Leveraging open-source remote-sensing data and machine learning approaches for flood prediction and susceptibility mapping in Nam Ngum River Basin (NNRB), Lao PDR

STUDENT NAME: MANGKHASEUM Sackdavong

STUDENT ID: 215F5005

SUPERVISOR: Associate Professor Akitoshi Hanazawa

Department of Electrical and Space Systems Engineering
Graduate School of Engineering
Kyushu Institute of Technology

Thesis submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy in Engineering
September 2024

ABSTRACT

Frequent floods caused by monsoons and rainstorms have significantly affected the resilience of human and natural ecosystems in the Nam Ngum River Basin, Lao PDR. A cost-efficient framework integrating advanced remote sensing and machine learning techniques is proposed to address this issue by enhancing flood susceptibility understanding and informed decision-making. This study utilizes remote sensing geo-datasets and machine learning algorithms (Random Forest, Support Vector Machine, Artificial Neural Networks, and Long Short-Term Memory) to generate comprehensive flood susceptibility maps. The results highlight Random Forest's superior performance, achieving the highest train and test Area Under the Curve of Receiver Operating Characteristic (AUROC) (1.00 and 0.993), accuracy (0.957), F1-score (0.962), and kappa value (0.91), with the lowest mean squared error (0.207) and Root Mean Squared Error (0.043). Vulnerability is particularly pronounced in low-elevation and low-slope southern downstream areas (Central part of Lao PDR). The results reveal that 36– 53% of the basin's total area is highly susceptible to flooding, emphasizing the dire need for coordinated floodplain management strategies. This research uses freely accessible remote sensing data, addresses data scarcity in flood studies, and provides valuable insights for disaster risk management and sustainable planning in Lao PDR

In addition, the generated flood susceptibility map is used to analyze the possible effect on the different land use/land cover classes, populations and critical facilities. ANN and DNN outperform LSTM, achieving higher accuracy based on Receiver Operating Characteristics. The resulting flood susceptibility maps identify critical zones within the Nam Ngum River Basin at high risk of flooding, revealing that 36-53% of the basin area is highly susceptible, especially in low-elevation and low-slope regions. Additionally, 85-93% of the population is highly vulnerable to flooding within 261 to 296 km² of built-up area. Almost all of the critical facilities for health and education lie within the area, which is highly susceptible to flooding.

Keywords: Flood susceptibility modeling; Flood risk assessment, Open-source Datasets, Machine learning algorithm; Deep learning algorithm, Remote sensing, Nam Ngum River Basin (NNRB), Lao PDR

ACKNOWLEDGEMENT

First and foremost, I would like to take this opportunity to express my deepest sense of gratitude and sincere thanks to my highly respected and esteemed supervisor, Assoc. **Prof. Akitoshi HANAZAWA**, for his valuable guidance, encouragement, and support in completing this thesis. He has enlightened me with timely suggestions and his inspiration in machine learning and deep learning technique in our group discussions. His continuous encouragement supports my research to a successful completion. Also, I would like to special thank **Assoc. Prof. Sunil Duwal** and **Yogesh Bhattarai** at Khwopa College of Engineering for the for their help in data analysis and suggestions. This dissertation would not have been possible without his patience and persistent help

Foremost, I gratefully acknowledge the scholarship support from Japanese Government for awarding me the **Monbukagakusho (MEXT)** scholarship and **The United Nations Office for Outer Space Affairs PNST fellowship program** for the financial support throughout my doctoral degree at Kyushu Institute of Technology, Japan

I would also like to express my sincere thanks to Kyushu Institute of Technology for giving me the opportunity and a good study environment to complete my doctoral degree and to our professors and Kyutech Staffs, who were very cooperative, helpful, and forever there for us.

Finally, I would like to express my sincere thanks to my lovely wife, parents, friends and families who helped me directly or indirectly to complete this doctoral thesis. Their trust, encouragement and understanding helped me overcome many obstacles during my pursuit of a Ph.D. degree. Finally, I thank myself for never giving up.

Sackdavong MANGKHASEUM

Kyushu Institute of Technology, September 2024

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

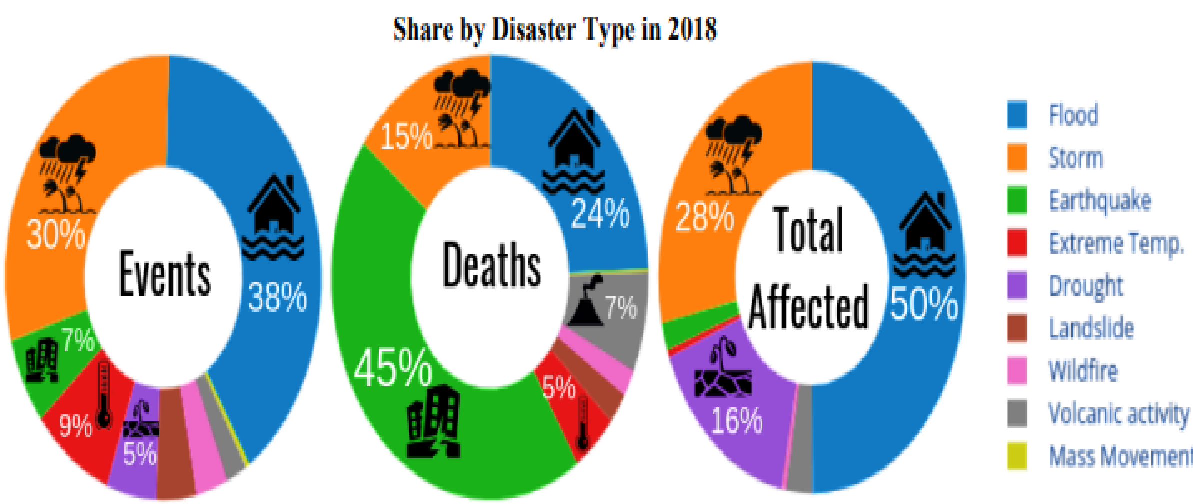
AHP	Analytical Hierarchy Process
ANFIS	Adaptive Neuro-Fuzzy Interface System
ANN	Artificial Neural Network
AUROC	Area Under the Curve of Receiver Operating Characteristic
DEM	Digital Elevation Model
DTR	Distance to River
DEMs	Digital Elevation Models
FR	Frequency Ratio
GIS	Geographic Information System
K4D	LAO Knowledge for Development
IGR	Information Gain Ratio
LN	Logistic Regression
LSTM	Long Short-Term Memory
LULC	Land use and land cover
MSE	Mean Square Error
MASL	Meters above sea level
MCDA	Multi-criteria decision analysis method
RMSE	Root Mean Square Error
NDVI	Normalized Difference Vegetation Index
NNRB	Nam Ngum River Basin
RS	Remote Sensing
RF	Random Forest
SAR	Synthetic Aperture Radar
SPI	Stream Power Index
SDGs	Sustainable Development Goals
SVM	Support vector machine
TWI	Topographic Roughness Index
WoE	Weights-of-Evidence
VV	Vertical Transmit and Vertical Receive
VIF	Variance Inflation Factor

