, IOP Publishing Ltd, 2019. https://doi.org/10.1088/1757-899X/513/1/012009.

[119] V.F. Chiorean, Determination of Matric Suction and Saturation Degree for Unsaturated Soils, Comparative Study - Numerical Method versus Analytical Method, IOP Conf. Ser. Mater. Sci. Eng. 245 (2017). https://doi.org/10.1088/1757-899X/245/3/032074.

[120] M.Y. Fattah, A.H. Al-lami, M. Ahmed, Effect of Initial Water Content on the Properties of Compacted Expansive Unsaturated Soil, (2015).

[121] N. Lu, M. Asce, W.J. Likos, M. Asce, Suction Stress Characteristic Curve for Unsaturated Soil, J. Geotech. Geoenvironmental Eng. 132 (2006) 131–142.

[122] C. W.W. Ng, B. Menzies, Advanced Unsaturated Soil Mechanics and Engineering, Taylor & Francis, New York, USA, 2007.

[123] R. Paul, Field Measurement of Soil Suction Using Thermal Conductivity Matric Potential Sensors, University of Saskatchewan, 1988.

[124] L. Ontrasio, L. Schilirò, Inferences on modeling rainfall-induced shallow landslides from experimental observations on stratified soils, Ital. J. Eng. Geol. Environ. 2 (2019). https://doi.org/10.4408/IJEGE.2018-02.O-06.

[125] B. Indraratna, Numerical analysis of matric suction effects of tree roots, 159 (2006) 77–90.

[126] D.G. Fredlund, H. Rahardjo, M.D. Fredlund, <Fredlund_Unsaturated_Soil_Mechanics_in_E.pdf>, John Wiley & Sons, Inc., New Jersey, 2012.

[127] T. Tsai, Influences of soil water characteristic curve on rainfall-induced shallow landslides, (2011) 449–459. https://doi.org/10.1007/s12665-010-0868-9.

[128] M. Jakob, I.T.O.I.T. Simpson, A regional real-time debris-flow warning system for the District of North Vancouver , Canada, (2012) 165–178. https://doi.org/10.1007/s10346-011-0282-8.

[129] M. Hong, J. Kim, S. Jeong, Rainfall intensity-duration thresholds for landslide prediction in South Korea by considering the effects of antecedent rainfall, Landslides. 15 (2018) 523–534. https://doi.org/10.1007/s10346-017-0892-x.

[130] L. Montrasio, R. Valentino, Experimental analysis and modelling of shallow landslides, Landslides. 4 (2007) 291–296. https://doi.org/10.1007/s10346-007-0082-3.

[131] Lumb P., Effect of rainstorms on slope stability, in: Symp. Hong Kong Soils, Hong Kong, 1962: pp. 73–87.

[132] Lumb P., Slope failures in Hong Kong, Q. J. Eng. Geol. 8 (1975) 31–65.

[133] P. Brand, E.W., Premchitt, J., Relationship between rainfall and landslides, in: 4th Int. Symp. Landslides, BiTech Publishers, Toronto, 1984: pp. 377–384.

[134] P.K. Finlay, P.J., Fell, R., Maguire, The relationship between the probability of landslide occurrence and rainfall, Can. Geotech. J. 34 (1997) 811–824.

[135] C.F.. Dai, F.C., Lee, Frequency–volume relation and prediction of rainfall-induced landslides., Eng. Geol. 59 (2001) 253–266.

[136] M. Marjanović, M. Krautblatter, B. Abolmasov, U. Đurić, C. Sandić, V. Nikolić, The rainfall-induced landsliding in Western Serbia: A temporal prediction approach using Decision Tree technique, Eng. Geol. 232 (2018) 147–159. https://doi.org/10.1016/j.enggeo.2017.11.021.

[137] H. Hong, H. Shahabi, A. Shirzadi, W. Chen, K. Chapi, B. Bin Ahmad, M.S. Roodposhti, A. Yari Hesar, Y. Tian, D. Tien Bui, Landslide susceptibility assessment at the Wuning area, China: a comparison between multi-criteria decision making, bivariate statistical and machine learning methods, Nat. Hazards. 96 (2019) 173–212. https://doi.org/10.1007/s11069-018-3536-0.

