Pan African University
Institute of Water and Energy Sciences

PAN-AFRICAN UNIVERSITY
INSTITUTE FOR WATER AND ENERGY SCIENCES
(including CLIMATE CHANGE)

Master Dissertation

Submitted in partial fulfilment of the requirements for the Master degree in

WATER ENGINEERING

Presented by

Benjamin BONKOUNGOU

FLOOD SUSCEPTIBILITY ASSESSMENT IN THE NAKAMBE BASIN, BURKINA FASO

Defended on 18/11/*2021* before the following commitees

Chair	Guemou Bouabdellah	Dr.	University of Ain Temouchent
Supervisor	Amos T. Kabo-bah	Prof.	UENR
External Examiner	Ashraf Mostapha	Prof.	Ain Chems University
Internal Examiner	Hamouda Boutaghane	Dr.	University of Annaba

TECHNICAL SHEET

Technical sheet	
Topic: Flood Susceptibility Assessment in the Nakambe Basin, Burkina Faso	
Author: Benjamin Bonkoungou	
Supervisor: Prof. Amos T. Kabo-bah	
Internship duration: July 1st, 2021 – September 30th, 2021	
Internship organization: WASCAL (West African Science Service Centre on Climate Change and Adapted Land Use)	
Internship Supervisor: : Dr. Kwame O. Hackman	
Research purpose: This study has been undertaken in order to develop a susceptibility map of flood hazards in the Nakambé basin, Burkina Faso, and recommend some mitigation and/or adaptation measures to reduce the impact on people and the environment.	
Software used	
R Python Microsoft Excel & Word	Google Earth Engine ArcGIS QGIS

STATEMENT OF THE AUTHOR

By my signature below, I declare that this thesis is my work. I have followed all ethical principles of scholarship in the preparation, data collection, data analysis, and completion of this thesis or dissertation. I have given all scholarly matter recognition through accurate citations and references. I affirm that I have cited and referenced all sources used in this document. I have made every effort to avoid plagiarism.

I submit this document in partial fulfillment of the requirement for a degree from Pan African University. This document is available from the PAU Library to borrowers under the rules of the library. I declare that I have not submitted this document to any other institution for the award of an academic degree, diploma, or certificate.

Scholars may use brief quotations from this thesis or dissertation without special permission if they make an accurate and complete acknowledgment of the source. The dean of the academic unit may grant permission for extended quotations or reproduction of this document. In all other instances, however, the author must grant permission.

Name: Benjamin Bonkoungou Signature: 

Date: 20/10/2021

Academic Unit: Water Engineering

PAU Institute: Institute of Water and Energy Science including Climate Change (PAUWES)

CERTIFICATION

This thesis has been submitted with my approval as the supervisor. It constitues the final version, and all corrections were added as recommended by the examination committee.

Signed  Date: November 20th, 2021

Prof. Amos T. Kabo-bah
Senior Lecturer/Head of department
Civil and Environmental Engineering
University of Energy and Natural Resources (UENR)

ACKNOWLEDGMENT

I gratefully acknowledge the African Union for the scholarship awards, the Director of Pan African University Institute of Water and Energy Sciences including Climate Change (PAUWES), and his team without forgetting the guest lecturers.

I sincerely thank my supervisor Prof. Amos T. Kabo-bah for his constant support, time, and guidance. The completion of this thesis work on this short time was never expected without his support.

I also want to thank all people and organizations who provided data for their contribution to this research. My sincere appreciation goes to WASCAL and particularly to Dr. Kwame O. Hackman for the assistance during my internship and throughout my data collection.

Last but not least, my grateful recognitions are passing to my family, my friends, and everyone I failed to mention here for all the encouragement and thoughtful advice. I am grateful.

DEDICATION

To all the internally displaced people because of natural disasters, insercurity, etc! Hold fast, however long and dark the night may be, there always comes an hour when at last the day breaks.

To our brave armed forces, the Defense and Security Forces (FDS) and the Volunteers for the Defense of the Homeland (VDP)! We sleep safely at night because you stand ready to visit violence on those who would harm us. You, fallen heroes, represent the character of our nation who has a long history of integrity, patriotism, and honor. There is no duty more urgent than that of returning thanks; we say thank you.

To my late uncle, Bonkoungou W. Pierre! You have been more than a father to us. Thank you for all the values you inculcated in us.

ABBREVIATIONS AND ACRONYMS

ANAM	National Agency of Meteorology of Burkina Faso
ArcGIS	Aeronautical Reconnaissance Coverage Geographic Information System
AUC	Area Under the Curve
BUMIGEB	National Soil Bureau of Burkina Faso
BUNASOLS	National Soil Bureau of Burkina Faso
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station
CRED	Centre for Research on the Epidemiology of Disasters
DEM	Digital Elevation Model
EM-DAT	Emergency Events Database
FR	Frequency Ratio
GEE	Google Earth Engine
GIRE	Integrated Management of Water Resources
GIS	Geographic Information System
LR	Logistic Regression
MERIT	Multi-Error-Removed-Improved-Terrain
LULC	Land Use and Land Cover
NDVI	Normalized Differential Vegetation Index
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
PERSIANN-CDR	Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks-Climate Data Record
QGIS	Quantum Geographic Information System
ROC	Receiver Operating characteristic Curve
SPI	Stream Power Index
TWI	Topographic Wetness Index
UNDRR	United Nations Office for Disaster Risk Reduction
UN-SPIDER	United Nations Platform for Space-based Information for Disaster Management and Emergency Response
WMO	World Meteorological Organization

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Flood is the most disastrous natural disaster in the world due to its devastating effects that endanger lives and cause property damage to the affected areas. Flood events are alarmingly increasing all over the world; and despite the recent efforts towards mitigation and management, vulnerability to flood damages is likely to continue to grow. This research mainly aims at identifying areas susceptible to flood through the application of two different models namely Frequency Ratio (FR) and Logistic Regression (LR) using remote sensing data (RS) and Geographical Information System (GIS). The study was conducted in the Nakambé Basin, Burkina Faso, and consisted of the use of thirteen predicting factors and 250 historical flood locations. The flood points were randomly split into 70% training and 30% validation through the Area Under Curve (AUC) method. The results indicated that the LR model performed best for the susceptibility mapping in the study area disclosing 93.1% of prediction rate compared to the FR model (80.6%). Moreover, the slope and elevation were found to be the most contributing factors in the study area. The outcome of this study can be used by the planners and policy makers to implement different management and mitigation measures with a view to the local environment.

Keywords: Flood, remote sensing, susceptibility mapping, logistic regression, frequency ratio

RÉSUMÉ

Les inondations sont les aléas naturels les plus catastrophiques au monde en raison de leurs effets dévastateurs qui mettent des vies en danger et causent des dommages matériels aux zones touchées. Elles augmentent de manière alarmante partout dans le monde ; et malgré les récents efforts d'atténuation et de gestion, la vulnérabilité aux dommages causés par ces inondations devrait continuer de croître. La présente recherche vise principalement à identifier les zones à risque d'inondation grâce à l'application de deux modèles différents à savoir le rapport de fréquence (FR) et la régression logistique (LR) à l'aide de données de télédétection (RS) et de système d'information géographique (SIG). L'étude a été menée dans le bassin du Nakambé, au Burkina Faso, et a consisté à l'utilisation de treize principaux facteurs de prédisposition dans l'occurrence spatiale des inondations et de 250 points historiques d'inondation. Les points ont été répartis aléatoirement en 70% pour l'apprentissage et 30% pour la validation par la méthode de l'aire sous la courbe (AUC). Les résultats indiquent que le modèle LR est le plus performant pour la cartographie de la susceptibilité dans la zone d'étude, avec un taux de prédiction de 93,1 %, comparé au modèle FR (80,6 %). De plus, la pente et l'élévation se sont avérées être les variables prédictives les plus contributifs dans la zone d'étude. Les résultats de cette étude peuvent être utilisés par les planificateurs et les décideurs politiques pour mettre en œuvre différentes mesures de gestion et d'atténuation en tenant compte de l'environnement local.

Mots clés : Inondation, Télédétection, cartographie de susceptibilité, régression logistique, rapport de fréquence

CHAPTER 1

“Water is the cause at times of life or death, or increase of privation, nourishes at times and at others does the contrary; at times has a tang, at times is without savour, sometimes submerging the valleys with great floods. In time and with water, everything changes.”

Leonardo da Vinci, circa 1500

1 INTRODUCTION

1.1 Background

Natural disasters are catastrophic events brought about by natural processes of the Earth resulting in widespread destruction to the environment and also loss of life. Ranging from floods, droughts, and hurricanes to tsunamis, volcanoes, and earthquakes, etc, these events are responsible for the death of over 60,000 people per year globally (Ritchie & Roser, 2014). Natural disasters have an alarmingly increasing trend worldwide. Over the past decade only, more than 300 catastrophes have been recorded around the globe affecting millions of people and cost billions (Prasad & Francescutti, 2016). However, the total number of death by natural disasters, as shown in Figure 1.1 is highly variable from year to year, some years pass with very few deaths before a large disaster event claims many lives. In some years, the death is low, often less than 10,000, but some years witnessed devastating events, inter alia, the 2004 Indian Ocean earthquake, and Tsunami, the 2008 Cyclone Nargis in Myanmar, the 2010 Port-au-Prince earthquake in Haiti. These events pushed global disasters deaths to over 200,000 – more than 0.4% of deaths in these years.

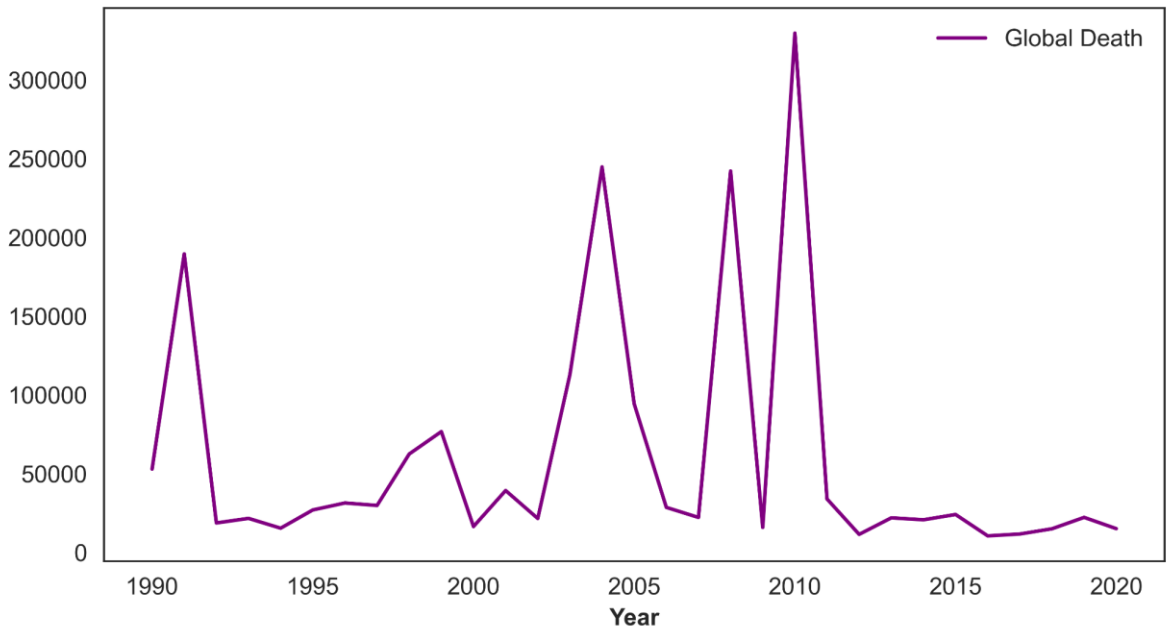

Figure 1.1. Global death from natural disasters, 1990 – 2020 (Source: based on EM-DAT/CRED data)
The data include all categories classified as "natural disasters". This includes drought, floods, extreme weather, extreme temperature, landslides, dry mass movements, wildfires, volcanic activity, and earthquakes.

In terms of occurrence, flooding is the most frequent among all-natural disasters. Figure 1.2. illustrates the general trend of flood occurrence all over the world over the past seven century. A significant increase can be observed, particularly in the last 50 years. This increase can also be observed in the number of people affected by floods and financial, economic, and insured damages.

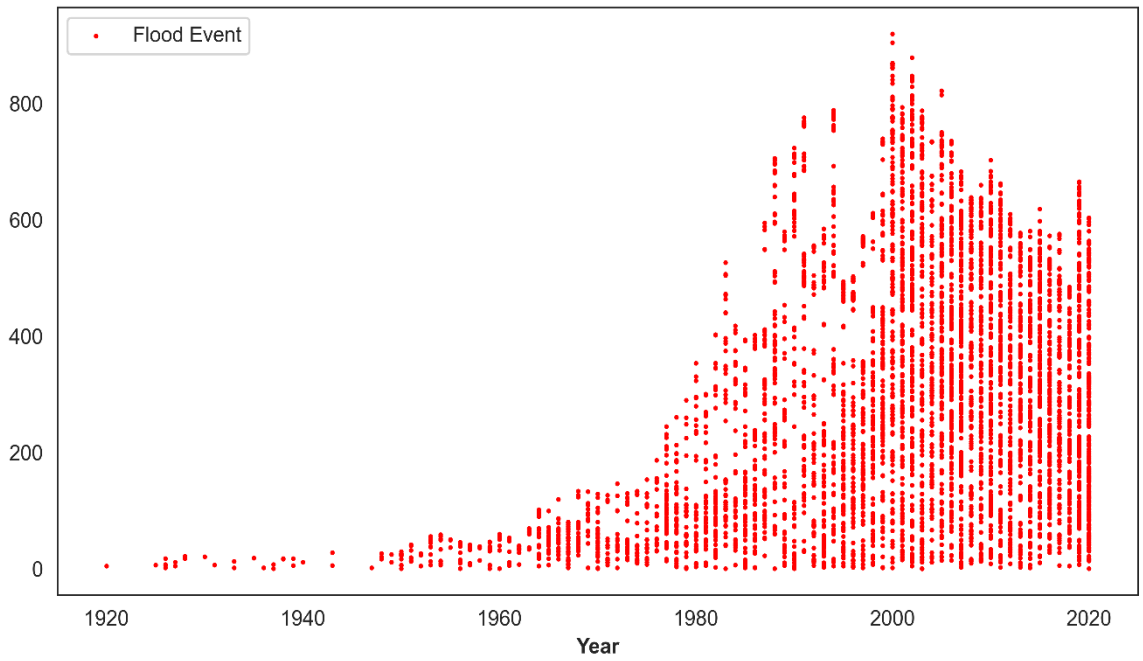

Figure 1.2. Number of reported flood events (Source: based on EM-DAT/CRED data)

Flooding is the most destructive type of natural disaster due to its devastating effects that endanger lives and cause property damage in the affected areas. The frequency and impacts (economic loss, human loss) of floods make them more hazardous than other natural disasters. According to the Emergency Events Database (EM-DAT), in 2020 alone, 33.2 million people were affected by floods, and the total damages were estimated at $51 million. Although the least developed nations suffer more from these because of their incapacity to effectively protect themselves, developed and industrialized countries also devastatingly experience them. In the last decade, there has been catastrophic flooding experienced in China, India, Bangladesh, Germany, Poland, Mozambique, the USA, and elsewhere (UNISDR, 2002). When floods occur, especially in developing countries, they have the potential of destroying decades of investments in infrastructure, thereby crippling the economic prosperity of the country and resulting in thousands of deaths, displaced, and

epidemics. The majority of these deaths are generally found within the most vulnerable member of the society, namely women and children. Yet, this tragedy can be avoided or significantly reduced through proper investment before during, and after the event in preparedness activities and associated infrastructure, flood plain policy development, effective watershed land use planning, flood forecasting, and warning systems, and response mechanisms. Figure 1.3 illustrates a success in the implementation of these management measures as a slowly decreasing trend can be observed in the total number of people affected in the last decade. This trend is encouraging but the fatalities still remain high in developing countries because of the disproportionate impact on the poor and the socially disadvantaged (Jha et al., 2012).

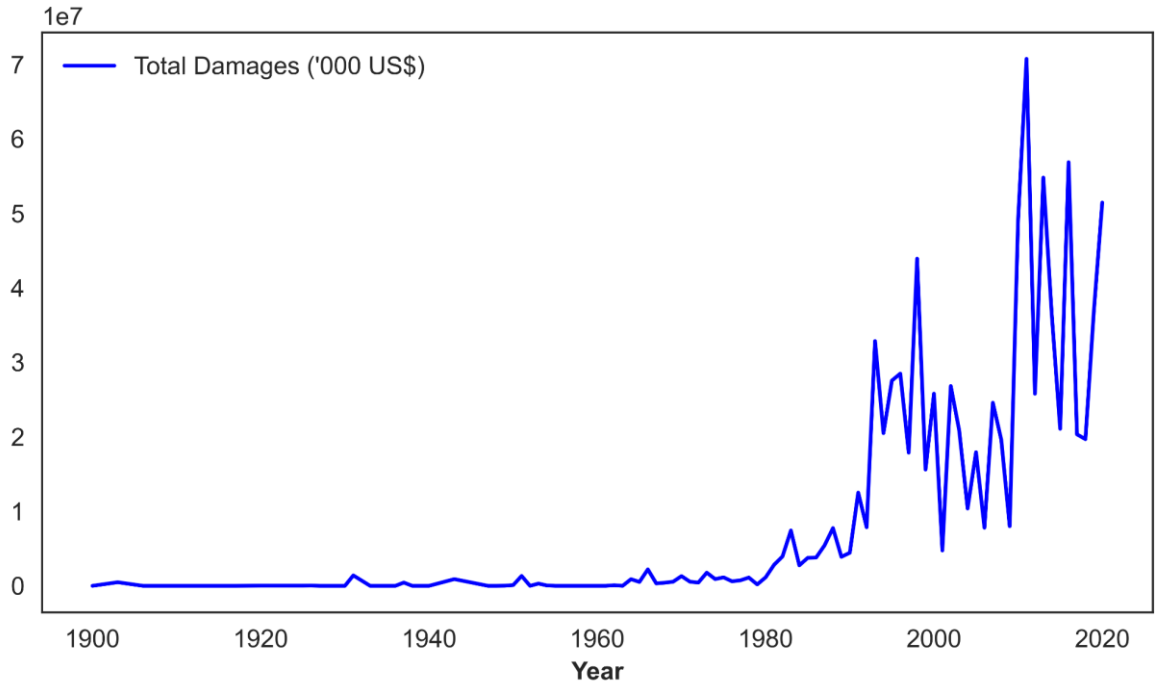

Figure 1.3. Total damages of flood events (Source: based on EM-DAT/CRED data)

1.2 Problem Statement

Flood events are alarmingly increasing all over the world; and despite the recent efforts towards mitigation and management, vulnerability to flood damages is likely to continue to grow (Şen, 2017). This trend can be explained by numerous factors such as the growing populations in and near floodplains, the loss of flood moderating wetlands, the increased runoff from paving over the soil, climate change, and inconvenient policies such as subsidies that encourage development in flood plains.

In the context of Burkina Faso, there have been recurrent flood events in the last decade. According to the international disasters database, since 2006 there have been flooding events recorded every year. Figure 1.4 shows an area inundated during the September 2009's flooding.

Figure 1.4. A floodplain area in Kongoussi submerged by flood, September 2009 (Source: SP/CONASUR)

Most recently in April 2020, all 13 regions of the country were affected by flooding with multiple properties destroyed and some loss of lives (Bakouan & Knutson, 2020). In fact, according to the United Nations Office for the Coordination of Humanitarian Affairs (OCHA), the economic loss can be summarized as followed: destroyed residential houses (3347), damaged residential houses (1656) and destroyed or damaged Internally Displaced Persons (IDP) emergency shelters (1790). They also noted the loss of animals (cattle, goats, sheep, asin, poultry) and more than 250 tons of food destroyed or washed away. Table 1.1

summarizes the total number of people affected by region. The northern center region (study area) and the Sahel were the most affected accounting for around 31% and 33% of the total damage respectively. In response, the government declared a state of natural disaster and mobilized the fund to tackle the issue. The Office reported that more than €13 million was needed as an investment.

Table 1.1. The total number of victims of the 2020 flooding by region.

Region	Men	Women	Children	Total	Percent (%)
Cascades	38	41	115	194	0.27
Centre	543	588	1630	2761	3.87
Centre-Nord	4256	4609	12756	21621	30.31
Centre-West	147	166	466	779	1.09
Centre-South	449	490	1349	2288	3.21
East	1267	1370	3797	6434	9.02
Hauts-bassins	187	159	401	747	1.05
Nord	41	83	147	271	0.38
Plateau central	1289	1395	3859	6543	9.17
Sahel	2385	13540	8049	23974	33.60
Sud-Ouest	94	101	280	475	0.67
Boucle du Mouhoun	874	941	2539	4354	6.10
Centre-Est	177	192	531	900	1.26
Total	11747	23675	35919	71341	100

Clearly, flooding can be considered as one of the most significant natural hazards in the country in terms of the number of affected populations, fatalities, and economic damage. A deeper understanding of flood hazards is essential for risk management and mitigation. Risk of flooding estimation and flood control are critical for identifying flood-prone areas in order to mitigate flood destruction by taking the appropriate steps, especially in the context of Burkina Faso where very few studies have been conducted. The available flood maps lack spatial variability while global datasets have resolutions too coarse to be relevant for local-scale risk assessment. Consequently, local disaster managers very often use traditional methods such as watermarks on buildings and media reports to identify flood hazard areas (Asare-Kyei et al., 2015). A country-level risk profile for Burkina Faso was developed in the context of the Integrated Management of Water Resources (GIRE) program in 2001 by the Ministry of Environment (MECV, 2007), but since then the country has witnessed

increasing severe extreme events. There is a need to continually update the information for better management, mitigation, and adaptation strategies.

1.3 Research Questions and Objectives

The general objective of this research is to develop a model-based methodological framework allowing for flood mapping in the Nakambé Watershed, Central Burkina Faso; and thus, assisting public administrators in making appropriate decisions. More specifically the following research questions will be addressed:

1. How accurately can the selected models predict flood susceptibility?
2. What are the high-risk zones and what are the possible reasons?
3. What flood management and mitigation practices can be adopted for the area?

This research thus aims to take a deeper look into the correlation between the flood hazard and the spatial distribution of some predisposing factors such as the geological, geomorphological, and hydrogeological features and soil/land cover characteristics. Particularly, the study has the following sub-objectives:

1. To identify the most important contributing factors for the study area
2. To test the flood susceptibility models' performance and analyze the model outputs
3. To identify suitable practices for risk management

This research apart from contributing to the scientific field will allow decision-makers to take an additional step towards risk mitigation and adaptation. For the land-use planners, it can help in the identification of the location of flood-prone areas. This information can be used to prescribe or advise building codes. The flood managers can use the maps to identify areas with high potential, in order to prioritize their protective efforts in case of an eventual flood event. It will also help the communities to know whether they live in a flood-prone area and what their options are in case of flooding.

1.4 Dissertation Structure

The dissertation is organized into five (05) chapters. The introductory chapter, chapter 1, is an overview of the research. It provides a solid background of the study, a concise explanation of the research objectives, questions, and relevance. Through an exhaustive literature review, chapter 2 covers the basic concepts of flood studies, flood risk modeling

and mapping, the use of Geographic Information System (GIS), remote sensing, and statistical tools for flood risk mapping or profiling. Chapter 3 highlights a detailed description of the study area, the Nakambé basin, Burkina Faso. It also describes the data used in this research and the different data analysis methods used. Chapter 4 addresses the first two research questions through a presentation, explanation, and discussion of the main findings. The final chapter, chapter 5, summarizes the conclusions and presents some recommendations for risk mitigation and adaptation in the study area. Lastly, chapters six and seven provide the reference and annexes respectively.

CHAPTER 2

"The main problem in Old Zongo and Abaase areas is the gutters. The gutters are not enough to carry the water when it rains heavily, and secondly, they pour so much rubbish in the gutters. So, when it rains heavily, where will the water go, it must flood the area. ….the way we build in this area too is a problem, that is why we are always trying to tell people here not to build in the waterway…"

A Resident in Asamankese, Eastern Region of Ghana (Jerome Glago, 2021)

2 LITERATURE REVIEW

2.1 Floods

2.1.1 Definition

The United Nations Platform for Space-based Information for Disaster Management and Emergency Response (UN-SPIDER) simply defines flooding as the presence of water where it is not wanted. A more comprehensive definition by hydrologists considers flooding as a sudden increase in water discharge that caused a sudden peak in the water level. It is thus a temporary condition of partial or complete inundation of normally dry land areas from overflow of inland or tidal waters from the unusual and rapid accumulation or runoff of surface waters from any source. In other terms, when a flood occurs, the water rises to overflow land that is not normally submerged (Di Baldassarre, 2010). This process is generally the consequence of many factors (and combinations thereof) such as heavy rain, ice melting, dam breaks, groundwater rises, etc. Although generally seen as a disaster due to its devastative impacts on the communities and properties, floods can also have positive impacts. They may provide relief to the people and ecosystems suffering from prolonged drought.

2.1.2 Causes of floods

Although defined as a natural disaster, the causes of flooding are not limited to natural processes. They are manifold but Knight et al. (2005) in their study broadly categorized these various causes into two headings namely natural causes and man-made causes.
The natural causes can be summarized as follows:

- Precipitation: extreme precipitation (rainfall, hail, and snowmelt) is one of the most common causes of flooding.
- Landslides: often due to slope instability, erosion, or seismic activity, the movement of an important mass of soil can lead to an overflow of water
- Storm surge: low pressure in the sea often raises tidal levels, higher than normal river levels. Storm surges are also associated with tropical cyclones, tsunamis, etc.

- High groundwater level: this speed up the runoff process)
- Climate change and ice melting: the rise in temperature affects the precipitation and the sea level. According to a recent report by the World Meteorological Organization (WMO), the actual global mean sea level is faster than that of any other time in the past 3,000 years, i.e., approximately 4.4 mm per year from 2000 to 2019 (Nerem et al., 2018; WMO, 2015) This acceleration in the sea level rise is predominantly due to anthropogenic forcing since 1970.

On the other hand, man-made causes are due to anthropogenic activities and can be summarized as follows:

- Dam failures due to catastrophic, overtopping, piping, etc.
- River and coastal flood defense embankment failures
- Floodplain encroachment: building on floodplains leads to loss of storage
- Change of land use: crop change, compaction of soil, deforestation, etc
- Inadequate planning controls within the whole catchment area
- Inadequate drainage capacity due to urbanization

2.1.3 Types of floods

According to their speed, geography, and causes of flooding, floods can be categorized into several types Each type has a unique impact in terms of how it happens, how it is forecast, the damages it can cause, and the measures of protection required. It is thus important to learn about the various types of floods and how they can affect properties and what measures can be taken. The literature on flood has highlighted various categories but the most common categorization is as follows:

2.1.3.1 Fluvial floods (river floods)

River floods are characterized by a gradual overflowing of riverbanks induced by heavy rain over a long period of time. The rain causes the river to be filled with too much water much more than the capacity of the river channel. As explained by Ali, 2018, river floods can be considered as an expected event because it usually occurs seasonally, generally during wet seasons; and the flooding process is generally relatively slow allowing the officials to take

measures such as evacuating the locals before the river overflows. This makes river floods rarely result in loss of lives yet they can lead to immense economic damage.

2.1.3.2 Pluvial floods (flash floods)

Contrary to river floods, flash floods occur suddenly without or with very little warning. The most common cause of pluvial floods is extreme rainfall. They occur when a large amount of water floods within a short period of time. They can therefore happen in any location, urban or rural; even in areas with no water bodies in the vicinity and not only areas located near a body of water. Flash floods can also happen due to a sudden release of water from an upstream levee or a dam. Because, they are mainly driven by an intense and high-velocity torrent of water, flash floods are very dangerous and destructive. In fact, the rapid water torrents can move large objects such as cars, rocks, and trees.

2.1.3.3 Coastal floods (storm surge)

In a simple definition, a coastal flood is when the coast is flooded by the sea. Coastal floods are caused by strong winds or storms that move towards the coast during high tide (storm surge) and tsunamis. This type of flood is often the greatest threat associated with windstorms. Coastal floods' intensity is affected by a number of elements inter alia the windstorm's strength, magnitude, speed, and direction. The geography of the onshore and offshore areas is also essential. Coastal flood models use this information, as well as data from previous storms in the region, to predict the probability and severity of a storm surge.

2.1.3.4 Urban Flood

Urban flooding is the flooding in urban areas usually caused by flash floods, or river floods, or coastal floods. The main causes of urban flooding are poorly planned urbanization, flood plains encroachment, and unsuitable land-use change. Large-scale urban expansion typically happens in the form of unplanned construction in floodplains, both coastal and inland locations, as well as other flood-prone places, when cities and towns swell and spread outwards to accommodate population growth. The lack of natural drainage in an urban area can also contribute to flooding as the drainage system will fail to absorb the water from

heavy rains. Because of the concentration of population in the urban environment, urban floods are usually more costly and difficult to manage (Jha et al., 2012).

2.2 Flood Modelling

2.2.1 Flood risk

Flood risk as described by Di Baldassarre (2010) is typically the combination of two components as expressed in Equation 2.1. These components are the probability or the likelihood of the flood happening and the impact (consequences) if it happens.

$$\text{Risk} = \text{Probability} * \text{Consequences} \quad (2.1)$$

On one hand, the likelihood of a flooding event is often described using return periods and can thus be misleading and confusing. In fact, a 1/50 year flood does not mean that a flood of the given magnitude only happens every 50 years but in reality, it shows that a flood of that magnitude has a 2 percent probability of happening in any one year. It is thus preferable to use the probability or the chance of flooding to facilitate the understanding by the public. On the other hand, the consequences can be described as the expected environmental, economic, and social losses associated with the event. A wide definition by Sagris et al., (2005) specifies two contributing factors to the consequences, the exposure, and the vulnerability. Exposure represents the number of people or things that may be affected by the event whereas vulnerability is a measure of the potential for people or things to be harmed. An illustration of a low exposure would be an area with very few people, and low vulnerability would be an area where people can evacuate quickly and easily. In such areas, a flood event is likely to have less severe consequences compared to areas with lots of people (high exposure) who have difficulty with evacuation (high vulnerability). The risk can thus be expressed in terms of hazard, exposure, and vulnerability (Equation 2.2). In simple terms, if a hazard or an event happens where there is no risk or no exposure, there is also no risk.

$$\text{Risk} = \text{Hazard} * \text{Exposure} * \text{Vulnerability} \quad (2.2)$$

As observed by many researchers, hazard, exposure, and vulnerability all appear to be on the rise, and will certainly continue to do so in the future unless proactive measures are

implemented (Jha et al., 2012). Table 2.1 highlights the key variables that are contributing to this increase.

Table 2.1: Causes of changes in hazard, exposure, and vulnerability (Jha et al., 2012)

Hazard	**Exposure**	**Vulnerabilities**
Climate change and variability Increased rainfall Changing rainfall patterns Sea level rise Land-use change Urban expansion Decrease permeability Poor drainage Urban Microclimate Aging infrastructure Land subsidence	The larger population in urban areas Higher density Development outside defense Uncontrolled development Floodplain occupation Lack of operations and maintenance Reliance on aid Land-use change	Poverty Changing demographics Building design without regard to flood risk Lack of preparedness Over-reliance on defenses Urban density Just in time management

Understanding the risk means that the consequences can be significantly mitigated and response and recovery can be more effective and efficient in case of eventual flooding. Flood risk may be controlled by addressing both the probability principally through structural defenses and the consequences of flooding, mainly non-structural measures. These issues are explored in detail in the last chapter in the recommendations section.

2.2.2 Flood modelling

A model is a simulation of the real world. Flood inundation models are numerical tools that help in the simulation of river hydraulics and floodplain inundation processes. It uses the actual hydrological input data, hydraulic features of the basin, and boundary conditions to mimic the real-world flood occurrences. In recent years, with the advances in technology allowing a tremendous reduction in computation time, flood simulations have been widely studied and various methodologies have been developed. Even now, different types of models are available and many approaches were conclusively researched by scholars and

researchers by using various static and dynamic models (1-D, 2-D, and 3-D) (Bates, 2012). In his briefing, Pender (2006) categorized hydraulic models according to the dimensionality of the represented processes, as shown in Table 2.2. The table also provides a summary of their potential application.

Table 2.2: Summary of hydraulics models (Pender, 2006)

Method	Features	Software	Applications
0D	No physical laws	ArcGIS, Delta mapper	Broad-scale assessment of flood extents and flood depths
1D	Solution of the 1-D equation	HEC-RAS, InfoWorks RS, Mike 11	Design scale modeling, which can be of order of tens to hundreds of km depending on catchment size (CHEN et al., 2008; Havnø et al., 1995)
$1D^{+}$	1D plus a flood storage cell approach to the simulation	HEC-RAS, InfoWorks RS, Mike 11	Design scale modeling, which can be of order of tens to hundreds of km depending on catchment size, also has the potential for broad-scale application if used with sparse cross-sectional data (Brunner, 2002)
$2D^{-}$	2D minus the law of conservation of momentum	LISFLOOD-FP, CA model	Large-scale modeling or urban inundation depending on cell dimensions (Shaw et al., 2021)
2D	Solution for the 2D shallow wave equations	TUFLOW, Mike21, TELEMAC, DIVAST	Large scale modeling of the order of tens of km. It May have the potential for use in broad-scale modeling if applied with coarse grids (Falconer & Lin, 2001; Galland et al., 1991)
$2D^{+}$	2D plus a solution for vertical velocities	TELEMAC 3D	Predominantly coastal modeling applications where 3D velocity profiles are important. Has also been applied to reach scale river modeling problems in research projects (Janin et al., 1992)
3D	Solution of the 3D Reynolds averaged Navier-Stokes equations	CFX, FLUENT, PHOENIX	Local predictions of 3D velocity fields in main channels and floodplains

These tools have been widely used and have proven to be very useful in flood risk mitigation and management (Di Baldassarre, 2010). Their ability to predict inundation extents, for example, can be used to reduce potential flood damage by:

- Promoting more appropriate land use and urban planning (where applicable);
- Raising flood awareness among people living in flood-prone areas;
- Discouraging new settlements in flood plains.

2.2.3 Remote sensing and geographical information system

Remote Sensing (RS) refers to the gathering of data on a given target without making physical contact with it. In Earth Sciences, the earth's surface or the atmosphere is observed from outside of space using satellites (space-borne) or from the air using aircrafts (airborne). The remotely sensed data have been a great aid in many disciplines including natural hazard studies such as floods. Presently, thousands of satellites are orbiting the Earth. This adventure of spatial remote sensing for the observation of the Earth all started in the United States in particular with the launch of LANDSAT-1 in 1972 and AVHRR in 1978, followed by SPOT-1 launched by France. Nowadays, data collected by satellites and radars such as LANDSAT, SENTINEL, MODIS, IRIS, RADARSAT, SPOT, AVIRIS, ASTER, ATLAS, ALOS, QUICKBIRD, TerraSAR-X, TanDEM-X are extensively used in a variety of disciplines (Gandhi et al., 2016).

In recent years, many studies have demonstrated the value of integrating RS and Geographic Information System (GIS). GIS is a computer-based system that handles georeferenced data and is one of the most important developed tools for hydraulic modeling (Huisman et al., 2009). It encompasses,

- Data capture and preparation,
- Data management (storage and maintenance),
- Data manipulation and analysis,
- Data presentation

In the GIS environment, the most familiar model is a map. A map is a miniature and graphic representation of the world, a static model. With recent advances in technology, many modeling applications use GIS as the database manager and visualization tools through the use of Windows Graphical User Interfaces (GUIs) making the output easier to understand by its users. The benefits of GIS integrated modeling are immense.

Many GIS software have been developed. Among the most commonly used software are:

- ArcGIS: developed by ESRI, it is the industry standard GIS software for users to create maps, perform spatial analysis, and manage data,
- QGIS: a free and open-source cross-platform desktop GIS that allows the users to view, edit, and analyze geospatial data. It integrates with other open-source GIS packages such as SAGA, GRASS, PostGIS, MapServer, etc,

- GRASS: a popular and well-known open-source software application which has raster and vector processing systems with data management and spatial modeling system,
- ILWIS: a multi-functionality GIS and Remote sensing software which has the capacity of model building,
- SPRING: a GIS and Remote sensing image processing software with an object-oriented model facility,
- Etc.

Although these tools have improved with recent advances in technology, they mainly target expert users and often lack sufficient resources for efficient computing. Besides, they are often slow in completing the complex geographic analysis (Zhu et al., 2021). But, recent years have seen the development of cloud computing systems allowing the users to capture, manipulate, analyze the different spectral and spatial data directly on the cloud, thus significantly speeding up the process. Google Earth Engine (GEE) is a perfect example of such platforms. GEE is a sophisticated online platform for large-scale cloud-based remote sensing data processing. It gathers the world's satellite imagery and geospatial datasets — trillions of scientific measurements spanning nearly four decades — and makes them available online with tools for scientists, independent researchers, and nations to mine this massive data warehouse to detect changes, map trends, and quantify differences on the Earth's surface. Because everything is done on the cloud, there is no longer a need to download raw images and the platform provides a variety of constantly updated data sets. One of the major benefits of the platform lies in its incredible computing speed because processing is outsourced to Google servers. The Earth Engine Code Editor, one of GEE's products, is a web-based integrated development environment that makes building complicated geospatial processes easier (*Google Earth Engine*, n.d.).

2.3 Flood Susceptibility Mapping

Previous research have established the fact that floods cannot be prevented (Huang et al., 2008; Khosravi et al., 2019; Malik et al., 2020). Although it is impossible to prevent floods, precise predictions and flood control can be obtained by applying appropriate procedures and conducting adequate studies. Recent years have thus seen a shift in flood mitigation and prevention techniques from a "flood defense" strategy focused on limiting the hazard through

structural measures to a "flood management" approach based on extensive risk assessment studies and cost-benefit assessments (Dottori et al., 2018). Flood susceptibility mapping (FSM) is one of the essential instruments that have played an important role in this shift (Arabameri et al., 2019).

Hazard mapping can be defined as a process or an activity which aims to establish geographically where and to what extent particular phenomena are likely to pose a threat to people, property, infrastructure, and economic activities. Flood hazard maps, therefore, designate different zones classifying the intensity of danger related to the probability of occurrences. As highlighted by Vojtek & Vojteková (2019) in their paper, mapping, and analysis of flood susceptibility is one of the most essential parts of early warning systems or methods for prevention and mitigation of future flood events because it enables the identification of the most susceptible regions based on physical factors that influence the propensity for flooding. In general, FSM entails determining flood-prone regions based on flood conditioning/causative factors.

Many recent studies have made use of data-driven models including but not limited to Bivariate Statistical Models (BSM), Multi-Criteria Decision Making (MCDM), Machine Learning (ML), Artificial Intelligence (AI), Computational Intelligence (CI), Soft Computing (SC), Data Mining (DM), Knowledge Discovery in Databases (KDD), and Intelligent Data Analysis (IDA) models in various branches of prediction sciences including FSM (M. Ahmadlou et al., 2019; Khosravi et al., 2019; Solomatine et al., 2008).

An comprehensive review of recent studies shows that most research on flood susceptibility has been carried out using GIS-based models (Mind'je et al., 2019). These analytical GIS-based models fall into two categories: statistical and machine learning (Xiao et al., 2019). Each model has its own advantages and disadvantages. Therefore, for better prediction, a variety of models must first be applied to a specific region, and only then can one be chosen for usage following an assessment of the models' predictive potential (Khosravi et al., 2018). Khosravi et al. (2019) in a recent study carried out a comparative assessment of flood susceptibility modeling using different MCDM and ML Methods. They considered the well-known VIKOR (Vlse Kriterijuska Optamizacija I Komoromisno Resenje), TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) and SAW (Simple Additive Weighting) MCDM techniques and the Naïve Bayes Trees (NBT) and Naïve Bayes (NB) ML techniques. All the models had Area Under the Receiver Operating Characteristic curve (AUC) more than 95%, thus demonstrating the strong flood prediction capability of the studied models. Malik et al. (2020), also carried out a comparative study on River

Dwarkeswar and Koiya River of Bengal Basin, India, using four models i.e.; Analytical Hierarchy Processes (AHP), Knowledge-Driven (KD), Fuzzy Logic (FL), and Logistic Regression (LR). Fourteen causal factors were used for the assessment, inter alia normalized differential vegetation index, rainfall, stream power index, land use, and land cover, topographic wetness index, geology, soil, slope, elevation, drainage density, plan curvature, profile curvature, aspect and distance from river. The results showed a successful prediction capability with LR having the highest rate (91.6%) followed by FL (89.3%), AHP (86.9%), KD (84.1%). Other models have also been successfully investigated for the natural hazard susceptibility modeling in many places in the world such as Artificial Neural Network (ANN), Frequency Ratio (FR) (Xiao et al., 2019), Support Vector Machine (SVM) (Roy et al., 2020), Random forest (RF) (Xiao et al., 2020), Evidential Belief Function (Arabameri et al., 2019), etc. Though many of these methods have been successfully applied, to date no single method has been recognized as being best the best (Costache & Tien Bui, 2019). Hybrid and ensemble models have been gaining more attention in recent years. Costache & Tien Bui (2019) in their recent study investigated the prediction potential of a number of artificial intelligence hybrid and ensemble models. Six models namely Multilayer Perceptron neural network-Frequency Ratio (MLP-FR), Multilayer Perceptron neural network -Weights of Evidence (MLP-WOE), Rotation Forest-Frequency Ratio (RF-FR), Rotation Forest-Weights of Evidence (RF-WOE), Classification and Regression Tree-Frequency Ratio (CART-FR), and Classification and Regression Tree-Weights of Evidence (CART-WOE) were investigated. All the models showed a good performance with an AUC ranging from 86.8%(the RF-FR model) to 93.9% (the RF-WOE model).

The common conclusion is that although preventing the occurrence of natural hazards and particularly flood hazards is not practically possible, accurate predictions, as well as flood control, may be achieved by employing appropriate methods and performing proper analyses (Arabameri et al., 2019).

2.4 The Context of the Study

In the context of Africa in general and of Burkina Faso in particular, flood-related fatalities by floods as well as the resulting economic losses, have grown substantially over the last half-century. Di Baldassarre et al. (2010) suggest that intensive and unplanned human settlements in flood-prone areas are the key contributors to this increased flood risk in Africa

while Asare-Kyei et al. (2015) believe that this trend is due to the high variability in rainfall patterns. For example, in 2007, a series of unusually abundant rainfall events caused severe floods in West Africa and other parts of Sub-Saharan Africa, affecting more than 1.5 million people and resulting in the destruction of farmland, personal properties, infrastructure, epidemic disease outbreaks, and human deaths (BBC, n.d.; Braman et al., 2013; Paeth et al., 2011). Similar floods hit an estimated 940,000 people in twelve West African nations in 2009, killing 193 people and damaging property worth $152 million (Asare-Kyei et al., 2015).

The number of fatalities caused by floods in Africa during the period 1960 – 2020, summarized in Figure 2.1, shows the dramatic increase in the number of deaths during the last 30 years.

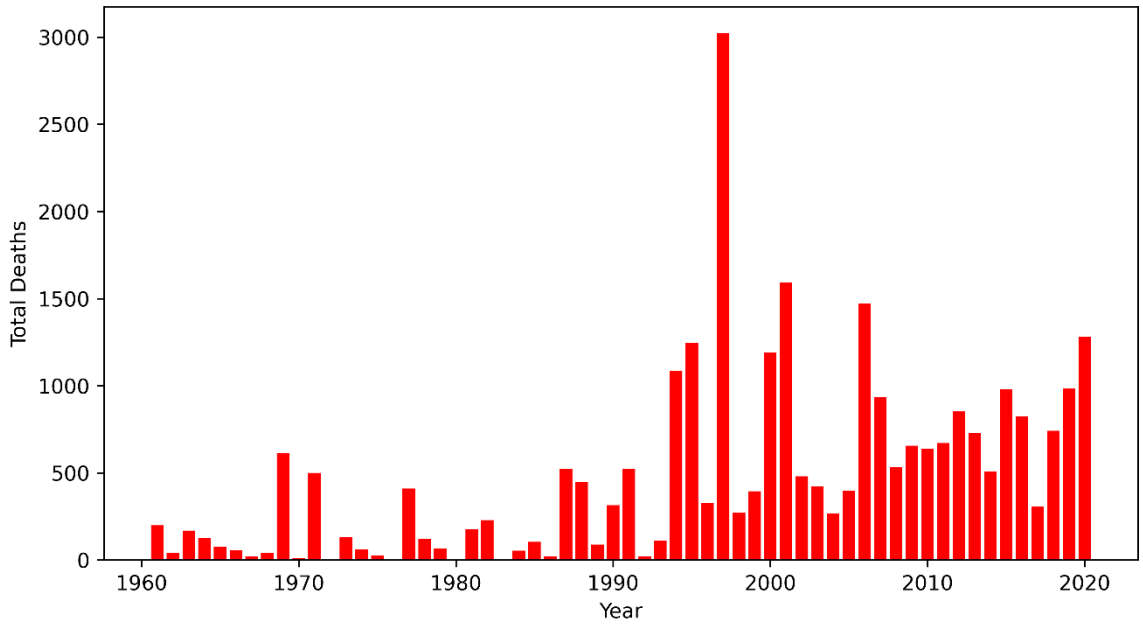

Figure 2.1. Total deaths caused by floods in Africa, 1960-2020 (Source: based on EM-DAT/CRED data)

Despite the increase in flooding events, the associated damages, and the belief that the frequency of occurrence will likely continue to grow, limited research has been conducted on the subject. There has been no attempt to define the flood intensity boundaries at the community level or to designate flood-prone regions. The usage of flood hazard maps for disaster management is almost non-existent. Even when available, the maps lack spatial variability while global datasets have resolutions too coarse to be relevant for local-scale risk assessment. Thus, traditional techniques of identifying possibly impacted regions during flood disasters, such as watermarks on buildings, local knowledge, and media reports, have

been used for many (Asare-Kyei et al., 2015). Yet, hazard maps have proven very useful in flood management around the world, particularly in developed countries where the maps have been used in:

- Early identification of vulnerable populations and elements;
- Spatial planning to avoid development in flood-prone areas;
- Served as a database for the implementation of a flood insurance scheme;
- Raising public awareness about flood-prone areas (Moel et al., 2009)
- etc.