[138] X. Sun, J. Chen, Y. Bao, X. Han, J. Zhan, W. Peng, Landslide susceptibility mapping using logistic regression analysis along the Jinsha river and its tributaries close to Derong and Deqin County, southwestern China, ISPRS Int. J. Geo-Information. 7 (2018) 1–29. https://doi.org/10.3390/ijgi7110438.

[139] K. Martinović, K. Gavin, C. Reale, C. Mangan, Rainfall thresholds as a landslide indicator for engineered slopes on the Irish Rail network, Geomorphology. 306 (2018) 40–50. https://doi.org/10.1016/j.geomorph.2018.01.006.

[140] S. Peruccacci, M.T. Brunetti, S.L. Gariano, M. Melillo, M. Rossi, F. Guzzetti, Rainfall thresholds for possible landslide occurrence in Italy, Geomorphology. 290 (2017) 39–57. https://doi.org/10.1016/j.geomorph.2017.03.031.

[141] D.P. Kanungo, S. Sharma, Rainfall thresholds for prediction of shallow landslides around Chamoli-Joshimath region, Garhwal Himalayas, India, Landslides. 11 (2014) 629–638. https://doi.org/10.1007/s10346-013-0438-9.

[142] M. Kuradusenge, S. Kumaran, M. Zennaro, Rainfall-induced landslide prediction using machine learning models: The case of ngororero district, rwanda, Int. J. Environ. Res. Public Health. 17 (2020) 1–20. https://doi.org/10.3390/ijerph17114147.

[143] L.M. I, L.S. I, A. Terrone, Physical and numerical modelling of shallow landslides, Landslides. (2015). https://doi.org/10.1007/s10346-015-0642-x.

[144] M. Ahmadi-adli, N.K. Toker, N. Huvaj, Prediction of seepage and slope stability in a flume test and an experimental field case, Procedia Earth Planet. Sci. 9 (2014) 189–194. https://doi.org/10.1016/j.proeps.2014.06.022.

[145] Y. Chen, Soil – Water Retention Curves Derived as a Function of Soil Dry Density, c (2018) 3–19. https://doi.org/10.3390/geohazards1010002.

[146] A. Saha, S. Sekharan, U. Manna, Evaluation of Capacitance Sensor for Suction Measurement in Silty Clay Loam, Geotech. Geol. Eng. 3 (2020). https://doi.org/10.1007/s10706-020-01297-3.

[147] S. Chersich, A Simplified Approach to Assess the Soil Saturation Degree and Stability of a Representative Slope Affected by Shallow Landslides in Oltrep ò, (2018). https://doi.org/10.3390/geosciences8120472.

[148] A. Farouk, L. Lamboj, J. Kos, A Numerical Model to Predict Matric Suction Inside

Unsaturated Soils, A Numer. Model to Predict Matric Suction Insid. Unsaturated Soils. 44 (2004). https://doi.org/10.14311/588.

[149] J. Ho, K.T. Lee, Performance evaluation of a physically based model for shallow landslide prediction, Landslides. (2016). https://doi.org/10.1007/s10346-016-0762-y.

[150] M.T. Abraham, D. Pothuraju, N. Satyam, Rainfall thresholds for prediction of landslides in Idukki, India: An empirical approach, Water (Switzerland). 11 (2019). https://doi.org/10.3390/w11102113.

[151] A. Rosi, T. Peternel, M. Jemec-Auflič, M. Komac, S. Segoni, N. Casagli, Rainfall thresholds for rainfall-induced landslides in Slovenia, Landslides. 13 (2016) 1571–1577. https://doi.org/10.1007/s10346-016-0733-3.

[152] G. Martelloni, S. Segoni, R. Fanti, F. Catani, Rainfall thresholds for the forecasting of landslide occurrence at regional scale, Landslides. 9 (2012) 485–495. https://doi.org/10.1007/s10346-011-0308-2.

[153] M.T. Abraham, N. Satyam, S. Kushal, A. Rosi, B. Pradhan, S. Segoni, Rainfall threshold estimation and landslide forecasting for Kalimpong, India using SIGMA model, Water (Switzerland). 12 (2020). https://doi.org/10.3390/W12041195.