In the context of Burkina Faso, a country-level risk profile was developed in the context of the GIRE program in 2001 by the Ministry of Environment (MECV, 2007), but since then the country has faced increasing severe extreme events. There is a need to continually update the information for better management, mitigation, and adaptation strategies.

Recently, Amadou, 2019 used a statistical method, Annual Serie Maximum, to analyze the patterns of flow data in the Nakambé basin but did not carry out a flood risk analysis. Other studies have looked at other aspects such as water management issues in the Nakambe basin (Traore, 2012), the relationship between rainfall and vegetation indexes (Diello et al., 2005). However, the literature shows no comprehensive study that has come out with flood risk maps showing the various levels of flood risk in the area.

CHAPTER 3

“Hydrologists have demonstrated that the meanderings of a creek are a necessary part of the hydrologic functioning. The flood plain belongs to the river…”

Aldo Leopold, A Sand County almanac, 1966

3 METHODOLOGY

3.1 Study Area

The Nakambe East is a subbasin of the National Basin Nakambé. In fact, Burkina Faso is located on three (03) international basins which are the Volta, the Niger, and the Comoé. The three basins are themselves subdivided in Burkina Faso into 4 national basins i.e., Nakanbé, Mouhoun, Niger, and Comoé. Finally, at a lower level, the national basins are subdivided into 17 national subbasins (MEE, 2001). Table 3.1 summarizes the basins at the different levels (international, national, subnational).

Table 3.1: Characteristics of basins in Burkina Fasso

International Basin	**National Basin**	**National Sub Basin**	**Area (km^2)**
COMOE	COMOE	Comoé - Léraba	17 590
NIGER	NIGER	-	83 442
		Beli	15 382
		Gorouol	7 748
		Dargol	1 709
		Faga	24 519
		Sirba - Gouroubi	11 946
		Bonsoaga	7 231
		Dyamangou	3 759
		Banifing	5 441
VOLTA			172 968
	NAKAMBE	-	81 932
		Pendjari - Kompienga	21 595
		Nakambé	41 407
		Nazinon	11 370
		Sissili	7 55
	MOUHOUN	-	91 036
		Mouhoun supérieur	20 978
		Mouhoun inférieur	54 802
		Sourou	15 256
BURKINA FASO			274000

The Nakambé with its 81,932 km^2 and its four sub-basins is the most populated basin in Burkina Faso with a population density of 53 inhabitants / km^2 (MEE, 2001). The basin has the highest number of dams. Around 400 dams were constructed in the basin, the most important of which are Bagré, Kompienga, Ziga, and the Toécé dam. Some of the dams are without agricultural vocation. The Kompienga, for example, is a hydroelectric dam and the Ziga is also used for the water supply of Ouagadougou.

The Nakambe River is one of the main tributaries of the Volta River. The first intermittent flows in the Nakambe occur in May. They become permanent flows in July and August at the Wayen station (20,800 km^2 watershed) and larger downstream when they reach Bagré (33,120 km^2) (Amadou, 2019). Figure 3.1 illustrates the study area, the Nakambe East basin, delineated using ArcGIS 10.7.1, with the two meteorological stations used in the study.

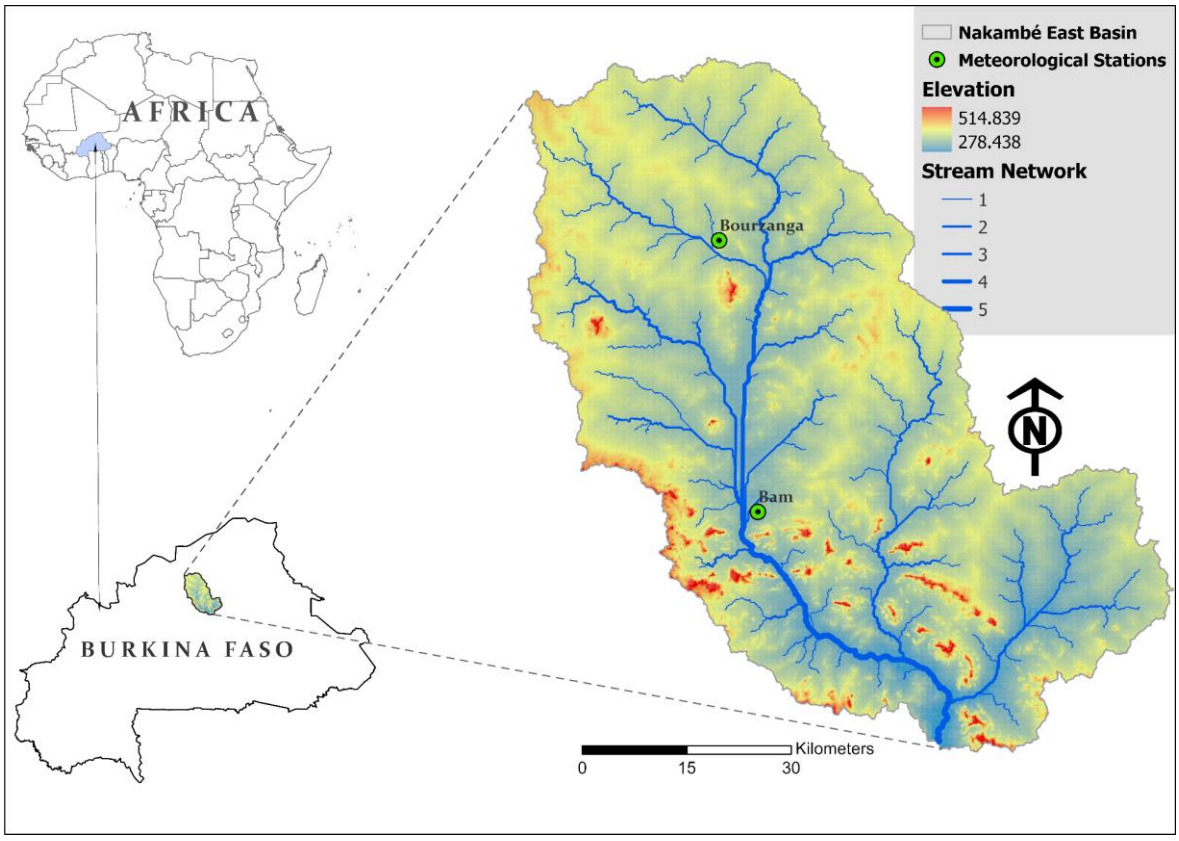

Figure 3.1: The Study Area, the Nakambé East Basin

3.2 Methods

3.2.1 Choice of model

The academic literature avails a wide range of models for flood susceptibility mapping (Malik et al., 2020). These models have been widely and successfully applied. For this research, given the data availability and the correctness of methods, the bivariate-statistics-based Frequency Ratio (FR) and the multivariate statistical and machine learning Logistic Regression (LR) methods have been used.

3.2.2 Frequency ratio

FR is a bivariate statistical analysis (BSA) method widely used in susceptibility mappings such as landslide and flood mapping (Shafapour Tehrany et al., 2019). It represents the quantitative relationship between flood occurrence and different conditioning factors. The relationship between the flood occurrence area and the factors could be deduced from the relationship between areas where flooding had not occurred and the causal factors. Despite the availability of several sophisticated machine learning methods, this method is still widely used by researchers and practitioners. It is easy to implement and yield accurate results (Rahmati et al., 2016).

The process involves the computing of the Frequency Ratio (FR), the Relative Frequency (RF), and the Prediction Rate (PR) of each factor. Microsoft Excel was used for the different computations. Equation (3.1) shows the computation of FR for a single variable (Shafapour Tehrany et al., 2019). The higher the ratio, the stronger the link between flooding and the conditioning factor, and vice versa.

$$FR = \frac{\frac{N_{pix}(SX_i)}{\sum_{i=1}^{m} SX_i}}{\frac{N_{pix}(X_j)}{\sum_{j=1}^{n} N_{pix} X_j}} \quad (3.1)$$

where $N_{pix}(SX_i)$ represents the number of flooded pixels in class i of the factor X; $N_{pix}(X_j)$ the total number of pixels within factor Xj; m the number of classes in factor Xi; and n is the

number of factors in the study area. Equation (3.2) shows a simplified version of Equation (3.1).

$$\mathrm{FR} = \frac{\%\ \text{target occurrence in each subcategory}}{\%\ \text{category of an independent factor}} = \frac{\text{points in factor class/total points}}{\text{factor class area / total area}} \tag{3.2}$$

The relative frequency (RF) of each class is computed following Equation (3.3) to normalize to FR weights. The derived RF values were used to reclassify each conditioning factor.

$$\mathrm{RF} = \frac{\text{Factor class FR}}{\sum \text{factor classes FR}} \tag{3.3}$$

A further step proposes a modification to the conventional FR method (Niraj Baral et al., 2021). It considers the mutual interrelationships among the factors by computing the predictor rate (PR) as follows:

$$\mathrm{PR} = \frac{\mathrm{RF_{min}} - \mathrm{RF_{min}}}{(\mathrm{RF_{max}} - \mathrm{RF_{min}})\mathrm{Min}} \tag{3.4}$$

Where, RFmax and RFmin are the maximum and the minimum RF among the classes within a factor while (RFmax – RFmin) Min is the minimum value among all the factors considered.

3.2.3 Logistic regression

Logistic Regression (LR) is based on data mining and is one of the influential supervised algorithms used for classification. For flood susceptibility mapping, it can be expressed as the probability of flood occurrence divided by the probability of no flood occurrence (Nachappa et al., 2020). It is based on the logistic function $\phi(z)$, which has the following definition:

$$\phi(z) = \frac{1}{1+e^{-z}} \tag{3.5}$$

where z is net input and can be computed using the equation below.

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n \tag{3.6}$$

x represents the different predicting factors while β is the weight or the coefficients of the LR model. In terms of probability, any pixel that is susceptible to flooding is represented by

the estimated probability of occurrence (p), which is indicated as the conditional probability in the LR model by the following expression (Mind'je et al., 2019):

$$p = \ln\left(\frac{p}{1-p}\right) = \frac{1}{1+e^{-z}} \tag{3.7}$$

Where p represents the probability (0 or 1) and z the linear combination of dependent variables defined in Equation (3.6).

To estimate the values of β (intercept and coefficients) the factors were evaluated in the R statistical environment, which was then used to map flood susceptibility. Figure 3.2 shows the general architecture of the LR model. The output has a value of 0 or 1. 0 denotes a 0% chance of a flood occurring, whereas 1 denotes a 100% chance of a flood occurring.

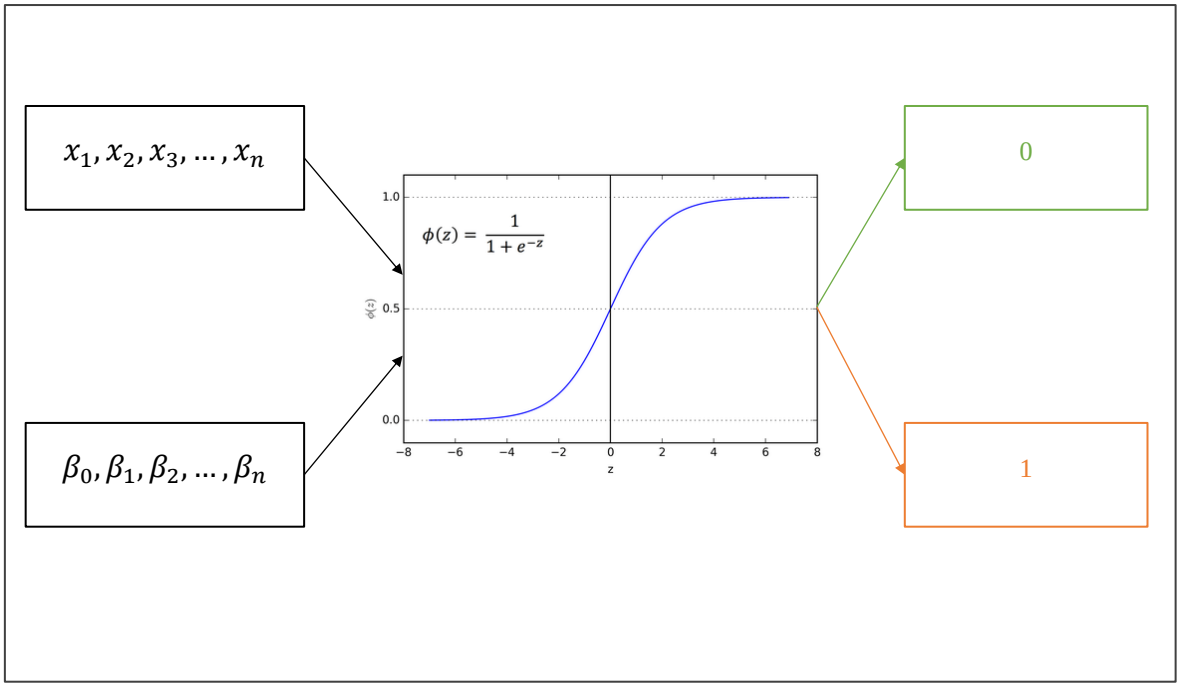

Figure 3.2. Architecture of the LR model

3.2.4 The predictive factors

The choice of factors is an important step in a predictive mapping (Costache & Tien Bui, 2019). The variables are very essential in determining and delineating flood-prone regions in any region (Malik et al., 2020). In the present study, a review of recent papers has been conducted to analyze and select the commonly used predicting factors for flood susceptibility mapping. Table 3.1 presents a list of 21 papers with the studied factors. Based

on this review 13 factors were selected including elevation, slope angle, slope aspect, curvature, distance from rivers, drainage density, Normalized Differential Vegetation Index (NDVI), Stream Power Index (SPI), Topographic Wetness Index (TWI), land use and land cover (LULC), geology, soil classes and rainfall for the basin. The factors have been described in section 3.3.

Table 3.1: Predictive factors used in the previous literature

Reference	S	E	A	R	PlC	PrC	TRI	TPI	TWI	NDVI	DD	DRi	DRo	LU	Ge	So	SPI	STI	MFI	CN	FA
Mind'je et al., 2019	*	*	*	*		*			*	*		*	*								
Costache & Tien Bui, 2019	*	*	*	*	*	*			*			*		*	*						
Malik et al., 2020	*	*	*	*	*	*			*	*	*			*	*	*	*				
Khosravi et al., 2019	*	*		*	*				*	*		*		*	*	*	*	*			
Arabameri et al., 2019	*	*	*		*				*	*	*	*		*	*	*	*				
Kanani-Sadat et al., 2019	*	*						*	*	*									*	*	*
Chapi et al., 2017	*	*		*	*				*	*	*	*		*	*		*				
Vojtek et al., 2021	*	*									*	*			*					*	*
Ahmadlou et al., 2020	*	*	*	*	*				*		*	*	*		*	*					
Vojtek & Vojteková, 2019	*	*									*	*			*					*	*
Swain et al., 2020	*	*		*		*	*		*	*	*	*		*		*	*				
Chen et al., 2019	*	*	*	*	*					*		*		*	*	*	*	*			
Shafapour Tehrany et al., 2017	*	*	*	*		*			*	*		*	*	*	*	*	*	*			
Tehrany et al., 2019	*	*	*			*	*		*			*	*	*	*	*	*	*			
Nachappa et al., 2020	*	*	*	*					*	*		*	*	*	*	*	*				
Lee et al., 2017	*	*			*				*			*			*	*	*				
Roy et al., 2020	*	*		*					*		*			*	*	*	*				
Wang et al., 2019	*	*	*	*		*			*	*		*	*	*	*	*	*	*			
Bui et al., 2019	*	*	*	*	*				*	*	*				*	*	*				
Hong et al., 2018	*	*	*	*		*			*	*		*		*	*	*	*	*			
Tehrany et al., 2013		*		*		*			*			*		*	*	*	*				
Shafapour Tehrany et al., 2019	*	*	*		*			*	*			*	*	*	*	*	*	*			

S-Slope, E-Elevation/Altitude, A-Aspect, R-Rainfall/Precipitation, PlC-Plan Curvature, PrC-Profil Curvature, TRI-Topographic Roughness Index ,TPI-Topographic Position Index, TWI-Topographic Wetness Index, NDVI- Normalized Difference Vegetation Index, NDSI- Normalized-Difference Snow Index For Agriculture, DD-Drainage Density, DRi-Distance from river, Dro-Distance from roads, LU-Land use land cover, Ge-Geology/Lithology, So-Soil, SPI- Stream Power Index, STI- Sediment Transport Index, MFI-Modified Fournier Index, CN-Curve number, FA-Flow Accumulation.

3.2.5 Model validation

In the present study, the area under the receiver operating characteristic curve (ROC-AUC) was used to assess the models' performance. AUC is a well-known quantitative approach of assessing accuracy that has been widely used to analyze prediction and success rates (Shafapour Tehrany et al., 2017). ROC is a probability curve that plots the TPR (true positive rate) or specificity against the FPR (false positive rate) or sensitivity at various threshold values using the equation (3.4).

$$\mathrm{AUC} = \frac{\sum \mathrm{TP} + \sum \mathrm{TN}}{\mathrm{P+N}} \tag{3.4}$$

Where P represents the total number of floods and N is the total number of non-floods; TP (true positive) and TN (true negative) are the number of pixels that are correctly classified (Mind'je et al., 2019). In other terms, the sensitivity, specificity would mean determining the proportion of flood pixels that were correctly identified by the model. AUC ranges between 0 and 1. When it is close to 1, then the model is well able to distinguish between all the flood and the non-flood class points correctly.

3.3 Data Availability

3.3.1 Flood inventory

A well-documented flood inventory of the study area is a primary requirement in flood susceptibility mapping. An inventory of historical floods can be attempted through the catalog of existing information such as historical archives (review of scientific studies, technical reports, humanitarian reports, university theses, newspapers, etc.) or by the use of satellite images or aerial photographs. In this research, a combination of both methods has been used to generate historical flood maps based on the UN-SPIDER recommended practices on flood mapping and damage assessment (United Nation of Outer Space Affairs, 2020). Thus, after an investigation of the CRED database and different reports and newspapers such as ReliefWeb, the events of 2009 and 2020 were considered for the mapping because of the magnitude of the events and the damages thereof. The map was

generated on GEE and exported unto QGIS for the creation of 250 random points. The points were split into training (70%) and validation(%30). Figure 3.3 illustrates the training and validation datasets.

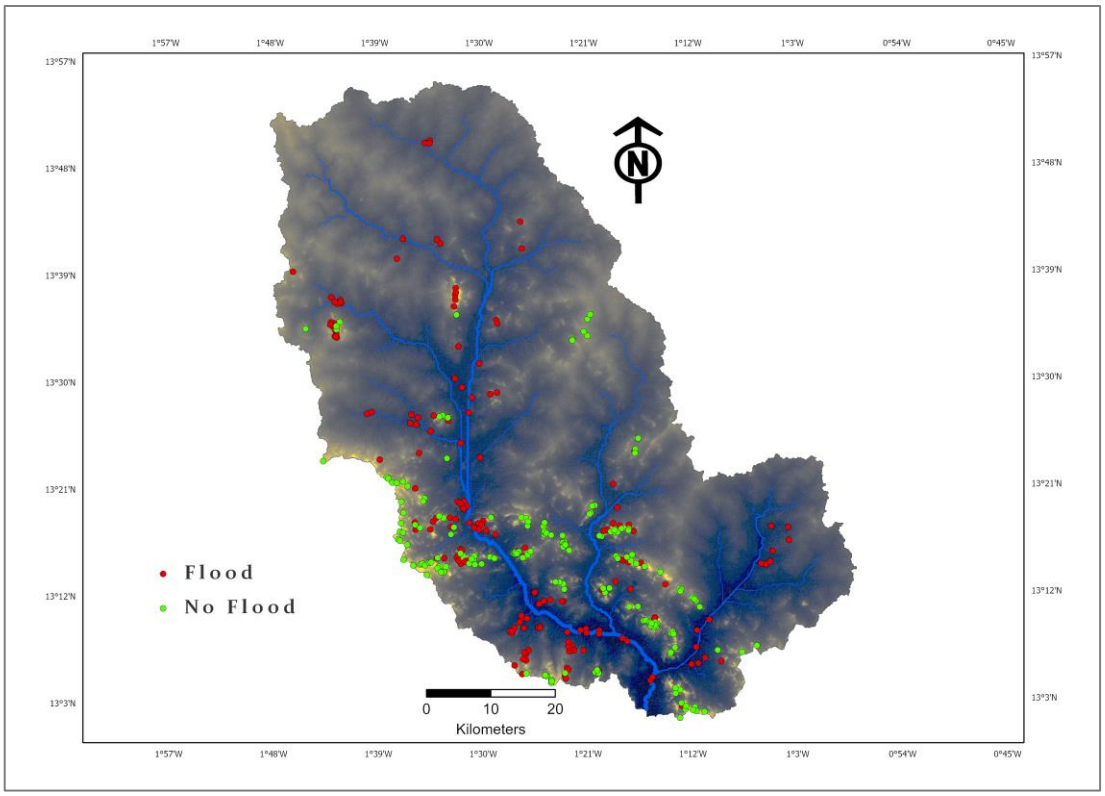

Figure 3.3. Flood inventory map

3.3.2 Elevation

Elevation has been identified in many research as one of the most important factors contributing to flooding (Malik et al., 2020). In fact, many other topographical features inter alia slope, aspect, profile and plan curvature, and TWI are derived from the elevation. Nowadays, many global digital elevation models (DEMs) are freely available and have enabled flood studies in data-sparse regions. Yet, their resolution and their vertical errors often pose an obstacle to an accurate estimate of flood hazards. It is thus important to carefully consider a model with high accuracy. After careful consideration of the most widely used open-access DEMs, the MERIT (Multi-Error-Removed-Improved-Terrain) DEM was selected for this study. In their comparative study, McCabe et al. (2018) conclude that from the freely available DEMs such as SRTM, ALOS, ASTER, Bare Earth, Earthenv,

GMTED, TanDEM-X, Viewfinder Panorama, etc, MERIT is "*the most comprehensive error removal from SRTM to date*". According to H. Chen et al. (2018) MERIT has better performance in flood models compared to SRTM which is the most widely used global DEM. The MERIT DEM was created by removing significant error components from previous DEMs NASA SRTM3, JAXA AW3D, Viewfinder Panoramas baseline DEMs, and other supplementary data (Yamazaki et al., 2017). The model was acquired using the GEE and reclassified into 5 classes using the Jenks natural break method. They are numerous techniques for reclassification in GIS but Natural break classification is one of the most widely used methods in natural hazard mapping for categorizing various classes of conditioning factors as well as susceptibility maps (Mohammad Ahmadlou et al., 2021). This method finds real classes in the data. This is beneficial because it produces maps that accurately depict data trends. Quantile, Geometric Interval, Standart Deviation, Equal Interval, Manual Interval are other methods of splitting. Figure 3.4 shows the reclassified elevation model.

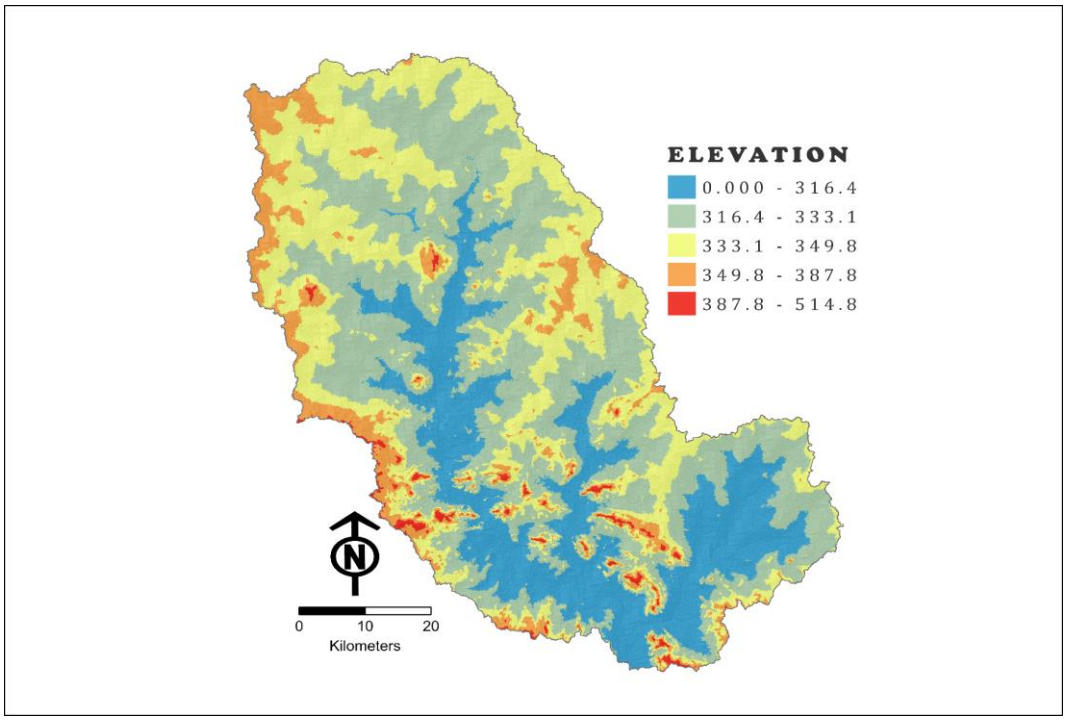

Figure 3.4. Elevation map

3.3.3 Slope angle

Slope is an important factor. It determines the quantity and intensity of water accumulation and percolation. Because water travels from higher to lower altitudes, the quantity of surface runoff and infiltration is influenced by the slope. Flooding may occur more quickly in flat regions at low altitudes than in steeper sloped areas at higher elevations (Hong et al., 2018). In the present study, the slope was derived from the DEM and reclassified in 5 classes as shown in Figure 3.5.

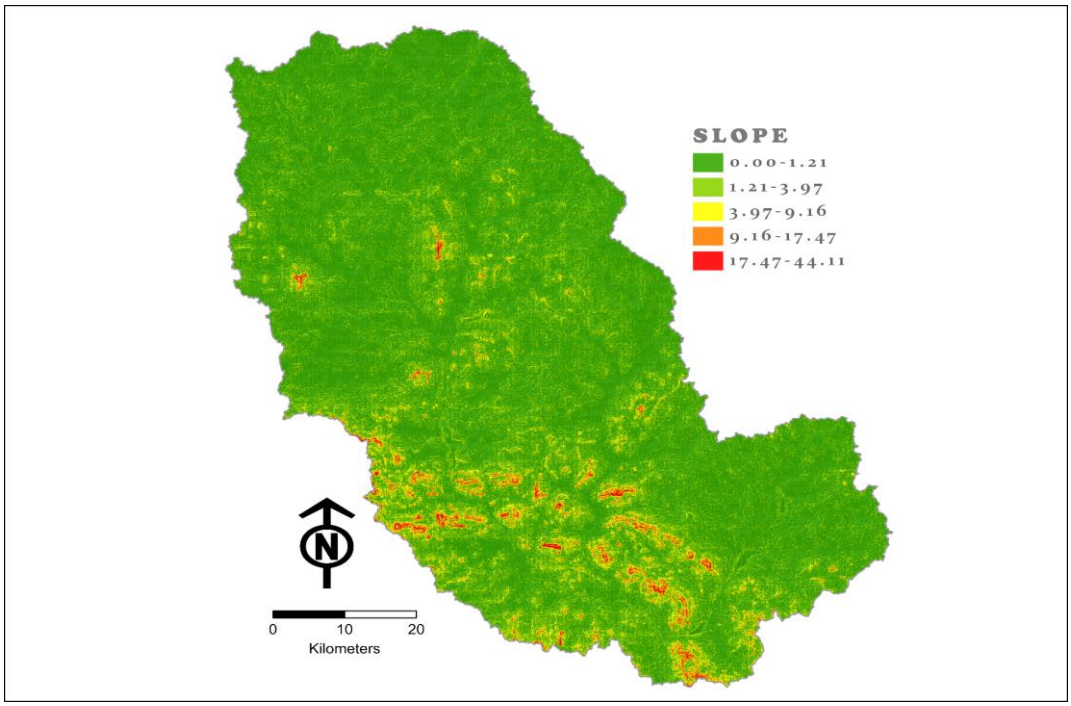

Figure 3.5. Slope angle map

3.3.4 Slope aspect

Aspect is the orientation of slope, measured clockwise in degrees from 0 to 360. It defines the direction of flow. Because various aspects get varying amounts of solar radiation, the degree of plant covering, surface weathering, and surface evaporation will all impact the occurrence of floods. Rahmati et al. (2016) suggest that aspect has an impact on hydrologic processes, local meteorological conditions, physiographic trends, and soil moisture patterns.

In this study, the aspect was derived from the DEM using the ArcGIS Spatial Analyst Tool Apect, and distinguished into 9 categories as seen in the aspect map in Figure 3.6.

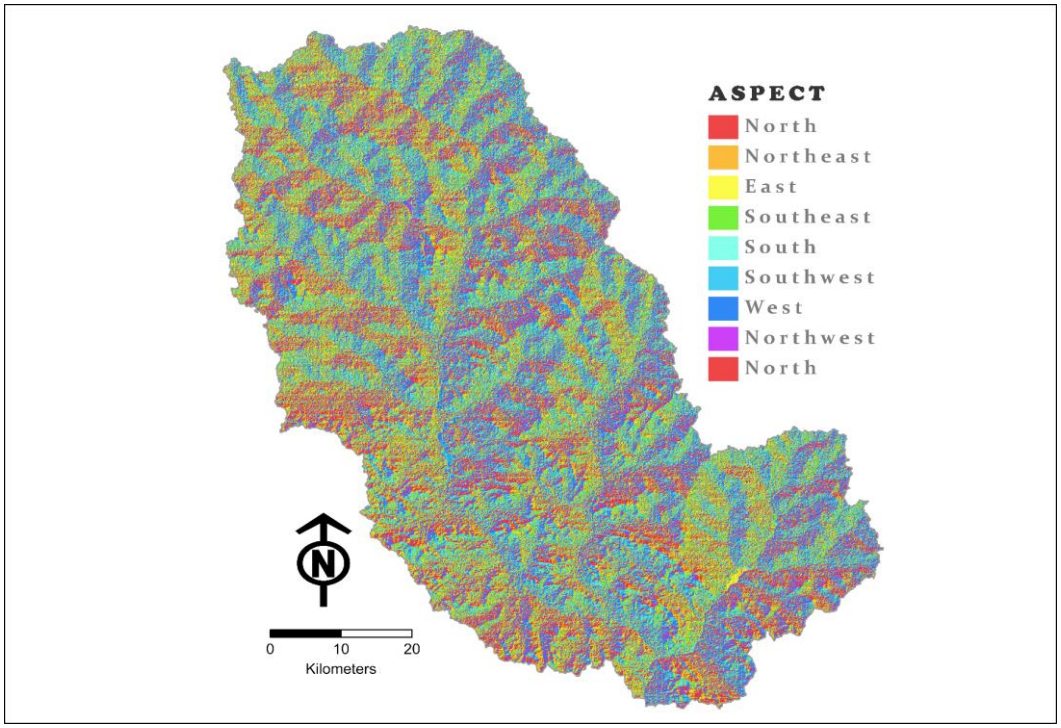

Figure 3.6. Slope aspect map

3.3.5 Curvature

Curvature is defined as the second derivative of a surface, in other terms the slope of the slope (Kimerling, 2012). The curvature is a composite of the plan curvature (parallel to the direction of the maximum slope) and the profile curvature (perpendicular to the direction of the maximum slope) which respectively affect the convergence or dispersion of flow across a surface, and the acceleration or deceleration of flow. The curvature map was computed in ArcGIS using the Spatial Analyst Tool Curvature and reclassified into three classes as indicated in Figure 3.7. The positive curvature (0.001 - 8.18) indicates that the surface is upwardly convex at that cell while the negative curvature (- 7.14 to - 0.001), indicates that the surface is upwardly concave at that cell. The null value (- 0.001 - 0.001), indicates a flat surface.

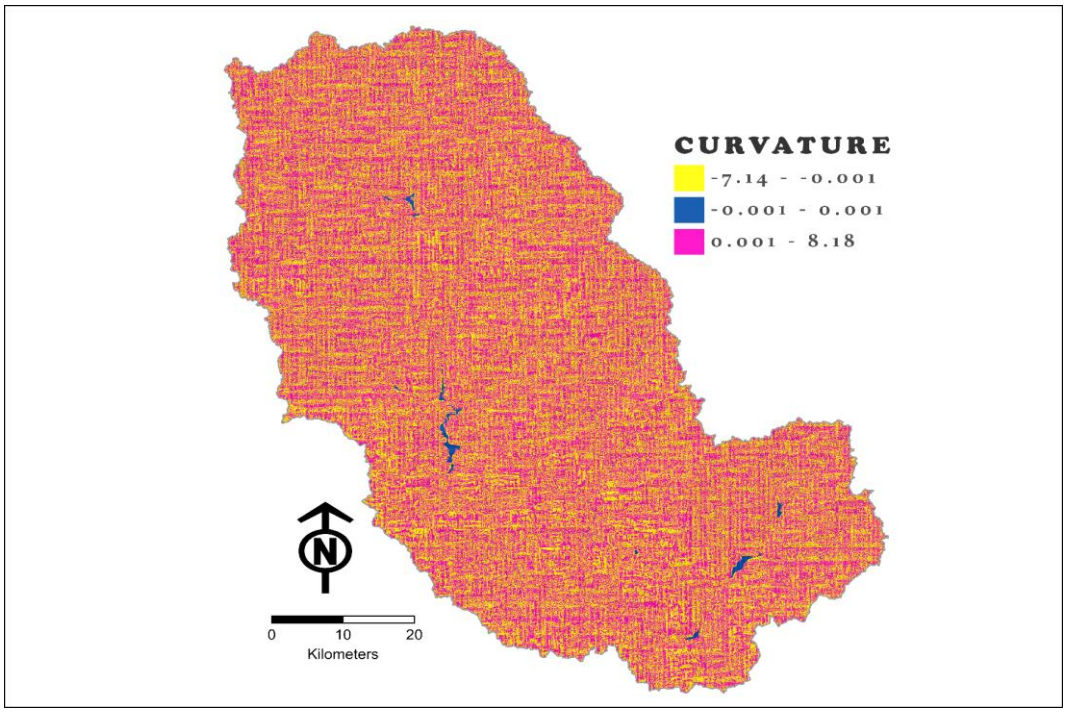

Figure 3.7. Curvature map

3.3.6 SPI

SPI is a measure of the erosive power of flowing water. SPI is calculated based upon slope and contributing area. A higher SPI indicates a larger surface runoff capacity, whereas a lower SPI indicates a lower surface runoff capacity. Thus, a heavy rainfall in a low SPI zone will significantly contribute to flooding (Malik et al., 2020). The SPI was computed in ArcGIS using the Raster Calculator following Equation 3.8.

$$SPI = A_s \tan\beta \quad (3.8)$$

Where, As represents the specific catchment area in square meter and β the slope in degree. The SPI was reclassified in 5 groups using the Jenks natural break method (Figure 3.8).

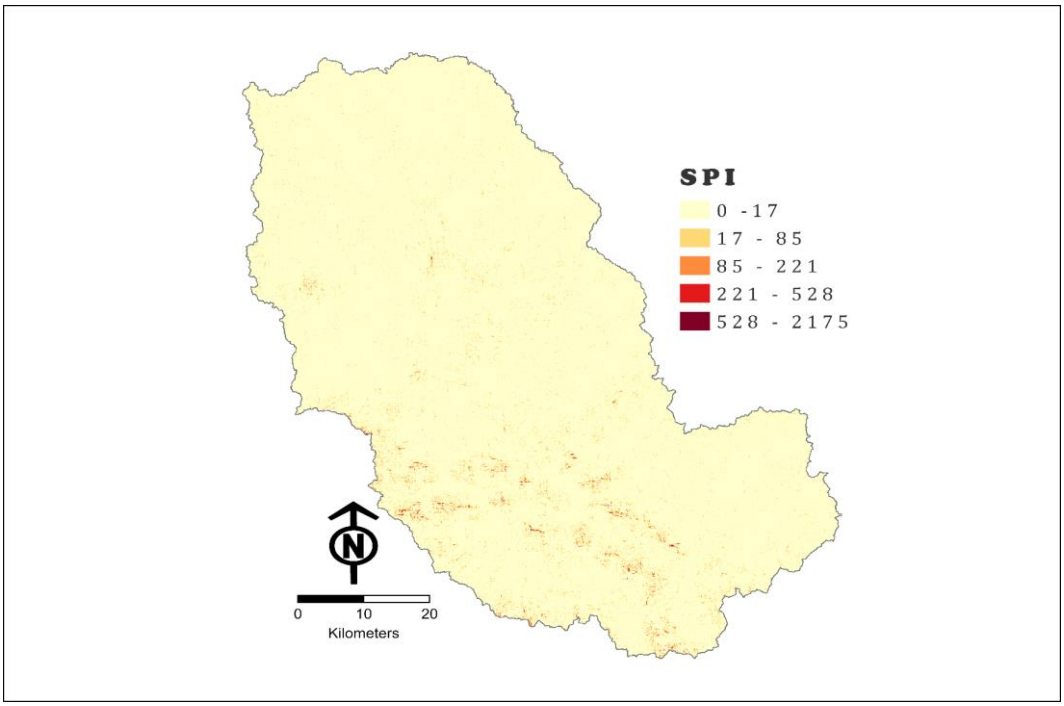

Figure 3.8. Stream power index map

3.3.7 TWI

TWI has been considered by many authors in flood studies. Also known as compound topographic index (CPI), TWI is a measure of the tendency of an area to accumulate water. It controls the geographical distribution and depletion of surface runoff (Samanta et al., 2018). Previous research suggest that there is a positive correlation between TWI and flood susceptibility (Malik et al., 2020; Wang et al., 2019). The TWI was derived in ArcGIS using the Raster Calculator following Equation 3.9.

$$\mathrm{TWI} = \ln\left(\frac{A_s}{\tan\beta}\right) \tag{3.9}$$

Where As represents the specific catchment area (m^2m^1) and β the slope in degree. High TWI values are thus associated with regions prone to water accumulation (large contributing drainage areas) and defined by a low slope angle. Low TWI values, on the other hand, are linked with well-drained dry regions (steep slopes) (Mattivi et al., 2019). In this work, the TWI values were divided into five classes (Figure 3.9).

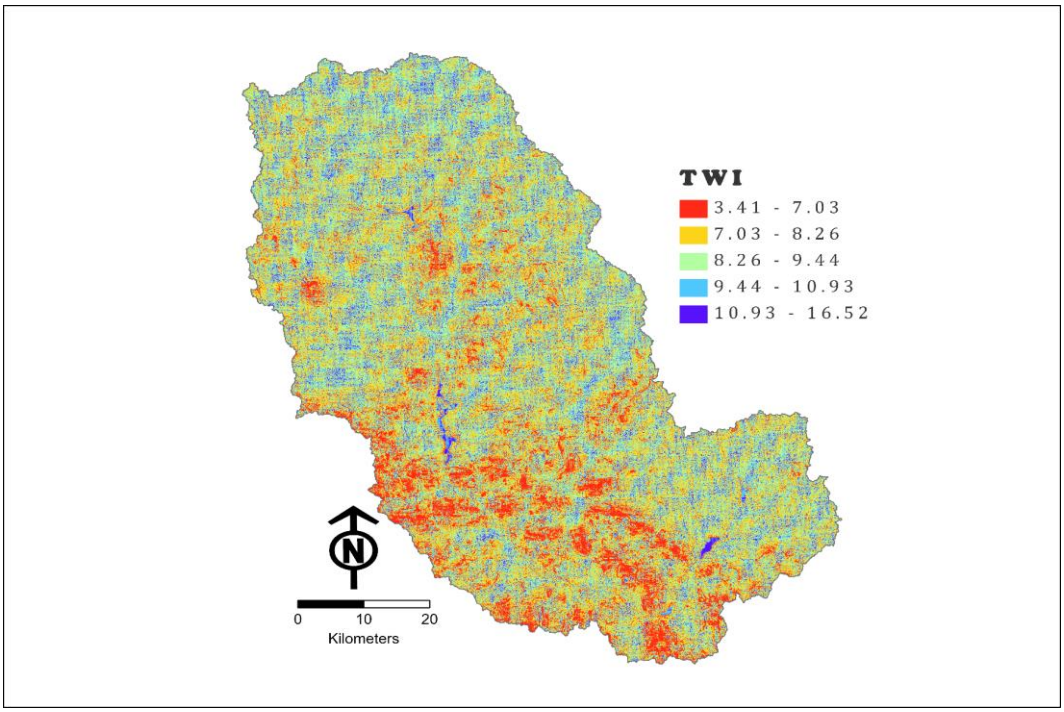

Figure 3.9. Topographic wetness index map

3.3.8 Rainfall

Rainfall is one of the most relevant predictive factors in susceptibility studies. Previous research have established a positive correlation between rainfall and flooding. As a result, excessive rainfall is closely associated with a high likelihood of flooding, whereas low rainfall areas are free of floods. (Arabameri et al., 2019; Bui et al., 2019; Costache & Tien Bui, 2019; Khosravi et al., 2019; Malik et al., 2020; Mind'je et al., 2019).

The climate of the study area is tropical, generally sunny, and hot. The temperature is generally high throughout the year with a variation interval of 15 ^{o}C to 39 ^{o}C. The climate is characterized by three to five months of rainfall with increasing variability. During the dry season, from October to May (sometimes longer) no rain falls and most rivers dry up.

In this research work, first, three (03) global precipitation datasets, for the year 1990-2020 and for the two meteorological stations (Bam and Bourzanga), were extracted from GEE. These are the *CHIRPS Daily: Climate Hazards Group InfraRed Precipitation with Station Data (version 2.0 final)* dataset (Funk et al., 2015), the *ERA5 Daily aggregates - Latest climate reanalysis produced by ECMWF / Copernicus Climate Change Service* dataset (Store, 2017), and the *PERSIANN-CDR: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks-Climate Data Record* dataset (Ashouri et al.,

2015). Daily rainfall data for the same period were also collected from the National Agency of Meteorology of Burkina Faso (ANAM - meteoburkina.bf) to refine the gridded data. A correlation analysis between the ground data and the three datasets was performed and the results are given in Table 3.2.

Table 3.2: Correlation analysis between the ground data and the gridded data

Dataset	CHIRPS	ERA5	PERSIANN
Correlation	0.931	0.886	0.911

The correlation coefficients given in the table measure the level of the association between the ground data and the gridded data, and thus suggest that CHIRPS gridded data has the strongest association. CHIRPS gridded data was therefore chosen to build a regression model and apply to the year 2020 rainfall data for correction. The linear model ($R^2 = 0.87$) built in R is given in Equation 3.10.

$$\alpha = 0.97618\beta + 0.45102 \tag{3.10}$$

Where α represents the ground data and β the CHIRPS gridded data. The downloaded CHIRPS rainfall raster for the year 2020 was thus refined following Equation 3.10 using ArcGIS Raster Calculator and reclassified into five classes (mm) as shown in Figure 3.11. The rainfall processing flowchart is shown in Figure 3.10.