[154] G. Jordanova, S.L. Gariano, M. Melillo, S. Peruccacci, M.T. Brunetti, M.J. Auflič, Determination of empirical rainfall thresholds for shallow landslides in slovenia using an automatic tool, Water (Switzerland). 12 (2020) 1–15. https://doi.org/10.3390/w12051449.

[155] L. Piciullo, S.L. Gariano, M. Melillo, M.T. Brunetti, S. Peruccacci, F. Guzzetti, M. Calvello, Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides, Landslides. 14 (2017) 995–1008. https://doi.org/10.1007/s10346-016-0750-2.

[156] M. Melillo, M.T. Brunetti, S. Peruccacci, S.L. Gariano, A. Roccati, F. Guzzetti, A tool for the automatic calculation of rainfall thresholds for landslide occurrence, Environ. Model. Softw. 105 (2018) 230–243. https://doi.org/10.1016/j.envsoft.2018.03.024.

[157] Y. Zhao, Y. Li, L. Zhang, Q. Wang, Groundwater level prediction of landslide based on classification and regression tree, Geod. Geodyn. 7 (2016) 348–355. https://doi.org/10.1016/j.geog.2016.07.005.

[158] J. Uwihirwe, M. Hrachowitz, T.A. Bogaard, Landslide precipitation thresholds in Rwanda, (2020). https://doi.org/10.1007/s10346-020-01457-9.

[159] USGS, Landslide Types and Processes, Highw. Res. Board Spec. Rep. (2004) 1–4. https://doi.org/Fact Sheet 2004-3072.

[160] E. Monsieurs, O. Dewitte, A. Demoulin, A susceptibility-based rainfall threshold approach for landslide occurrence, Nat. Hazards Earth Syst. Sci. Discuss. (2018) 1–25. https://doi.org/10.5194/nhess-2018-316.

[161] A. Terzis, A. Anandarajah, K. Moore, I.-J. Wang, Slip surface localization in wireless sensor networks for landslide prediction, Proc. Fifth Int. Conf. Inf. Process. Sens. Networks - IPSN '06. (2006) 109. https://doi.org/10.1145/1127777.1127797.

[162] A. Dikshit, N. Satyam, B. Pradhan, S. Kushal, Estimating rainfall threshold and temporal probability for landslide occurrences in Darjeeling Himalayas, Geosci. J. 24 (2020) 225–233. https://doi.org/10.1007/s12303-020-0001-3.

[163] L. Schilirò, G.P. Djueyep, C. Esposito, G.S. Mugnozza, The Role of Initial Soil Conditions in Shallow Landslide Triggering : Insights from Physically Based Approaches, 2019 (2019).

[164] T.W.J. Van Asch, B. Yu, W. Hu, The Development of a 1-D Integrated Hydro-Mechanical Model Based on Flume Tests to Unravel Different Hydrological Triggering Processes of Debris Flows, (2018). https://doi.org/10.3390/w10070950.

[165] Republic of Rwanda Gakenke District Environmental and Social Management Plan (ESMP) Quality Basic Education for Human Capital Development (QBE), 2019.

[166] D. Karemera, Landslide victims laid to rest | The New Times | Rwanda, (2016). https://www.newtimes.co.rw/section/read/199730 (accessed June 15, 2020).

[167] R. Umurengezi, Gakenke: Family of eight that perished in landslide buried, (2020). https://www.newtimes.co.rw/news/gakenke-family-eight-perished-landslide-buried (accessed June 15, 2020).

[168] TRTWORLD, At least 65 killed in flooding, landslides in Rwanda, (2020). https://www.trtworld.com/africa/at-least-65-killed-in-flooding-landslides-in-rwanda-36116 (accessed June 15, 2020).

[169] G.G. Benineza, I. Rwabudandi, M.J. Nyiransabimana, Landslides hazards assessment using geographic information system and remote sensing: Gakenke District, IOP Conf. Ser. Earth Environ. Sci. 389 (2019). https://doi.org/10.1088/1755-1315/389/1/012015.

[170] M.J. Claude, N.V. Martin, M. Abias, M. Francoise, U. Johnson, K. Tonny, U. Martine, Mapping Landslide Susceptibility and Analyzing Its Impact on Community Livelihoods in Gakenke District, Northern Rwanda, J. Geosci. Environ. Prot. 08 (2020) 41–55. https://doi.org/10.4236/gep.2020.85003.