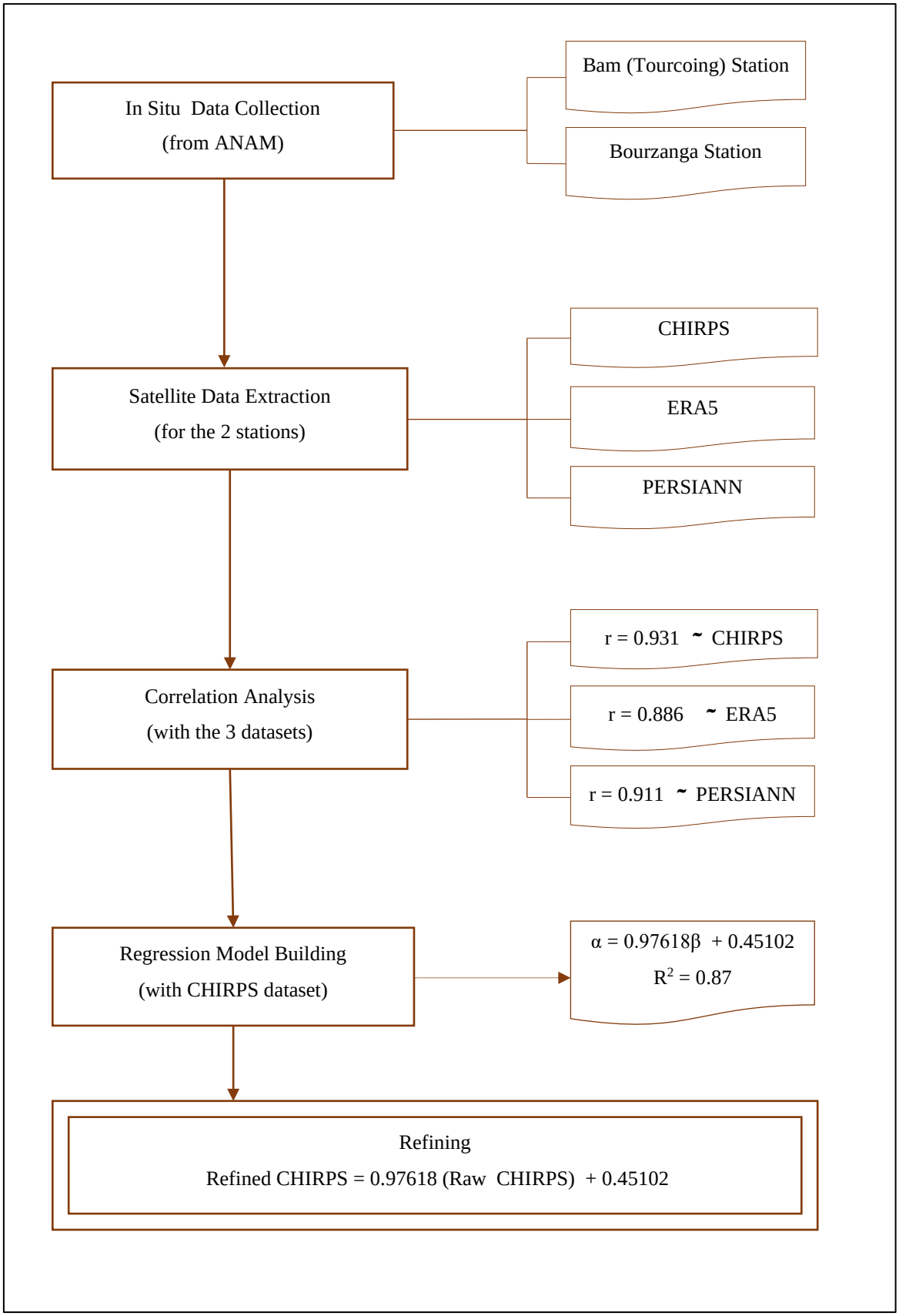

Figure 3.10. Rainfall data processing flowchart

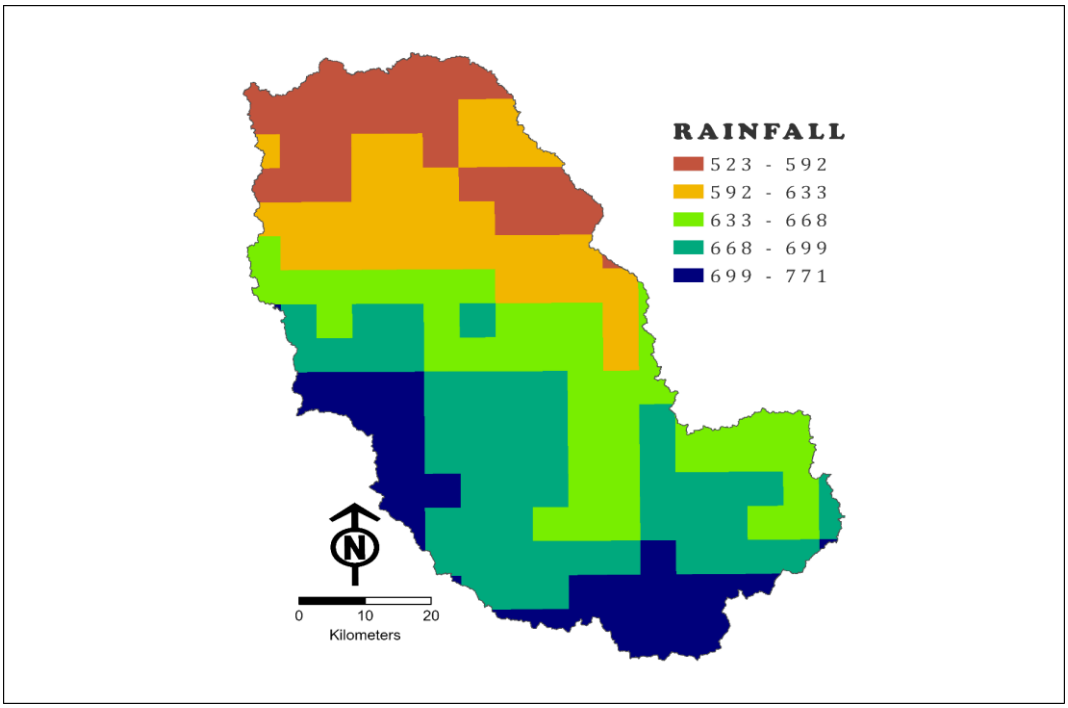

Figure 3.11. Rainfall map (mm)

3.3.9 NDVI

NDVI is a measure of surface reflectance that provides a quantitative estimate of vegetation growth and biomass. Since vegetation plays a big role in soil and water conservation and slope protection, it has a certain impact on the occurrence of flood hazards. A number of studies on flood mapping have included NDVI in the processing (Table 3.1). In this research, a Landsat 8 Level 2, Collection 2, Tier 1 image collection was acquired from GEE and used to calculate the NDVI. In total 26 images of the year 2020, with a cloud cover less than 4, were collected and reduced to the median value and used for the NDVI computation. The NDVI was thus generated in GEE following Equation 3.12 and then classified into 5 categories by ArcGIS Jenks natural breaks method. The reclassified map is shown in Figure 3.12.

$$\mathrm{NDVI} = \frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+\mathrm{RED}} \tag{3.11}$$

For Landsat 8 Level 2 Collection 2, the index is calculated using bands 5 and 4.

$$\mathrm{NDVI} = \frac{\mathrm{SR\_B5}-\mathrm{SR\_B4}}{\mathrm{SR\_B5}+\mathrm{SR\_B4}} \tag{3.12}$$

The NDVI has a value range of -1 to 1. Negative values (close to -1) represent water. Values close to zero (-0.1 to 0.1) are typically associated with barren regions of rock, sand, or snow. Low, positive values reflect shrub and grassland, whereas high positive values (close to 1) suggest temperate and tropical rainforests.

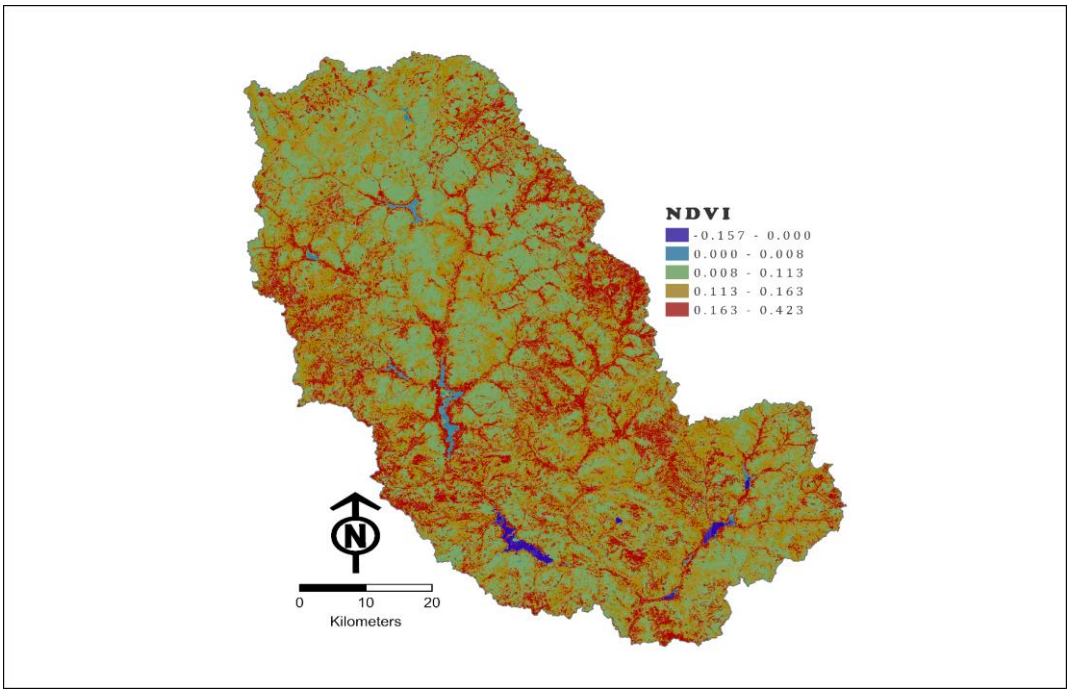

Figure 3.12. Normalized difference vegetation index map

3.3.10 LULC

It has been proven that there is a negative link between the occurrence of flooding and the density of vegetation. Rainfall on bare terrain decreases quickly in comparison to farming and woodland regions. This implies that impermeable surfaces in urban areas produce more storm runoff than equivalent regions covered by mass vegetation and forest (Tehrany et al., 2013). A LULC map is thus an important input in susceptibility mapping. In this research, the Esri Land Cover 2020 was acquired from the Living Atlas and processed. This opensource dataset is a ten-class global LULC map for the year 2020 at 10-meter resolution derived from ESA Sentinel-2 imagery. The most common landuse types within the study area are farm and farmland, vegetation, settlement area, and water body. Figure 3.13 shows the different LULC classes.

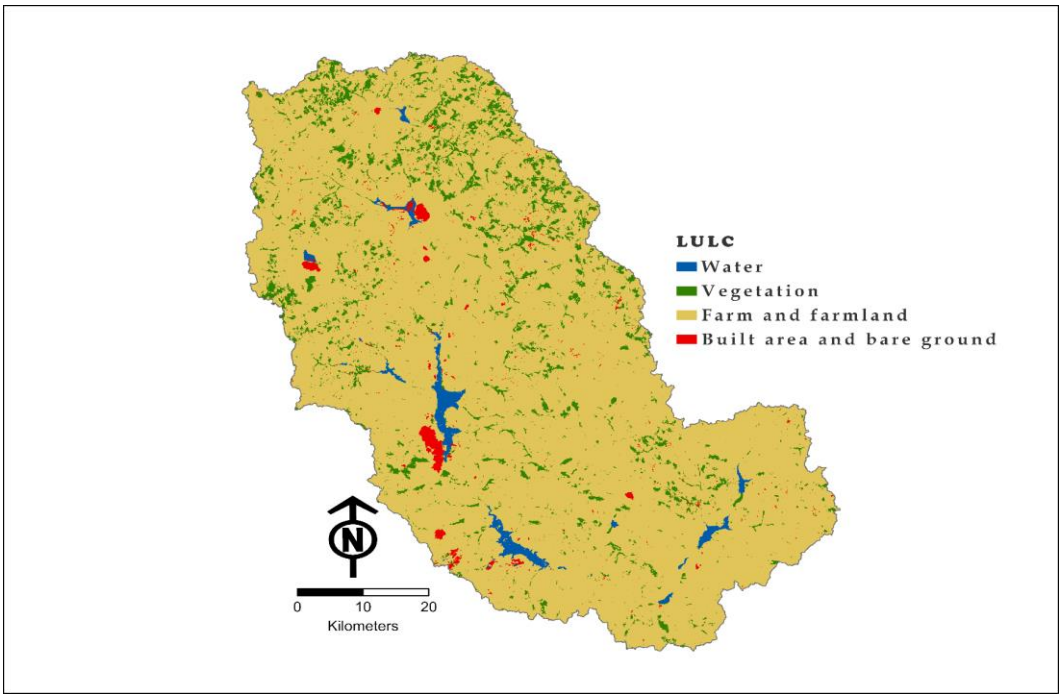

Figure 3.13. Land use land cover map

3.3.11 Soil Classes

Soil data are very useful for determining excess precipitation and infiltration. Soil type determines the degree of drainage and urbanization (Hong et al., 2018). The soil factor as shown in Table 3.1 has been considered by a great number of authors in their research work. The soil map in the present study was acquired from the National Soil Bureau of Burkina Faso (BUNASOLS - bunasols.bf). The study area is characterized by 6 different types of soil series as seen in Figure 3.14.

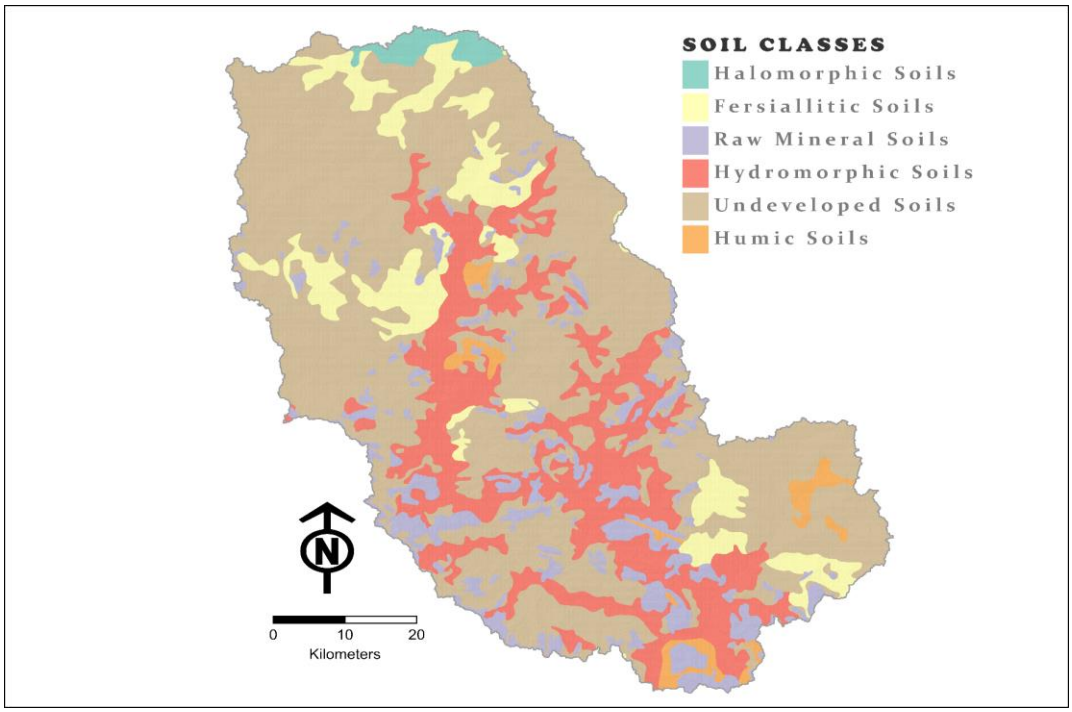

Figure 3.14. Soil classes map

3.3.12 Geology

Geology is believed to play a major role in flood susceptibility investigations because of the various susceptibilities of lithological units to active hydrological processes (Rahmati et al., 2016). Different geological features may be used to analyze the type of infiltration, runoff generation, and channel qualities of a river and its basin (Miller, 1990), which can then be used to predict the vulnerability of a river system to floods. Several authors in the past have considered geology in their susceptibility studies (W. Chen et al., 2019; Nachappa et al., 2020; Roy et al., 2020; Shafapour Tehrany et al., 2017; Tehrany et al., 2019). In the present study, a geological map was acquired from the Bureau of Mines and Geology of Burkina Faso (BUMIGEB - bumigeb.bf). The study area is characterized by 4 different geological features as can be seen in Figure 3.15.

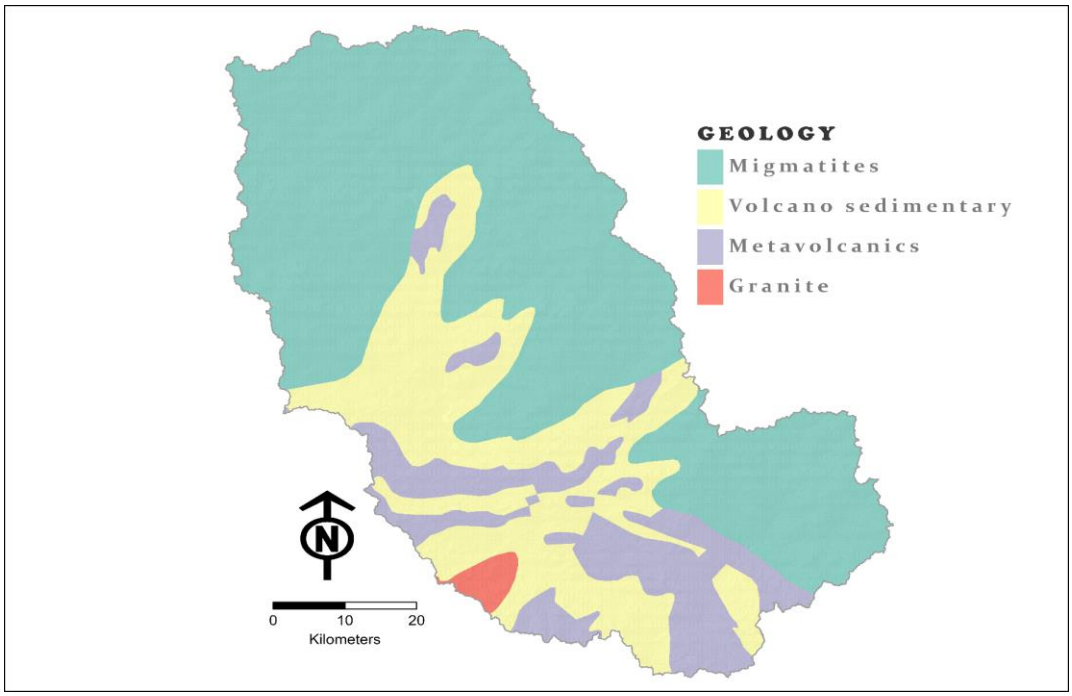

Figure 3.15. Geological map

3.3.13 Drainage density

Drainage density is the channel length per unit basin area (km km^{-2}). It represents the amount of channels in a watershed. During runoff in a drainage system, concentrated flow develops, and when the channelized flow exceeds the channel potential, excess water discharges in the surrounding area, causing floods (Malik et al., 2020). Drainage density affects infiltration rate. High drainage density values are associated with lower infiltration rates, high sediment yield conveyance via river network, higher surface flow velocity, and high flood peaks (Dragičević et al., 2018). As a result, drainage density would be one of the most important factors in determining a region's flooding susceptibility. The drainage was computed in ArcGIS using the Spatial Analyst Tool Line Density following Equation 3.13 and reclassified into 5 groups as seen in Figure 3.16.

$$DD = \frac{L}{A_s} \qquad (3.13)$$

Where L is the total length of channels and As the basin area.

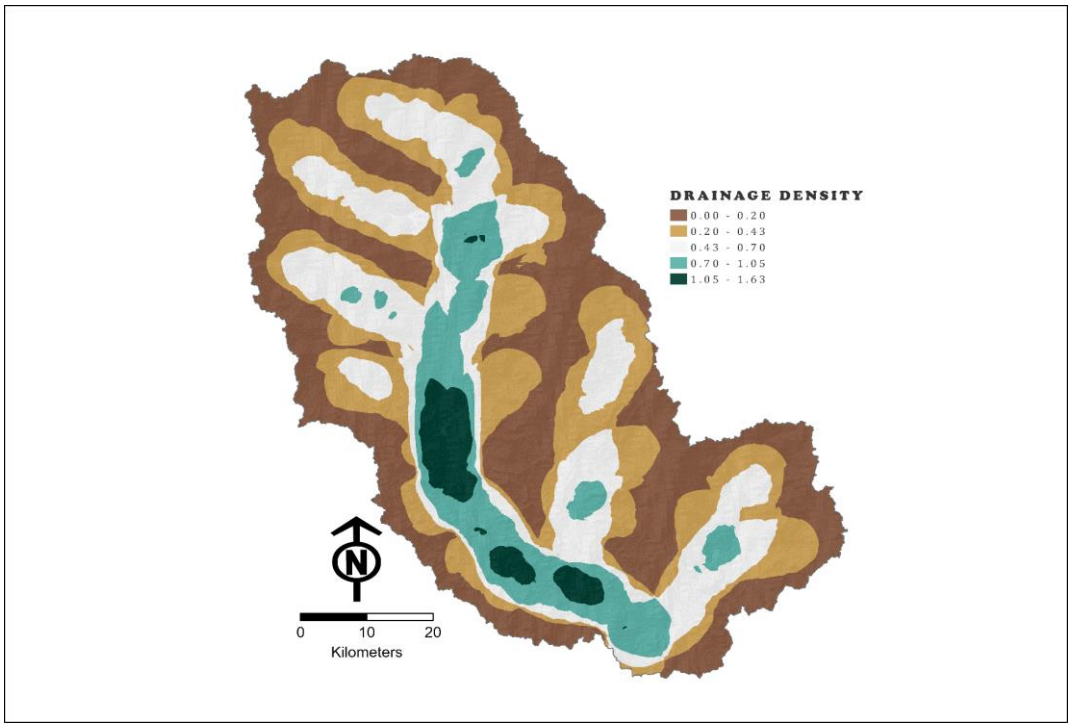

Figure 3.16. Drainage density map

3.3.14 Distance from river

The distance from the river is an important factor in flood susceptibility assessment (Malik et al., 2020). When heavy causes streamflow to grow, it reaches the capacity limit of the stream and becomes a flood. Thus, a distance closer to the river implies and higher flooding susceptibility while a more distant place is less prone to flooding. The distance from the river was computed in ArcGIS using the river network generated from the DEM. The distance to the river raster (values in m) was thus generated using the Euclidean distance tool and divided into 5 classes as seen in Figure 3.17.

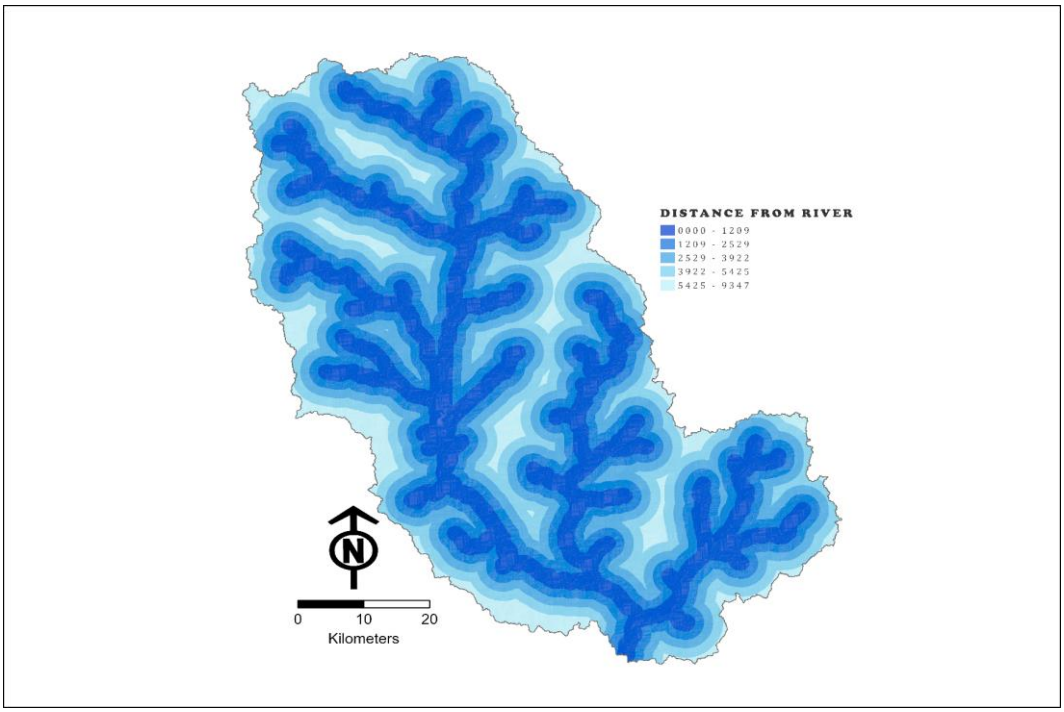

Figure 3.17. Distance from river map (m)

3.4 Methodological Framework

A detailed flowchart is provided in Figure 3.17 to demonstrate the steps and process of the data analysis in preparing the maps of the research area's flood susceptibility.

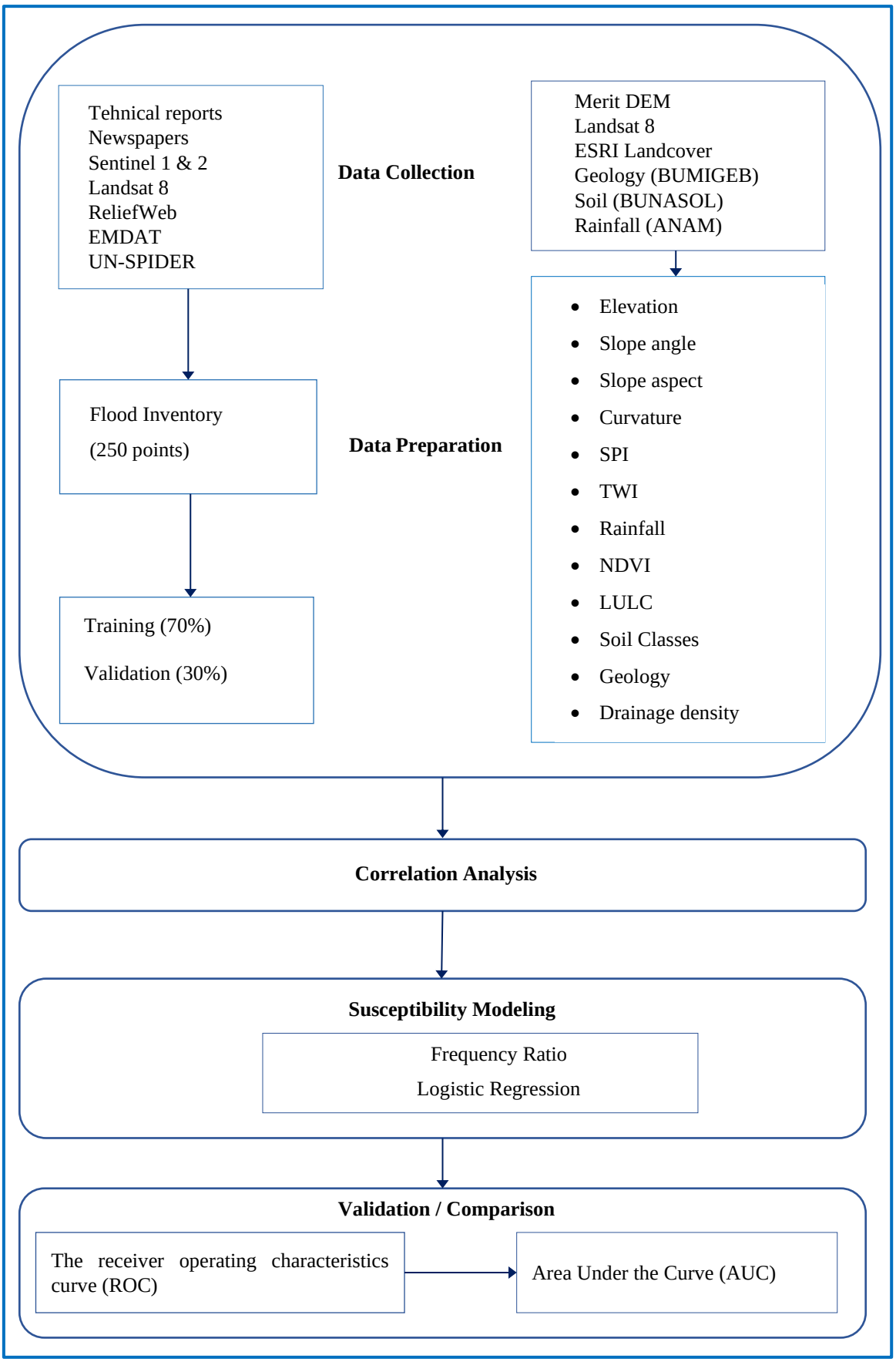

Figure 3.18. Flowchart of the proposed framework

CHAPTER 4

“When the flood submerges the whole country, no raindrop may feel responsible. Finally things had lost their weightiness.”

Erik Pevernagie

4 RESULTS AND DISCUSSIONS

4.1 Frequency Ratio

4.1.1 The role of conditioning factors

The correlation between each class of variables and floods was investigated using the FR method presented in section 3.2.2. The relative frequency of each variable was computed in Excel. The results are presented in Table 4.1. As can be seen in the table, the 0.000 - 316.4 m height class, the 0–1.21 slope class, and the highly drained areas (1.05 - 1.63) were among the most important classes that were assigned the highest weights by the FR method. On the contrary, the elevation above 333 m class received the lowest weight. In other terms, the probability of flooding increases as the elevation decreases. This finding is in agreement with previous research (Mohammad Ahmadlou et al., 2021; Shafizadeh-Moghadam et al., 2018). Tehrany et al. (2013) also reported that as the slope increases, the probability of occurrence increases in the lower parts of the catchment.

Another factor which directly affects the occurrence of floods in our study area is the distance from rivers. The results show that the closest the regions are to the rivers the most likely they are to flood. Indeed, the first class (0-1200 m) is associated with the highest probability, 35%. Besides, as seen in Table 4.2, proximity to river network is regarded as one of the most important parameters in this study by the FR model. This pattern is expected because most floodings are caused by an overflow of water from the river banks during heavy rain(Chapi et al., 2017).

It has been demonstrated in previous research that the higher the rainfall, the higher the probability of flood occurrence will be (Shafizadeh-Moghadam et al., 2018). Our findings are also in line with this because the last two classes in the rainfall factor show the highest contribution, as seen from the table.

The roles of the geology and the land cover factors were also quite remarkable. The Migmatites class has the highest weight. This finding was expected because this class alone represents 60% of the study area. Besides, as expected, water bodies displayed the most impacts on flooding and showed the probability of flood occurrence. On the other hand, as expected, the vegetation class showed a lower impact on the flood because the vegetation is

known to slow down the water flow and increase the penetration rate thus preventing flooding.

The role of NDVI was also remarkable. Since lower NDVI are associated with water it was expected to get a higher probability in lower NDVI classes. With a glimpse into Table 4.1, it can be found that indeed that classes 1 and 2 alone represent 86% of the factor's total relative frequency.

The SPI and the TWI are also two effective hydrological factors influencing the occurrence of floods in the study area. As can be seen from the table, higher SPI values are associated with high flood probability, thus corroborating previous findings reports from Chapi et al., (2017). On the other hand, lowest TWI values show high probability values.

Table 4.1. Spatial relationship between floods and conditioning factors

Factor	Factor Classes	No. of points	Class Area	RF
Slope Angle	0.00 - 1.21	75600	23352	0.83
	1.21 - 3.97	37800	70798	0.14
	3.97 - 9.16	24300	199877	0.03
	9.16 - 17.47	15300	1342461	0.00
	17.47 - 44.11	4500	3790564	0.00
Slope Aspect	North	10800	400808	0.10
	Northeast	11700	520454	0.08
	East	27900	868488	0.12
	Southeast	10800	533615	0.08
	Southeast	20700	836301	0.09
	Southwest	11700	522524	0.08
	West	23400	838406	0.10
	Northwest	20700	501298	0.15
	North	19800	405158	0.18
Curvature	-7.14 - -0.001	85500	2688631	0.39
	-0.001 - 0.001	1800	79027	0.28
	0.001 - 8.18	70200	2674722	0.32
Distance from River	0000 - 1209	66600	1477363	0.35
	1209 - 2529	40500	1357993	0.23
	2529 - 3922	30600	1213893	0.20

	3922 - 5425	14400	950301	0.12
	5425 - 9347	5400	442830	0.10
Drainage Density	0.00 - 0.20	18000	1971073	0.03
	0.20 - 0.43	16200	1418678	0.04
	0.43 - 0.70	48600	1226044	0.14
	0.70 - 1.05	41400	618449	0.23
	1.05 - 1.63	33300	208136	0.56
Elevation	0.000 - 316.4	71100	151936	0.82
	316.4 -333.1	32400	456912	0.12
	333.1 -349.8	26100	1608436	0.03
	349.8 -387.8	21600	2139978	0.02
	387.8 - 514.8	6300	1085118	0.01
Geology	Migmatites	36000	56087	0.84
	Volcano sedimentary	67500	800053	0.11
	Metavolcanics	47700	1376491	0.05
	Granite	6300	3209701	0.00
LULC	Water	23400	57539	0.52
	Vegetation	7200	82837	0.11
	Farm and farmland	114300	396689	0.36
	Built area	12600	4908477	0.01
NDVI	-0.157 - 0.000	2700	24794	0.18
	0.000 - 0.008	12600	30955	0.68
	0.008 - 0.113	57600	1780212	0.05
	0.113 - 0.163	53100	2599811	0.03
	0.163 - 0.423	31500	1015858	0.05
Rainfall	523 - 592	4500	795181	0.04
	592 - 633	14400	1031813	0.11
	633 - 668	32400	1291717	0.19
	668 - 699	74700	1586861	0.37
	699 - 771	31500	736750	0.33
Soil	Halomorphic Soils	9000	69915	0.20
	Fersiallitic Soils	30600	114968	0.42

	Raw Mineral Soils	42300	523831	0.13
	Hydromorphic Soils	72000	530122	0.21
	Undeveloped Soils	3600	1118371	0.01
	Humic Soils	54000	3085125	0.03
SPI	0 - 17	140400	5216602	0.04
	17 - 85	12600	184069	0.11
	85 - 221	2700	21359	0.20
	221 - 528	1800	4469	0.64
	528 - 2175	0	553	0.00
TWI	3.41 - 7.03	44100	615512	0.45
	7.03 - 8.26	43200	1898497	0.14
	8.26 - 9.44	40500	1603415	0.16
	9.44 - 10.93	23400	895259	0.16
	10.93 - 16.52	6300	414369	0.09

4.1.2 Flood susceptibility map by FR

After determining the relative frequency of each factor, their predictive rates (PR) were established following Equation 3.4. Table 4.2 presents the results.

Table 4.2. Predictive rate of the different conditioning factors

No	Factor	PR
1	Slope Angle	7.73
2	Slope Aspect	1.00
3	Curvature	6.70
4	Distance from River	8.43
5	Drainage Density	9.03
6	Elevation	7.55
7	Geology	7.82
8	LULC	4.81
9	NDVI	6.01
10	Rainfall	3.00
11	Soil	3.88
12	SPI	6.01
13	TWI	3.27

The differents PR were then used in ArcGIS to derive the final susceptibility map following Equation 4.1.

$$\text{FSM} = (773) * \text{Slope Angle} + (100) * \text{Slope Aspect} + (670) * \text{Curvature} + (843) * \text{Distance from rivers} + (903) * \text{Drainage Density} + (755) * \text{Elevation} + (782) * \text{Geology} + (481) * \text{LULC} + (601) * \text{NDVI} + (300) * \text{Rainfall} + (388) * \text{Soil"} + (601) * \text{SPI} + (327) * \text{TWI} \quad (4.1)$$

Figure 4.1 provides the reclassified prediction map by FR method. The literature gives no standard method for the reclassification of the resulting susceptibility maps. The results were classified into four classes namely low, medium/moderate, high and very high susceptibility levels using the ArcGIS quantile classification method. This methode distributes the values equally across the class interval, giving unequal class widths but the same frequency of observations per class. Roy et al. (2020) argued that this method is better than the commonly used natural breaks where some classes might have limited or excessive number of values.

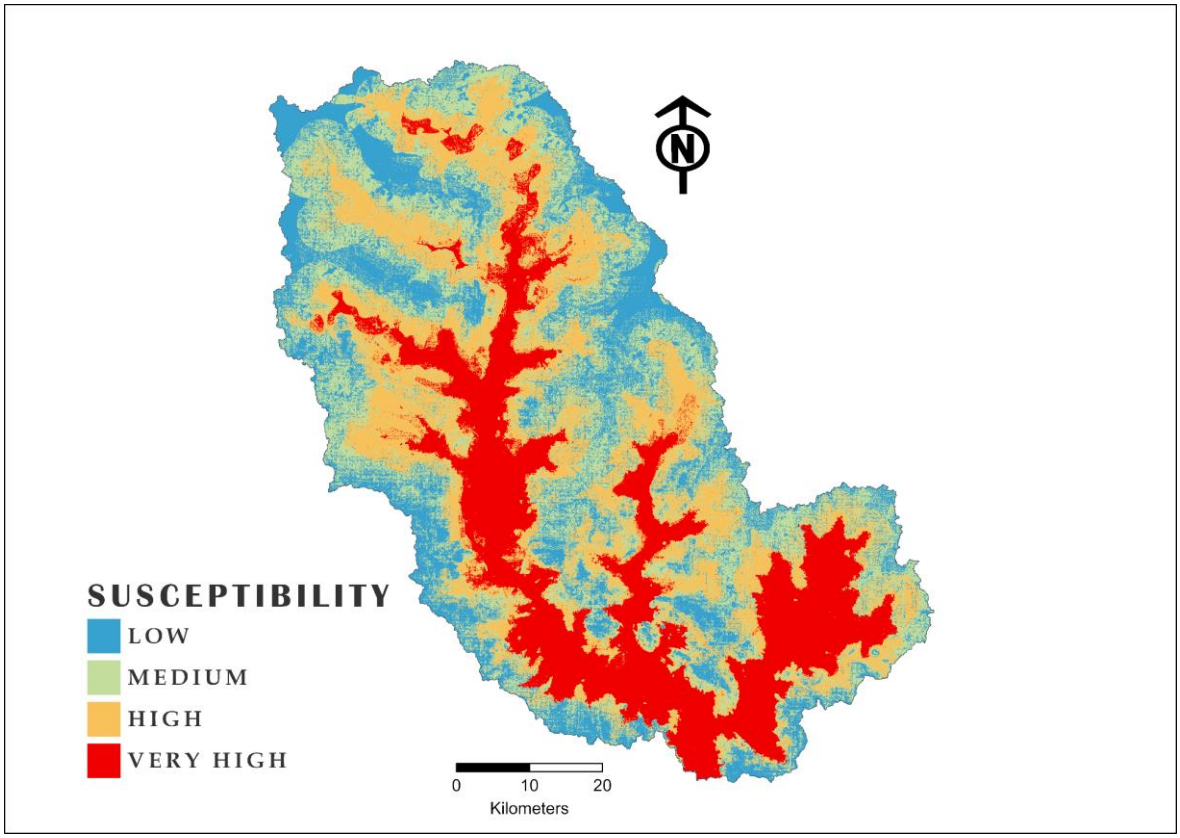

Figure 4.1. Flood susceptibility map with frequency ratio

The map closely follows the trend of drainage density and distance from rivers factors. This was expected because as can be observed from Table 4.2, these two factors have the highest weights predicted by the FR model.

4.1.3 FR model assessment

The accuracy of the model's outputs was assessed using the area under the receiver operating characteristic (ROC) curve (AUC). The result indicates that the model is fairly capable of predicting flood-prone areas in the study area with an AUC value of 80.5%. Moreover, as already mentioned in the methodology, the inventory dataset was split into 70% for training and 30% for testing. The testing resulted in a slight decrease in the model perfomance with an AUC value equals to 78.1%. Figure 4.2 illustrates the ROC curves for the training and validation datasets.

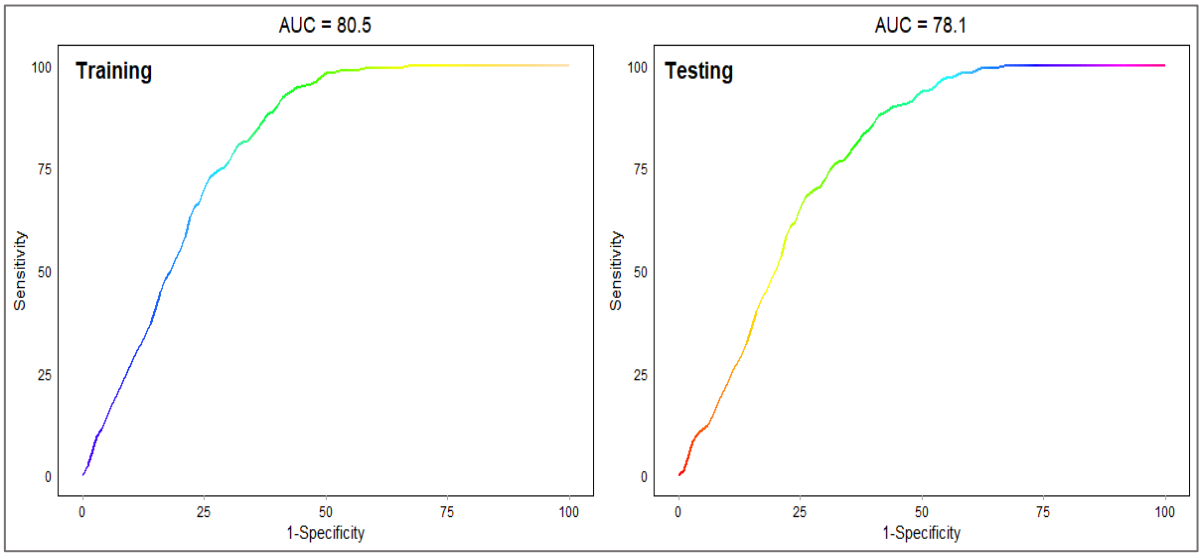

Figure 4.2. Frequency ratio model's ROC curves

4.2 Logistic Regression

4.2.1 Correlation analysis

Figure 4.3 illustrates a correlogram of a pairwise correlation analysis which was performed on the training data and the conditioning factors using the Pearson method. Pearson correlation is one of the most common types of correlations, with coefficient values ranging from +1 to -1, where +1 indicates a perfect positive relationship, -1 a perfect negative relationship, and 0 no relationship. From the correlogram, it can be seen that among the factors, the elevation and the slope show the highest correlation coefficients. The negative correlation between these factors and the training dataset explain the fact that regions with high slope and elevation are less likely to flood; and is consistent with previous research (Hong et al., 2018).

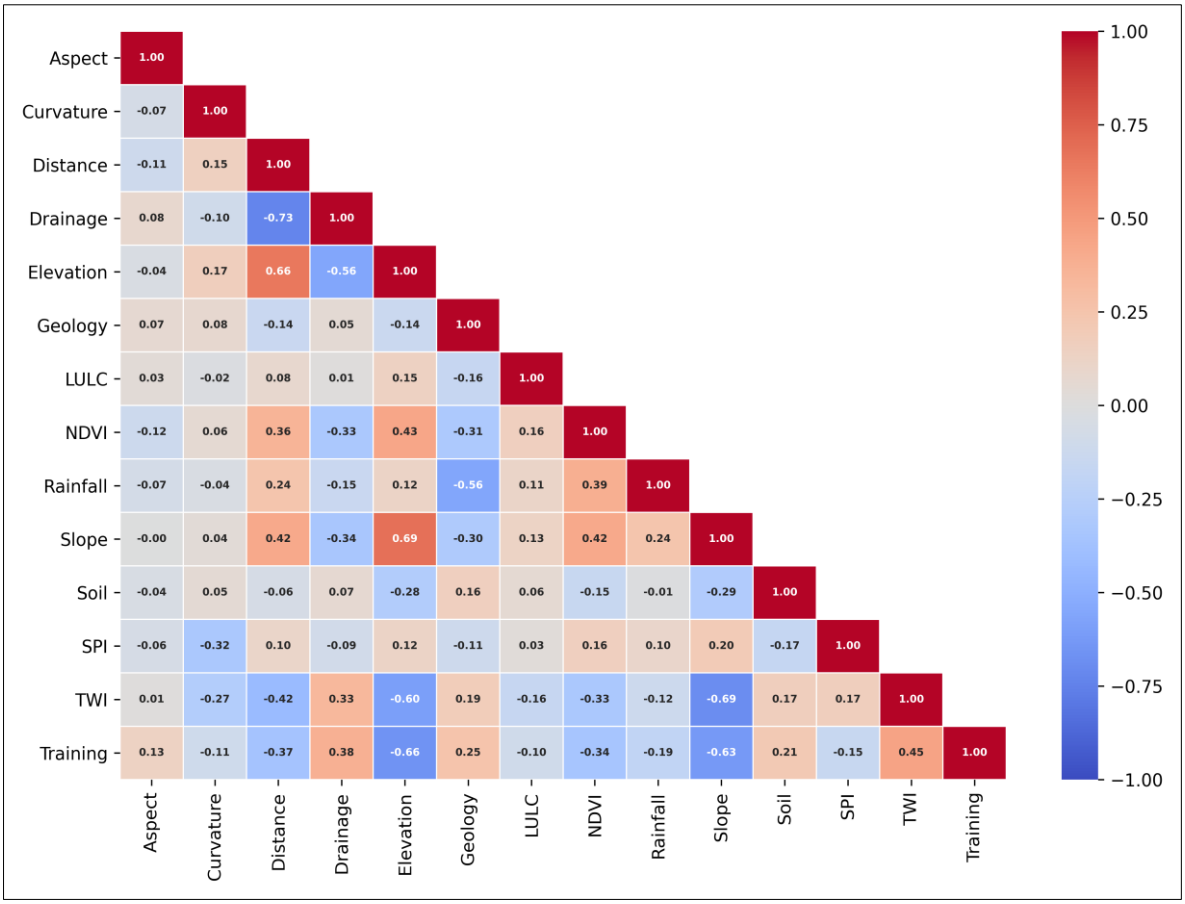

Figure 4.3. Pearson correlation matrix

Despite the fact that each of the thirteen parameters studied is directly connected to flood susceptibility, either with a positive relationship or a negative relationship, there may be some correlation between them. The analysis findings may be skewed if partially unnecessary index components are included without evaluation or adjustment. It was, therefore, important to test for correlation between the factors themselves to ensure that they are sufficiently independent of one another. The results (Figure 4.3) show that the correlation coefficients are not very high. There is, therefore, an acceptable level of independence between the factors. The TWI and the elevation have a strong relationship with the slope (-0.69 and +0.69) which is expected because the slope is directly connected to the elevation; higher elevations will result in step slopes. Besides, since TWI describes the effect of topographic changes on surface water runoff, when a slope is flater, topographic changes will have a greater impact on runoff, thus the inverse relationship with the slope. All the thirteen factors were consequently used to build the prediction model.

4.2.2 LR model assessment

The logit model (logistic regression model) was built in R using the Generalized Linear Model (GLM) family. The model result is given in Annex 4.1. The accuracy of the model's outputs was assessed using the area under the receiver operating characteristic (ROC) curve (AUC) also plotted in R as shown in Figure 4.4. The training dataset recorded 93% accuracy while the validation dataset performed 92%.