[171] G.T. Harilal, D. Madhu, M.V. Ramesh, D. Pullarkatt, Towards establishing rainfall thresholds for a real-time landslide early warning system in Sikkim, India, Landslides. 16 (2019) 2395–2408. https://doi.org/10.1007/s10346-019-01244-1.

[172] R. Sehler, J. Li, H. Ye, Investigating Relationship Between Soil Moisture and Precipitation Globally Using Remote Sensing Observations, (2019) 106–118. https://doi.org/10.1111/j.1936-704X.2019.03324.x.

[173] Y. Zhang, J. Liu, X. Xu, Y. Tian, Y. Li, Q. Gao, The response of soil moisture content to rainfall events in semi-arid area of Inner Mongolia, 2 (2010) 1970–1978. https://doi.org/10.1016/j.proenv.2010.10.211.

[174] M. Fereidoon, M. Koch, L. Brocca, Predicting rainfall and runoff through satellite soil moisture data and SWAT modelling for a poorly gauged basin in Iran, Water (Switzerland). 11 (2019). https://doi.org/10.3390/w11030594.

[175] B. Zhao, Q. Dai, D. Han, H. Dai, J. Mao, L. Zhuo, G. Rong, Estimation of soil moisture using modified antecedent precipitation index with application in landslide predictions, Landslides. 16 (2019) 2381–2393. https://doi.org/10.1007/s10346-019-01255-y.

[176] A. Wicki, P. Lehmann, C. Hauck, S.I. Seneviratne, P. Waldner, M. Stähli, Assessing the potential of soil moisture measurements for regional landslide early warning, Landslides. 17 (2020) 1881–1896. https://doi.org/10.1007/s10346-020-01400-y.

[177] G. Crosta, Regionalization of rainfall thresholds: An aid to landslide hazard evaluation, Environ. Geol. 35 (1998) 131–145. https://doi.org/10.1007/s002540050300.

[178] F.H. Ali, N. Osman, Shear strength of a soil containing vegetation roots, Soils Found. 48 (2008) 587–596. https://doi.org/10.3208/sandf.48.587.

[179] F.H. Ali, N. Osman, Shear Strength of a Soil Containing Vegetation Roots, Soils Found. 48 (2008) 587–596. https://doi.org/https://doi.org/10.3208/sandf.48.587.

[180] C. Maffra, R. Sousa, F. Sutili, R. Pinheiro, The Effect of Roots on the Shear Strength of Texturally Distinct Soils, Floresta e Ambient. 26 (2019). https://doi.org/10.1590/2179-8087.101817.

[181] T.H. Wu, Effect of Vegetation on Slope Stability., Transp. Res. Rec. (1984) 37–46. https://doi.org/10.1016/0148-9062(85)92272-7.

[182] Y. Li, C. Hu, L. Jian, R. Zhao, C. Li, Evaluation of the stability of vegetated slopes according to layout and temporal changes, J. Mt. Sci. 18 (2021) 275–290. https://doi.org/10.1007/s11629-020-6022-6.

[183] S. Wang, M. Zhao, X. Meng, G. Chen, R. Zeng, Q. Yang, Y. Liu, B. Wang, Evaluation of the effects of forest on slope stability and its implications for forest management: A case study of Bailong River Basin, China, Sustain. 12 (2020). https://doi.org/10.3390/su12166655.

[184] L. Piciullo, M. Calvello, J.M. Cepeda, Territorial early warning systems for rainfall-

induced landslides, Earth-Science Rev. 179 (2018) 228–247. https://doi.org/10.1016/j.earscirev.2018.02.013.

[185] A. Wicki, M. Stähli, Performance analysis of regional landslide early warning based on soil moisture simulations, (2021).

[186] UNDRR, Early warning system | UNDRR, United Nations Off. Disaster Risk Reduct. (2020). https://www.undrr.org/terminology/early-warning-system (accessed January 26, 2021).

[187] W. Meteorological, O. Wmo, Multi-hazard Early Warning Systems : A Checklist, (2017).