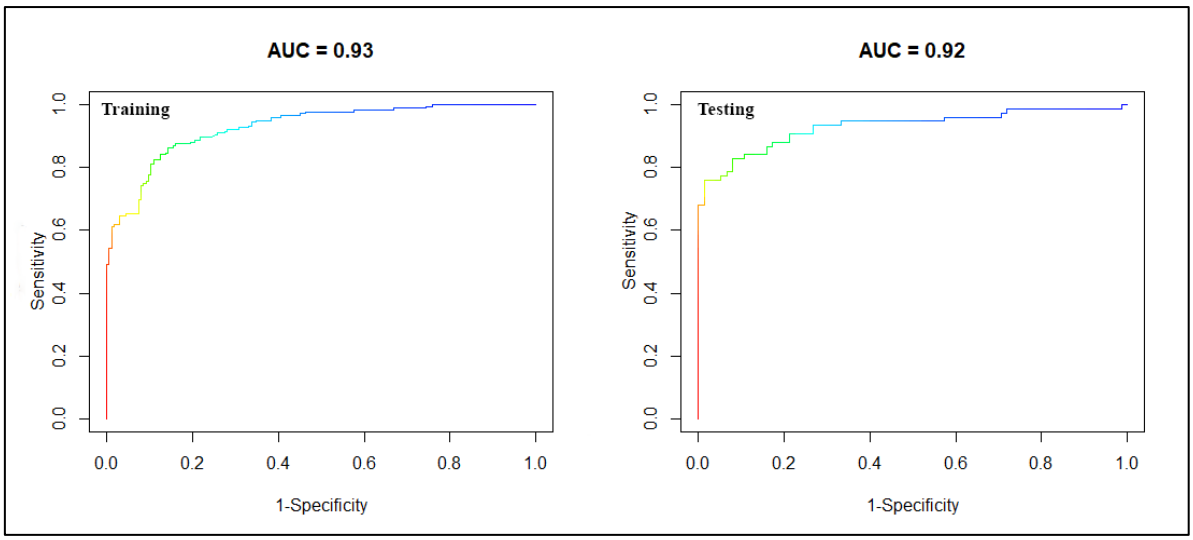

Figure 4.4. Logistic regression model's ROC curves

4.2.3 Flood suceptibility map by LR

The susceptibility map was developed in GIS following Equations (3.6) and (3.7). Equation 4.2 was thus obtained by replacing the intercepts and the values of the coefficients from Annex 4.1. Using Raster Calculator, the final map was derived and reclassified into four zones as shown in Figure 4.5.

$$Y = 4.8493342 + (-0.7348411) * \text{Slope} + (0.2018542) * \text{Aspect} + (-0.2094457) * \text{Curvature} + (0.7484634) * \text{Distance from Rivers} + (0.4734837) * \text{Drainage Density} + (-1.5360690) * \text{Elevation} + (0.0000005) * \text{Geology} + (-0.0000002) * \text{LULC} + (-0.0056608) * \text{NDVI} + (0.0839466) * \text{Rainfall} + (-0.0000001) * \text{Soil} + (-0.4638204) * \text{SPI} + (0.0611361) * \text{TWI} \quad (4.2)$$

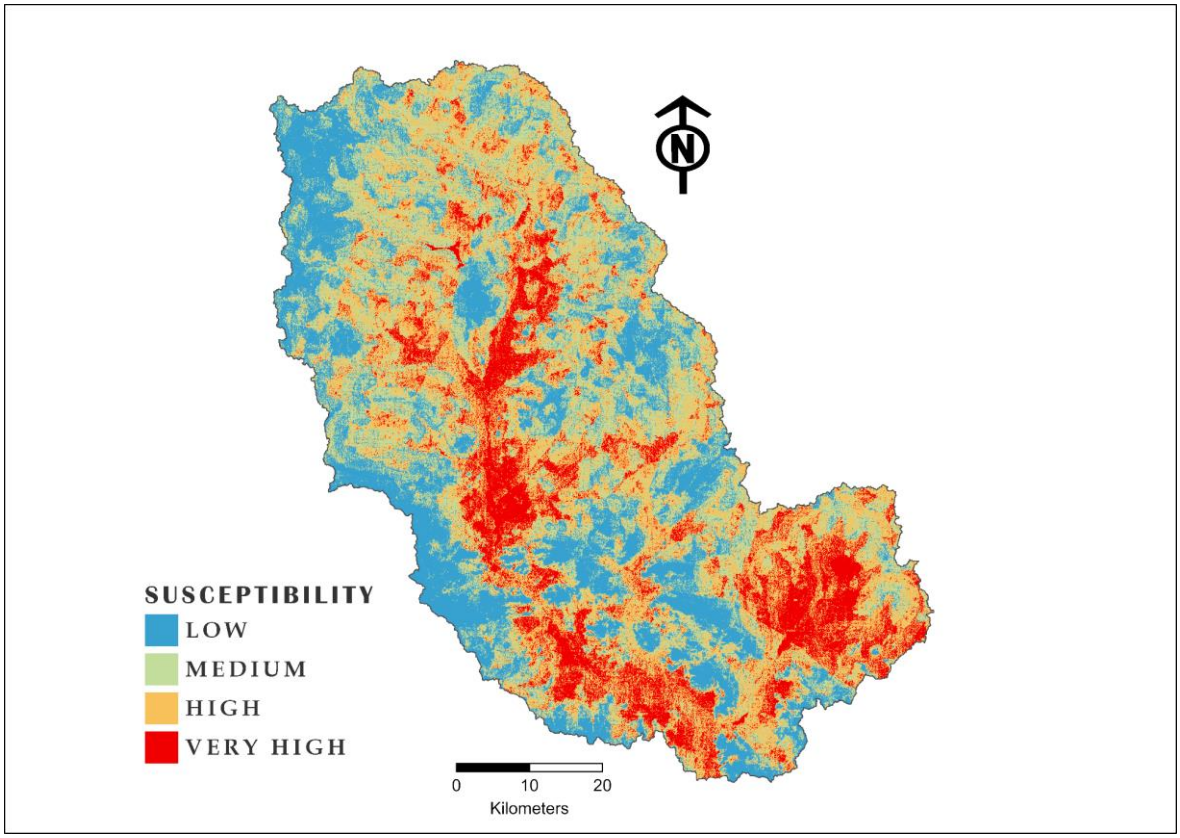

Figure 4.5. Flood susceptibility map with logit regression

The results show that the predicted map closely follows the trend of the elevation map, in that flat areas (Class 1 and 2) are the very highly susceptible areas.

4.2.4 Dominance Analysis

Dominance analysis is a method used to determine the relative importance of predictors in a regression analysis. Importance is defined as a qualitative comparison between pairs of predictors. In other terms, one factor is more important than another, if it contributes more to the prediction of the response variable in all possible subset models. The dominance analysis was performed in R using the recommended four indices, i.e., McFadden (r2.m), Cox and Snell (r2.cs), Nagelkerke (r2.n), and Estrella (r2.e), from the dominance analysis package.

Figure 4.6 illustrates the proportion of each conditioning factor to flood occurrence based on the Nagelkerke (r2.n) index. The elevation, the slope, and the TWI were found to be the most conditioning factors in the study area, with variable importance values of 23.6%, 16.8%, and 7.5%, respectively. This finding explains the verisimilitude between the predicted map and the elevation map. On the other hand, LULC and curvature with variable importance values less than 1%, do not seem to contribute much to the flooding in the study area. This finding was is a bit unexpected because different land uses, especially urbanization, may quickly lead to rapid land change which in turn contributes to extreme events. It is probably due to the nature of the analyzed data. In fact, class 4 of the LULC factor which is the build-up area represents only 1.06% of the total class and can only affect the model as much.

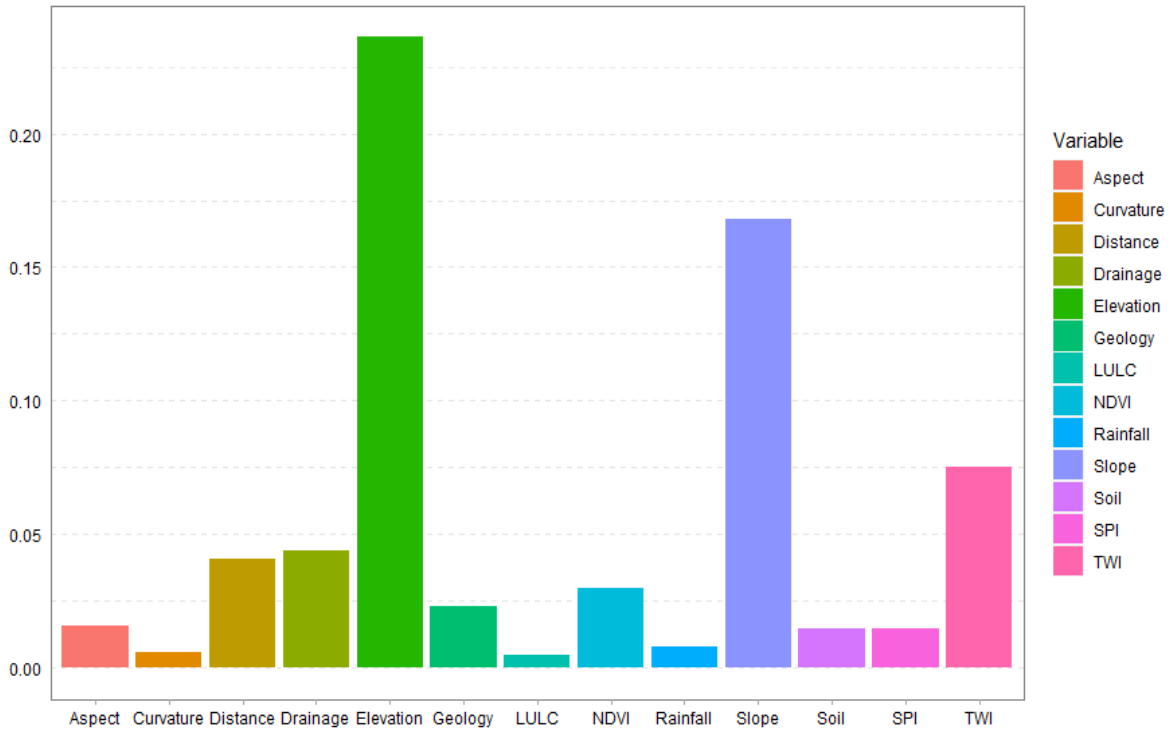

Figure 4.6. Variable importance of the flood conditioning factors

4.3 Model Comparison

The results show that the logistic regression model (AUC = 93%) by far performs better than the frequency ratio model (80.5%) in the study area. In this research, two models were used to explain the flood susceptibility of the region. Assessments of these models were produced to determine the appropriate flood susceptibility models. Evaluation of the models reveals that the AUC of LR is 93% while FR showed an AUC equals to 80.5%. It is thus evident that LR model has been the most effective approach for the Nakambé basin flood susceptibility study.

CHAPTER 5

"A man receives warning about the return of rains in the Somali Region of Ethiopia. Mobile phones are now increasingly used to communicate warnings and coordinate preparation activities…"

United Nations Office for Disaster Risk Reduction (UNDRR, 2020)

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 General Conclusions and Limitations

Floods are one of the greatest natural disasters that cannot be completely prevented. As the population grows, so do human encroachment and climate change. It is thus critical to consider and plan a flood vulnerability map in which appropriate management measures and policies for a given area can be implemented through scientific assessment. Different types of models were used for such analysis in many regions. For example, to produce a flood susceptibility map, Roy et al.(2020) used Support Vector Machine (SVM), Random Forest (RF), and Biogeography Based Optimization (BBO) model for flood susceptibility mapping in the Ajoy river basin, India, Rahmati et al. (2016) used RF model in Golestan watershed, Iran, Gudiyangada Nachappa et al. (2020) applied MCDA in Salzburg, Austria and Mind'je et al. (2019) applied FR for a national scale flood susceptibility mapping in Rwanda. In this analysis, two susceptibility models namely FR and LR were used for the susceptibility mapping of the Nakambé east basin. Although many factors have been researched in the literature for their contribution to flooding, not all factors are compatible with a single model. Even so, some factors might be insignificant in their contribution to flood susceptibility. Therefore, identifying suitable factors for the susceptibility study using some logical and experimental approaches is one of the foremost important tasks in a susceptibility study. In this research, thirteen (13) factors were selected based on an exhaustive review of previous studies. These factors include elevation, slope angle, slope aspect, curvature, distance from rivers, drainage density, Normalized Differential Vegetation Index (NDVI), Stream Power Index (SPI), Topographic Wetness Index (TWI), land use, and land cover (LULC), geology, soil classes. The built models were assessed and validated using ROC-AUC. LR was found to perform far better than FR with an AUC value equals 93% and compared to 80.5% for FR. This result is in agreement with that of Malik et al. (2020) of where the LR model was able to predict flood susceptibility with a precision of 91.6%. The ultimate advantage of machine learning models is not surprising. This is because it involves building an analytical system based on data training rather than on pre-defined explicit build rules (as is the case with statistical models).

Although this study can help planners to take appropriate measures, which will be discussed in the next section, it has some limitations. One of the limitations is the staticity of the

model. Indeed, the present study only considers static factors for the prediction. Further studies can explore the use of dynamic models, or the use of future scenarios for the predicting factors such as future rainfall scenarios estimated from Global Circulation Models (GCM) data, future LULC scenarios simulated from models such as land change modeler (LCM), Cellular Automata (CA), etc. In the same manner other factors such as the drainage, soil textures can be projected with suitable methods. The scenario outputs can then be incorporated into the model to find suitable algorithms for the region.

5.2 Contributions to Knowledge

This study can be used as a model for flood modelers working on complicated floodplain systems in data-scarce places like Burkina Faso, where limited flood records can make complex physically-based hydraulic models difficult to apply. The proposed method is a viable alternative for the hydraulic model. As a result, the flood danger maps developed have ramifications for flood managers, urban planners, and decision-makers. Given the study area's history of flooding events, immediate flood mitigation measures must be implemented.

It is however worth noting that our findings are solely applicable to the current case study and should be generalized with caution to other places.

5.3 Recommendations

Despite the heavy and continued economic losses resulting from floodings, many countries, particularly in Africa, are still lagging behind with regards to allocating resources for mitigation or prevention. In fact, domestic resources devoted to this cause are underfinanced. Evidence from sixteen African countries from the United Nations Office for Disaster Risk Reduction (UNDRR) shows that, on average, only four percent of national planned budgets is related to disaster risk reduction (UNDRR, 2020). Figure 5.1 illustrates the 16 country risk-sensitive budget review reports developed by UNDRR in 2020.

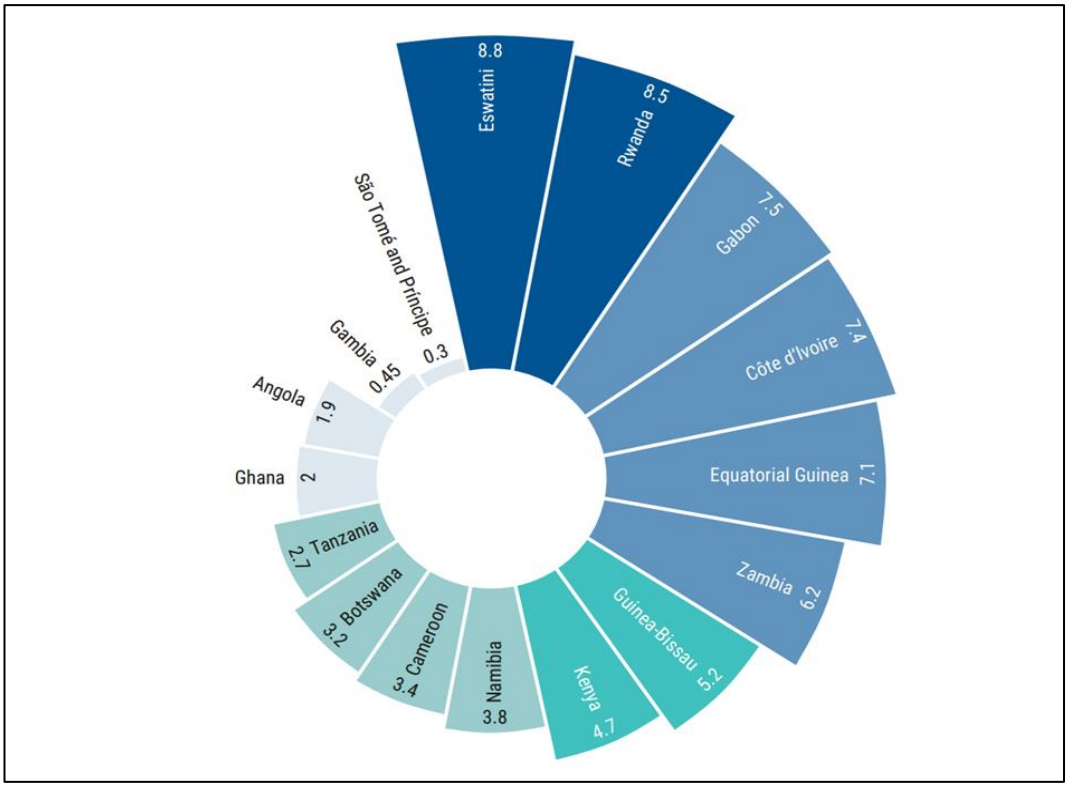

Figure 5.1. Proportion of national budgets spent on disaster risk reduction (source: UNDRR) The budget reviews cover three to five financial years depending on the country (except for that of Cameroon, which covers only one financial year).

It is worth noting that despite the recurrent flooding events in Burkina Faso, the country is not among the 16 countries in Africa. This fact highlights the dire need for investment. Notwithstanding the challenging context of persistent technical and financial capacity constraints in this country, few measures can be proposed for mitigation and disaster risk reduction. As it is with most disasters, an effective risk management should be broken down into major phases such as anticipation, crisis management, and prevention or post-crisis management. Most flood risk management strategies in our countries can be likened to the policy of *"le médecin après la mort"*, a french proverb that portrays the lateness in taking action. In other words, most strategies are focus on crisis management or tackling the effect of the disaster after it has happened. An effective strategy should start with anticipation.

5.3.1 Forecasting and anticipation

This is a process based on modeling phenomena and regular observations of typical variables. Anticipating a phenomenon makes it possible to prepare the means to tackle the

situation. In a practical sense, it is at this stage that the weather services intervene to anticipate the meteorological event and the flood warning services for the flood announcement in case of exceeding the warning numbers. Alert messages should be effectively communicated to the population and the competent authorities.

5.3.2 Crisis Management

This stage aims to manage the ongoing crisis, to ensure the safety of people and goods. The competent authorities should have the resources at hand to better manage this phase. Most time, it is after establishing the required resources that the authorities begin to seek funds as was the case with the Permanent Secretariat of the National Council for Emergency Relief and Rehabilitation (SP/CONASUR) during the 2020's flooding event. For the emergency response plan, a financial need of 9.8 billion CFA francs was estimated. The government declared the state of natural disaster and decided to release 5 billion CFA francs. The deficit was still to be covered (Bakouan & Knutson, 2020). This situation highlights the need for a consistent national planned budget related to disaster risk reduction and management.

5.3.3 Post-crisis management and prevention

By drawing lessons from past experiences it is possible to make recommendations to limit the effects of a similar event in the future. In this phase, research such as the present study can play a significant role. In fact, from the past floodings, the research has established zones of high risk of inundation. It can be observed that many places labeled as very highly susceptible are closed to river networks. Thus the following measure can be taken concerning these highly susceptible areas.

- Disaster risk reduction at community level: most often the community is not associated in floodplain management, the decisions are of top-down type. An effective management must take into consideration the active engagement of the local community, thus the necessity of involving them at each stage of the decision-making process regarding flood risk management. Many community actors have crucial knowledge, experience, and capacities in terms of building resilience, and can develop novel techniques to decreasing the dangers they encounter constantly. Local

organizations can take a series of actions including community-based risk-informed early actions, mainstreaming risk reduction education, and strengthening application of indigenous knowledge.

- Controlling settlements in flood-prone areas: One of the best ways to reduce future flood damages is to prevent development from occurring on flood-prone lands. Controlling this new development combined with measures to reduce damages to existing development will curb the rising social and economic losses that result from floods. Zoning of such lands (parks, nature areas or ecological reserves, etc) will be an effective approach. The local authority must update the laws or legislative documents concerning settlements in flood areas and enforce these laws. Experiences from Burkina Faso where people came back to the inundable areas to resettle after they have been relocated by the authorities show the necessity to effectively enforce the laws.
- In urban areas, it has been noticed that the conditions of the drainage are one of the main causes of flooding. Therefore, cleaning and emptying the drainage canal at the end of the dry season will prove to be an effective measure of mitigation. The canals can also be widened if need be, in order to increase the cross-sectional area so that the drainage system will be able to drain water.

6 REFERENCES

Ahmadlou, M., Karimi, M., Alizadeh, S., Shirzadi, A., Parvinnejhad, D., Shahabi, H., & Panahi, M. (2019). Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). *Geocarto International*, *34*(11), 1252–1272. https://doi.org/10.1080/10106049.2018.1474276

Ahmadlou, Mohammad, Al-Fugara, A., Al-Shabeeb, A. R., Arora, A., Al-Adamat, R., Pham, Q. B., Al-Ansari, N., Linh, N. T. T., & Sajedi, H. (2021). Flood susceptibility mapping and assessment using a novel deep learning model combining multilayer perceptron and autoencoder neural networks. *Journal of Flood Risk Management*, *14*(1), 1–22. https://doi.org/10.1111/jfr3.12683

Ali, A. bin M. (2018). Flood inundation modeling and hazard mapping under uncertainty in the Sungai Johor basin, Malaysia. In *Taylor & Francis Group*. Delft University of Technology.

Amadou, A. (2019). *Statistical analysis and flood frequency analysis for conceiving flood management Options in the Volta Basin (Case of Nakambe River Sub-Basin in Burkina Faso)*. Pan-African University Institute for Water and Energy Sciences (including Climate Change).

Arabameri, A., Rezaei, K., Cerdà, A., Conoscenti, C., & Kalantari, Z. (2019). A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran. *Science of the Total Environment*, *660*, 443–458. https://doi.org/10.1016/j.scitotenv.2019.01.021

Asare-Kyei, D., Forkuor, G., & Venus, V. (2015). Modeling flood hazard zones at the sub-district level with the rational model integrated with GIS and remote sensing approaches. *Water (Switzerland)*, *7*(7), 3531–3564. https://doi.org/10.3390/w7073531

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., & Prat, O. P. (2015). PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, *96*(1), 69–83.

Bakouan, F. Y., & Knutson, K. (2020). *BURKINA FASO - Inondations 2020 Rapport de situation N°01 11 septembre 2020*.

Bates, P. D. (2012). Integrating remote sensing data with flood inundation models: How far

have we got? In *Hydrological Processes* (Vol. 26, Issue 16, pp. 2515–2521). https://doi.org/10.1002/hyp.9374

BBC. (n.d.). *Million hit by floods in Africa*. Retrieved July 19, 2021, from http://news.bbc.co.uk/2/hi/africa/6998651.stm

Braman, L. M., van Aalst, M. K., Mason, S. J., Suarez, P., Ait-Chellouche, Y., & Tall, A. (2013). Climate forecasts in disaster management: Red Cross flood operations in West Africa, 2008. *Disasters*, *37*(1), 144–164.

Brunner, G. W. (2002). Hec-ras (river analysis system). *North American Water and Environment Congress \& Destructive Water*, 3782–3787.

Bui, D. T., Tsangaratos, P., Ngo, P. T. T., Pham, T. D., & Pham, B. T. (2019). Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. *Science of the Total Environment*, *668*, 1038–1054. https://doi.org/10.1016/j.scitotenv.2019.02.422

Chapi, K., Singh, V. P., Shirzadi, A., Shahabi, H., Bui, D. T., Pham, B. T., & Khosravi, K. (2017). A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environmental Modelling and Software*, *95*, 229–245. https://doi.org/10.1016/j.envsoft.2017.06.012

Chen, H., Liang, Q., Liu, Y., & Xie, S. (2018). Hydraulic correction method (HCM) to enhance the efficiency of SRTM DEM in flood modeling. *Journal of Hydrology*, *559*, 56–70. https://doi.org/10.1016/J.JHYDROL.2018.01.056

CHEN, M., WU, Y., LU, W., ZHU, J., & YANG, H. (2008). Characteristics and application of InfoWorks RS and FloodWorks software. *Hydro-Science and Engineering*, *4*.