[188] T.F. Fathani, D. Karnawati, W. Wilopo, An integrated methodology to develop a standard for landslide early warning systems, (2016) 2123–2135. https://doi.org/10.5194/nhess-16-2123-2016.

[189] S. Lacasse, Event tree analysis of Aknes rock slide hazard, 4th Can. Conf. Geohazards From Causes to Manag. 2 (2008) 551–558.

[190] B.N. Silva, M. Khan, K. Han, Internet of Things: A Comprehensive Review of Enabling Technologies, Architecture, and Challenges, IETE Tech. Rev. 35 (2018) 205–220. https://doi.org/10.1080/02564602.2016.1276416.

[191] D. Giordan, A. Wrzesniak, P. Allasia, The Importance of a Dedicated Monitoring Solution and Communication Strategy for an Effective Management of Complex Active Landslides in Urbanized Areas, (2019). https://doi.org/10.3390/su11040946.

[192] S. Segoni, L. Piciullo, S.L. Gariano, Preface: Landslide early warning systems: Monitoring systems, rainfall thresholds, warning models, performance evaluation and risk perception, Nat. Hazards Earth Syst. Sci. 18 (2018) 3179–3186. https://doi.org/10.5194/nhess-18-3179-2018.

[193] F. Nadim, E. Intrieri, Early Warning Systems for Landslides : Challenges and New Monitoring Technologies, (2011) 1–15.

[194] T.F. Fathani, D. Karnawati, W. Wilopo, An integrated methodology to develop a standard for landslide early warning systems, Nat. Hazards Earth Syst. Sci. 16 (2016) 2123–2135. https://doi.org/10.5194/nhess-16-2123-2016.

[195] H. Mittelbach, S.I. Seneviratne, A new perspective on the spatio-temporal variability of soil moisture: Temporal dynamics versus time-invariant contributions, Hydrol. Earth Syst. Sci. 16 (2012) 2169–2179. https://doi.org/10.5194/hess-16-2169-2012.

[196] A. Wicki, P. Lehmann, C. Hauck, S.I. Seneviratne, P. Waldner, M. Stähli, Assessing the potential of soil moisture measurements for regional landslide early warning, 2015

(2020) 1881–1896. https://doi.org/10.1007/s10346-020-01400-y.

[197] R. Fensholt, A.I. Abdelkerim, M.M.R.S. Eusuf, Study of soil moisture sensor for landslide early warning system : Experiment in laboratory scale Study of soil moisture sensor for landslide early warning system : Experiment in laboratory scale, (2016). https://doi.org/10.1088/1742-6596/739/1/012034.

[198] M. Kuradusenge, Rainfall and soil moisture data, (2021). https://aceiot.ur.ac.rw/?Martin-data.

[199] M. Kuradusenge, S. Kumaran, M. Zennaro, J.B. Minani, Risks Reduction of Rainfall-Induced Landslides-A Site-Specific Early Warning System (SSEWS) BT - ICT Systems and Sustainability, in: M. Tuba, S. Akashe, A. Joshi (Eds.), Springer Nature Singapore, Singapore, 2022: pp. 873–881.

List of Publications

1. M. Kuradusenge, S. Kumaran, M. Zennaro, J.B. Minani, Risks Reduction of Rainfall-Induced Landslides-A Site-Specific Early Warning System (SSEWS) BT - ICT Systems and Sustainability, in: M. Tuba, S. Akashe, A. Joshi (Eds.), Springer Nature Singapore, Singapore, 2022: pp. 873–881..

2. Martin Kuradusenge, Santhi Kumaran, Marco Zennaro, and Albert Niyonzima, *"Experimental Study of Site-Specific Soil Water Content and Rainfall Inducing Shallow Landslides: Case of Gakenke District, Rwanda,"* Geofluids, Wiley Hindawi, vol. 2021, p. 7194988, 2021, doi: 10.1155/2021/7194988.

3. M. Kuradusenge, S. Kumaran, and M. Zennaro, "*Rainfall-induced landslide prediction using machine learning models: the case of Ngororero District, Rwanda*," International Journal of Environmental Research and Public Health, MDPI, vol. 17, no. 11, p. 4147, 2020, doi: 10.3390/ijerph17114147.