Chen, W., Hong, H., Li, S., Shahabi, H., Wang, Y., Wang, X., & Ahmad, B. Bin. (2019). Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. *Journal of Hydrology*, *575*(February), 864–873. https://doi.org/10.1016/j.jhydrol.2019.05.089

Costache, R., & Tien Bui, D. (2019). Spatial prediction of flood potential using new ensembles of bivariate statistics and artificial intelligence: A case study at the Putna river catchment of Romania. In *Science of the Total Environment* (Vol. 691, pp. 1098–1118). https://doi.org/10.1016/j.scitotenv.2019.07.197

Di Baldassarre, G. (2010). Floods in a changing climate: Inundation modelling. In *Floods in a Changing Climate: Inundation Modelling*. https://doi.org/10.1017/CBO9781139088411

Di Baldassarre, G., Montanari, A., Lins, H., Koutsoyiannis, D., Brandimarte, L., & Blschl,

G. (2010). Flood fatalities in Africa: From diagnosis to mitigation. *Geophysical Research Letters*, *37*(22). https://doi.org/10.1029/2010GL045467

Diello, P., Mahe, G., Paturel, J.-E., Dezetter, A., Delclaux, F., Servat, E., & Ouattara, F. (2005). Relations indices de Végétation–Pluie au Burkina Faso: Cas du Bassin Versant du Nakambé. *Hydrological Sciences Journal*, *50*(2). https://doi.org/10.1623/hysj.50.2.207.61797

Dottori, F., Martina, M. L. V., & Figueiredo, R. (2018). A methodology for flood susceptibility and vulnerability analysis in complex flood scenarios. *Journal of Flood Risk Management*, *11*, S632–S645. https://doi.org/10.1111/jfr3.12234

Dragičević, N., Karleuša, B., & Ožanić, N. (2018). Improvement of Drainage Density Parameter Estimation within Erosion Potential Method. *Proceedings*, *2*(11), 620. https://doi.org/10.3390/proceedings2110620

Falconer, R. A., & Lin, B. (2001). Depth Integrated Velocity and Solute Transport (DIVAST) Model Reference Manual. *Hydro-Environmental Research Centre, School of Engineering, Cardiff University*.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & others. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 1–21.

Galland, J.-C., Goutal, N., & Hervouet, J.-M. (1991). TELEMAC: A new numerical model for solving shallow water equations. *Advances in Water Resources*, *14*(3), 138–148.

Gandhi, S. M., Sarkar, B. C., Infrared, T., Scales, L., & Deposits, A. M. (2016). Chapter 4 - Remote Sensing Techniques. In S. M. Gandhi & B. C. Sarkar (Eds.), *Essentials of Mineral Exploration and Evaluation* (pp. 81–95). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-805329-4.00011-9

Google Earth Engine. (n.d.). Retrieved July 15, 2021, from https://earthengine.google.com/#intro

Gudiyangada Nachappa, T., Tavakkoli Piralilou, S., Gholamnia, K., Ghorbanzadeh, O., Rahmati, O., & Blaschke, T. (2020). Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. In *Journal of Hydrology* (Vol. 590, p. 125275). Elsevier. https://doi.org/10.1016/j.jhydrol.2020.125275

Havnø, K., Madsen, M. N., Dørge, J., & others. (1995). MIKE 11-a generalized river modelling package. *Computer Models of Watershed Hydrology.*, 733–782.

Hong, H., Panahi, M., Shirzadi, A., Ma, T., Liu, J., Zhu, A. X., Chen, W., Kougias, I., & Kazakis, N. (2018). Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution. *Science of the Total Environment*, *621*, 1124–1141. https://doi.org/10.1016/j.scitotenv.2017.10.114

Huang, X., Tan, H., Zhou, J., Yang, T., Benjamin, A., Wen, S. W., Li, S., Liu, A., Li, X., Fen, S., & Li, X. (2008). Flood hazard in Hunan province of China: An economic loss analysis. *Natural Hazards*, *47*(1), 65–73. https://doi.org/10.1007/s11069-007-9197-z

Huisman, O., De By, R. A., & others. (2009). Principles of geographic information systems. In *ITC Educational Textbook Series* (Vol. 1).

Janin, J. M., Lepeintre, F., & Pechon, P. (1992). Telemac-3D: A finite element code to solve 3D free surface flow problems. In *Computer modelling of seas and coastal regions* (pp. 489–506). Springer.

Jerome Glago, F. (2021). Flood Disaster Hazards; Causes, Impacts and Management: A State-of-the-Art Review. In *Natural Hazards - Impacts, Adjustments and Resilience*. IntechOpen. https://doi.org/10.5772/intechopen.95048

Jha, A. K., Bloch, R., & Lamond, J. (2012). Cities and Flooding : A Guide to Integrated Urban Flood Risk Management for the 21st Century. In *Cities and Flooding*. World Bank. https://doi.org/10.1596/978-0-8213-8866-2

Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F., & Nasseri, M. (2019). A new approach to flood susceptibility assessment in data-scarce and ungauged regions based on GIS-based hybrid multi criteria decision-making method. *Journal of Hydrology*, *572*(February), 17–31. https://doi.org/10.1016/j.jhydrol.2019.02.034

Khosravi, K., Pham, B. T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I., & Tien Bui, D. (2018). A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Science of The Total Environment*, *627*, 744–755. https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.01.266

Khosravi, K., Shahabi, H., Pham, B. T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H. B., Gróf, G., Ho, H. L., Hong, H., Chapi, K., & Prakash, I. (2019). A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. *Journal of Hydrology*, *573*, 311–323. https://doi.org/10.1016/j.jhydrol.2019.03.073

Kimerling, A. J. (2012). *Map use : reading, analysis, interpretation*. Esri Press Academic.

Knight, D., Shamseldin, A., & Shamseldin, A. (2005). River Basin Modelling for Flood Risk Mitigation. In *River Basin Modelling for Flood Risk Mitigation* (1st Editio). CRC Press. https://doi.org/10.1201/9781439824702

Lee, S., Kim, J. C., Jung, H. S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomatics, Natural Hazards and Risk*, *8*(2), 1185–1203. https://doi.org/10.1080/19475705.2017.1308971

Malik, S., Chandra Pal, S., Chowdhuri, I., Chakrabortty, R., Roy, P., & Das, B. (2020). Prediction of highly flood prone areas by GIS based heuristic and statistical model in a monsoon dominated region of Bengal Basin. *Remote Sensing Applications: Society and Environment*, *19*(June), 100343. https://doi.org/10.1016/j.rsase.2020.100343

Mattivi, P., Franci, F., Lambertini, A., & Bitelli, G. (2019). TWI computation: a comparison of different open source GISs. *Open Geospatial Data, Software and Standards 2019 4:1*, *4*(1), 1–12. https://doi.org/10.1186/S40965-019-0066-Y

McCabe, M., Manfreda, S., Jarihani, B., Hawker, L., Bates, P., Neal, J., & Rougier, J. (2018). Perspectives on Digital Elevation Model (DEM) Simulation for Flood Modeling in the Absence of a High-Accuracy Open Access Global DEM. *Perspectives on Digital Elevation Model (DEM) Front. Earth Sci*, *6*, 233. https://doi.org/10.3389/feart.2018.00233

MECV. (2007). *Programme d'action National d'adaptation à la Variabilité et aux Changements Climatiques (Pana du Burkina Faso)*.

MEE. (2001). *Gestion Intégrée des Ressources en Eau: Etat des lieux des ressources en eau du Burkina Faso et de leur cadre de gestion*.

Miller, A. J. (1990). Flood hydrology and geomorphic effectiveness in the central Appalachians. *Earth Surface Processes and Landforms*, *15*(2), 119–134. https://doi.org/10.1002/ESP.3290150203

Mind'je, R., Li, L., Amanambu, A. C., Nahayo, L., Nsengiyumva, J. B., Gasirabo, A., & Mindje, M. (2019). Flood susceptibility modeling and hazard perception in Rwanda. *International Journal of Disaster Risk Reduction*. https://doi.org/10.1016/j.ijdrr.2019.101211

Moel, H. de, Alphen, J. van, & Aerts, J. (2009). Flood maps in Europe--methods, availability and use. *Natural Hazards and Earth System Sciences*, *9*(2), 289–301.

Nachappa, T. G., Ghorbanzadeh, O., Gholamnia, K., & Blaschke, T. (2020). Multi-hazard exposure mapping using machine learning for the state of Salzburg, Austria. *Remote*

Sensing, *12*(17), 1–24. https://doi.org/10.3390/RS12172757

Nerem, R. S., Beckley, B. D., Fasullo, J. T., Hamlington, B. D., Masters, D., & Mitchum, G. T. (2018). Climate-change–driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(9), 2022–2025. https://doi.org/10.1073/pnas.1717312115

Niraj Baral, Akhilesh Kumar Karna, & Suraj Gautam. (2021). Landslide Susceptibility Assessment Using Modified Frequency Ratio Model in Kaski District, Nepal. *International Journal of Engineering and Management Research*, *11*(1), 167–177. https://doi.org/10.31033/IJEMR.11.1.23

Paeth, H., Fink, A. H., Pohle, S., Keis, F., Mächel, H., & Samimi, C. (2011). Meteorological characteristics and potential causes of the 2007 flood in sub-Saharan Africa. *International Journal of Climatology*, *31*(13), 1908–1926.

Pender, G. (2006). Briefing: Introducing the flood risk management reserch consortium. *Proceedings of the Institution of Civil Engineers: Water Management*, *159*(1), 3–8. https://doi.org/10.1680/wama.2006.159.1.3

Prasad, A. S., & Francescutti, L. H. (2016). Natural Disasters. In *International Encyclopedia of Public Health* (pp. 215–222). Elsevier Inc. https://doi.org/10.1016/B978-0-12-803678-5.00519-1

Rahmati, O., Pourghasemi, H. R., & Zeinivand, H. (2016). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto International*, *31*(1), 42–70. https://doi.org/10.1080/10106049.2015.1041559

Ritchie, H., & Roser, M. (2014). Natural Disasters: Our World in Data. In *University of Oxford* (pp. 1–2). https://ourworldindata.org/natural-disasters

Roy, P., Chandra Pal, S., Chakrabortty, R., Chowdhuri, I., Malik, S., & Das, B. (2020). Threats of climate and land use change on future flood susceptibility. *Journal of Cleaner Production*, *272*, 122757. https://doi.org/10.1016/j.jclepro.2020.122757

Sagris, V., Lavalle, C., & Petrov, L. (2005). *Towards an Integrated Assessment of Climate Change-Induced Sea-Level Rise in the Baltic Sea: An Example for the City of Pärnu (Estonia)*. https://publications.jrc.ec.europa.eu/repository/handle/JRC31833

Samanta, R. K., Bhunia, G. S., Shit, P. K., & Pourghasemi, H. R. (2018). Flood susceptibility mapping using geospatial frequency ratio technique: a case study of Subarnarekha River Basin, India. *Modeling Earth Systems and Environment*, *4*(1), 395–408. https://doi.org/10.1007/s40808-018-0427-z

Şen, Z. (2017). *Flood modeling, prediction and mitigation* (1st ed. 20). Springer. https://doi.org/10.1007/978-3-319-52356-9

Shafapour Tehrany, M., Kumar, L., Neamah Jebur, M., & Shabani, F. (2019). Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. *Geomatics, Natural Hazards and Risk*, *10*(1), 79–101. https://doi.org/10.1080/19475705.2018.1506509

Shafapour Tehrany, M., Shabani, F., Neamah Jebur, M., Hong, H., Chen, W., & Xie, X. (2017). GIS-based spatial prediction of flood prone areas using standalone frequency ratio, logistic regression, weight of evidence and their ensemble techniques. *Geomatics, Natural Hazards and Risk*. https://doi.org/10.1080/19475705.2017.1362038

Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K., & Shirzadi, A. (2018). Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. In *Journal of Environmental Management* (Vol. 217, pp. 1–11). https://doi.org/10.1016/j.jenvman.2018.03.089

Shaw, J., Kesserwani, G., Neal, J., Bates, P., & Sharifian, M. K. (2021). LISFLOOD-FP 8.0: the new discontinuous Galerkin shallow-water solver for multi-core CPUs and GPUs. *Geoscientific Model Development*, *14*(6), 3577–3602.

Solomatine, D., See, L. M., & Abrahart, R. J. (2008). Data-Driven Modelling: Concepts, Approaches and Experiences. *Practical Hydroinformatics*, 17–30. https://doi.org/10.1007/978-3-540-79881-1_2

Store, C. C. C. S. C. D. (2017). *Copernicus Climate Change Service (C3 S)(2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate*.

Swain, K. C., Singha, C., & Nayak, L. (2020). Flood Susceptibility Mapping through the GIS-AHP Technique Using the Cloud. *ISPRS International Journal of Geo-Information*, *9*(12), 720. https://doi.org/10.3390/ijgi9120720

Tehrany, M. S., Jones, S., & Shabani, F. (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena*, *175*(December 2018), 174–192. https://doi.org/10.1016/j.catena.2018.12.011

Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *Journal of Hydrology*, *504*, 69–79. https://doi.org/10.1016/j.jhydrol.2013.09.034

Traore, R. (2012). *Eau, Territoire et Conflits : analyse des enjeux de la gestion communautaire de l'eau au Burkina Faso : l'exemple du bassin versant du Nakambé*.

Université de Toulouse.

UNDRR. (2020). *Highlights : Africa Regional Assessment Report 2020*.

UNISDR. (2002). Guidelines for Reducing Flood Losses. *United Nations - Headquarters (UN), Available on-Line at: Https://Www.Unisdr.Org/We/Inform/Publications/558*, 79.

United Nation of Outer Space Affairs. (2020). *Step-by-Step: Recommended Practice: Flood Mapping and Damage Assessment using Sentinel-1 SAR data in Google Earth Engine*. United Nation of Outer Space Affairs. https://un-spider.org/advisory-support/recommended-practices/recommended-practice-google-earth-engine-flood-mapping/step-by-step

Vojtek, M., & Vojteková, J. (2019). Flood Susceptibility Mapping on a National Scale in Slovakia Using the Analytical Hierarchy Process. *Water*, *11*(2), 364. https://doi.org/10.3390/w11020364

Vojtek, M., Vojteková, J., Costache, R., Pham, Q. B., Lee, S., Arshad, A., Sahoo, S., Linh, N. T. T., & Anh, D. T. (2021). Comparison of multi-criteria-analytical hierarchy process and machine learning-boosted tree models for regional flood susceptibility mapping: a case study from Slovakia. *Geomatics, Natural Hazards and Risk*, *12*(1), 1153–1180. https://doi.org/10.1080/19475705.2021.1912835

Wang, Y., Hong, H., Chen, W., Li, S., Panahi, M., Khosravi, K., Shirzadi, A., Shahabi, H., Panahi, S., & Costache, R. (2019). Flood susceptibility mapping in Dingnan County (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. *Journal of Environmental Management*, *247*(September 2018), 712–729. https://doi.org/10.1016/j.jenvman.2019.06.102

WMO. (2015). *The Global Climate in 2011–2015* (Issue 1179). Deutscher Wetterdienst.

Xiao, T., Segoni, S., Chen, L., Yin, K., & Casagli, N. (2020). A step beyond landslide susceptibility maps: a simple method to investigate and explain the different outcomes obtained by different approaches. *Landslides*, *17*(3), 627–640. https://doi.org/10.1007/s10346-019-01299-0

Xiao, T., Yin, K., Yao, T., & Liu, S. (2019). Spatial prediction of landslide susceptibility using GIS-based statistical and machine learning models in Wanzhou County, Three Gorges Reservoir, China. *Acta Geochimica*, *38*(5), 654–669. https://doi.org/10.1007/s11631-019-00341-1

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global

terrain elevations. *Geophysical Research Letters*, *44*(11), 5844–5853. https://doi.org/10.1002/2017GL072874

Zhu, A.-X., Zhao, F.-H., Liang, P., & Qin, C.-Z. (2021). Next generation of GIS: must be easy. *Annals of GIS*, *27*(1), 71–86.

7 APPENDIX

Annex 4.1. Logit resession model outputs

Factors	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	4.8493342	2.0395799	2.3776143	0.0174250
Aspect	0.2018542	0.0718496	2.8093973	0.0049634
Curvature	-0.2094457	0.1841952	-1.1370857	0.2555025
Distance	0.7484634	0.2470254	3.0299040	0.0024463
Drainage	0.4734837	0.1970316	2.4030856	0.0162574
Elevation	-1.5360690	0.2442677	-6.2884646	0.0000000
Geology	0.0000005	0.0000002	1.9484932	0.0513560
LULC	-0.0000002	0.0000001	-1.6211359	0.1049885
NDVI	-0.0056608	0.2610151	-0.0216875	0.9826972
Rainfall	0.0839466	0.2138819	0.3924906	0.6946958
Slope	-0.7348411	0.1895490	-3.8767878	0.0001058
Soil	-0.0000001	0.0000002	-0.8521105	0.3941528
SPI	-0.4638204	0.3549433	-1.3067448	0.1912994
TWI	0.0611361	0.3393553	0.1801536	0.8570320

Annex 4.2. R script for the logit model

````
---
title: "Logistic Regression"
output: html_notebook
author: "Benjamin Bonkoungou"
date: "13/08/2021"
editor_options:
  chunk_output_type: console
---

**Create Working Folder**

```{r}
dir.create("D:/COURSES/Thesis/LR")
dir.create("D:/COURSES/Thesis/LR/Library")
dir.create("D:/COURSES/Thesis/LR/Data")
setwd("D:/COURSES/Thesis/LR")
.libPaths("D:/COURSES/Thesis/LR/Library")
.libPaths()
```

**Installing the packages**

```{r}
packages = c("raster", "rgeos", "rgdal", "maptools", "ROCR",
"dominanceanalysis", "readxl", "ggplot2")
ipak = function(pkg){
    new.pkg = pkg[!(pkg %in% installed.packages()[, "Package"])]
    if (length(new.pkg))
        install.packages(new.pkg, dependencies = TRUE)
    sapply(pkg, require, character.only = TRUE)
}
ipak(packages)
```
````

Loading the data exported from ArcGIS

```{r}
library(raster)
Training = raster("Data/re_Training")
Elevation = raster("Data/re_Elevation.tif")
TWI = raster("Data/re_TWI.tif")
SPI = raster("Data/re_SPI.tif")
NDVI = raster("Data/re_NDVI.tif")
Slope = raster("Data/re_Slope.tif")
LULC = raster("Data/re_LULC.tif")
Soil = raster("Data/re_Soil.tif")
Geology = raster("Data/re_Geology.tif")
Curvature = raster("Data/re_Curvature.tif")
Drainage = raster("Data/re_DD.tif")
Distance = raster("Data/re_DR.tif")
Aspect = raster("Data/re_Aspect.tif")
Rainfall = raster("Data/re_Rainfall.tif")
```

Resampling the data to make sure they have the same extent

Then save them to a new folder, "Resampled" using "writeRaster" function

```{r}
Training = resample(Training, Elevation, resample='bilinear')
TWI = resample(TWI, Elevation, resample='bilinear')
SPI = resample(SPI, Elevation, resample='bilinear')
NDVI = resample(NDVI, Elevation, resample='bilinear')
Slope = resample(Slope, Elevation, resample='bilinear')
LULC = resample(LULC, Elevation, resample='bilinear')
Soil = resample(Soil, Elevation, resample='bilinear')
Geology = resample(Geology, Elevation, resample='bilinear')
Curvature = resample(Curvature, Elevation, resample='bilinear')
Drainage = resample(Drainage, Elevation, resample='bilinear')
Distance = resample(Distance, Elevation, resample='bilinear')
```

```
Aspect = resample(Aspect, Elevation, resample='bilinear')
Rainfall = resample(Rainfall, Elevation, resample='bilinear')
```

Stacking the resampled data and extracting the values

```{r}
stackList = list.files(path = "D:\\COURSES\\Thesis\\LR\\Resampled",
                       pattern = "tif$", full.names = TRUE)
Rasters = stack(stackList)
valueTable = getValues(Rasters)
valueTable = na.omit(valueTable)
valueTable = as.data.frame(valueTable)
# Export the values to text file, that's for double check or  further calculations
# write.table(valueTable, "valueTable.txt", sep = "\t")
```

Running the Logistic Regression Model

```{r}
training.fit = glm(Training~., data = valueTable, family = binomial)
options("scipen" = 100)  # display the results
summary(training.fit)
Predicted = predict(training.fit, type = "response") # Predicted values
#write.table(Predicted, "predictedValues.txt", sep="\t") # Export the predicted values
table(valueTable$Training, Predicted > 0.5) # Check the confusion matrix
```

Ploting the area under curve (AUC)

```{r}
library(ROCR)
# ROC summarizes the predictive power for all possible values of p > 0.5.
ROCRpred = prediction(Predicted, valueTable$Training)
ROCRperf = performance(ROCRpred, 'tpr','fpr')
plot(ROCRperf)
```

```
#The AUC is a perfect performance metric for ROC curve.
AUC = performance(ROCRpred, measure = "auc")
AUC = AUC@y.values[[1]]
print(AUC)
plot(ROCRperf, main = "AUC = 0.93", colorize = TRUE, text.adj = c(-0.2, 1.7))
```

Prediction Map

```{r}
Y = 4.8493342 + (-0.7348411)*Slope + (0.2018542)*Aspect + (-0.2094457)*Curvature + (0.7484634)*Distance + (0.4734837)*Drainage + (-1.5360690)*Elevation + (0.0000005)*Geology + (-0.0000002)*LULC + (-0.0056608)*NDVI + (0.0839466)*Rainfall + (-0.0000001)*Soil + (-0.4638204)*SPI + (0.0611361)*TWI

P = 1/(1+exp(Y*(-1)))
plot(P, col = colorRampPalette(c("red", "red", "white"))(30))
```

Perform a Dominance Analysis

```{r}
library(dominanceanalysis)
da.glm.fit()("names")  # "r2.m"  "r2.cs" "r2.n"  "r2.e"
da.LR = dominanceAnalysis(training.fit)
print(da.LR)

averageContribution(da.LR, fit.functions = "r2.m")   # McFadden (r2.m)
plot(da.LR, which.graph ="general", fit.function = "r2.m")

averageContribution(da.LR, fit.functions = "r2.cs")   # Cox and Snell (r2.cs)
plot(da.LR, which.graph ="general", fit.function = "r2.cs")

averageContribution(da.LR, fit.functions = "r2.n")   # Nagelkerke (r2.n)
plot(da.LR, which.graph ="general", fit.function = "r2.n")
```

```
averageContribution(da.LR, fit.functions = "r2.e")   # Estrella (r2.e)
plot(da.LR, which.graph ="general", fit.function = "r2.e")

bootLR100 = bootDominanceAnalysis(glm.LR, R = 100)  # Applying booth
summary(bootLR100, fit.functions="r2.n")
print(format(summary(bootLR100)$r2.m, digits=3), row.names = F))
```

Dominance Analysis Graph

```{r}
library(ggplot2)
library(readxl)
df = read_excel("Dominance Analysis  Results (r2.m).xlsx") #Previously saved
colnames(df) = c("Variable", "Indice")
ggplot(data = df, aes(reorder(Variable, Indice), Indice, fill = Variable))+
geom_bar(stat = "identity") +
scale_colour_gradient2() +
xlab(" ") +ylab(" ") +
theme(panel.background = element_rect(fill = "white", colour = NA),
      panel.border = element_rect(fill = NA, colour = "gray50", size = rel(1)),
      panel.grid = element_line(colour = "grey90", linetype = "dashed"),
      panel.grid.major.x = element_blank(),
      axis.text = element_text(colour = "black"),
      axis.ticks = element_line(size = 0, colour = "grey65"),
      axis.ticks.y = element_blank())
```