CHAPTER 1

INTRODUCTION

1.1 Background of the study

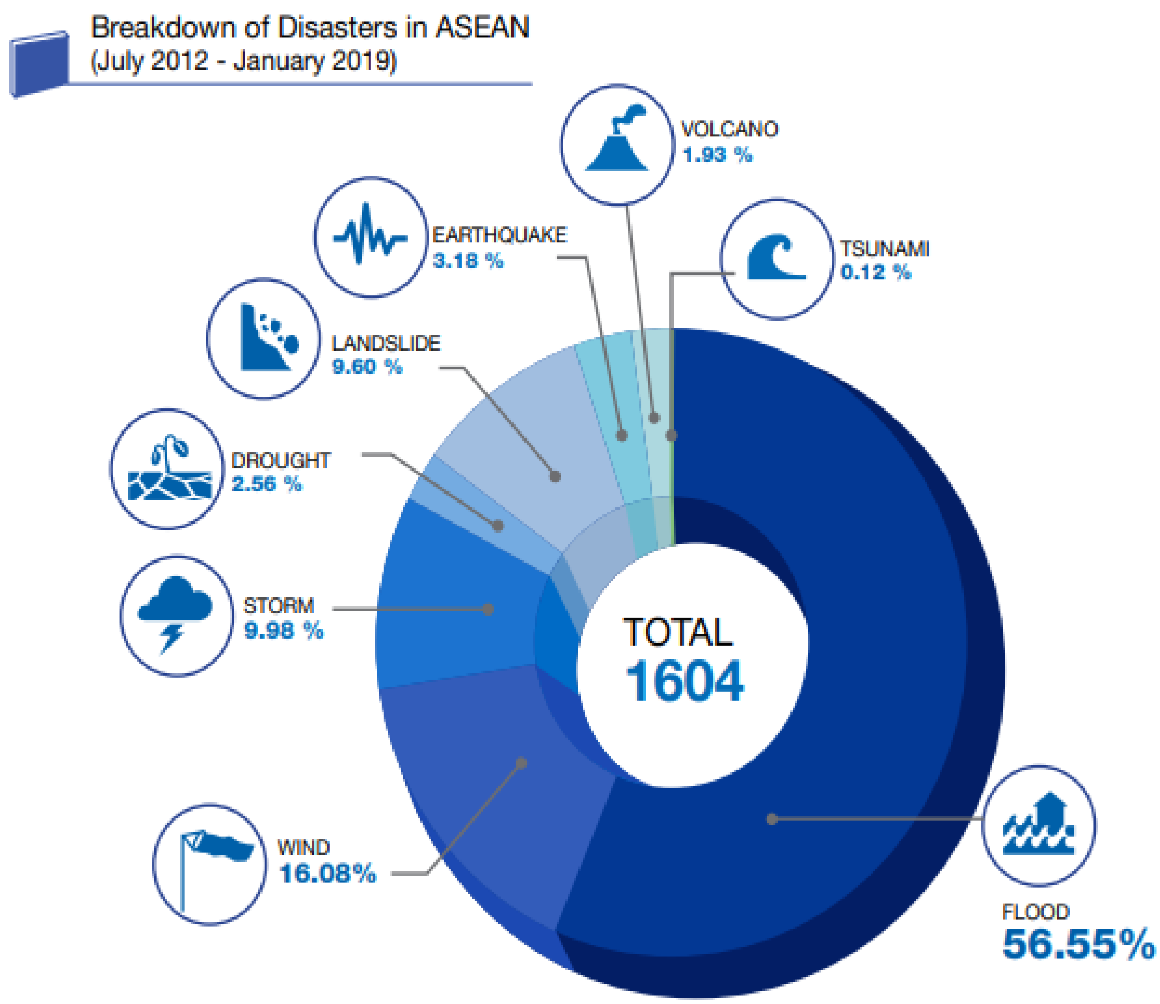
Natural disasters significantly affect human lives, property, and the environment, leading to economic losses. The Centre for Research on the Epidemiology of Disasters (CRED) reported in April 2019 that there were 315 disasters related to natural hazards worldwide in the previous year. These disasters resulted in 11,804 deaths, impacted more than 68.5 million people, and caused over US$131.7 billion in economic damages. The report highlighted that major disasters in Asia, South America, and Africa were predominantly floods and landslides. Floods accounted for 38% of these events, causing 24% of deaths and 50% of the economic damage, making them the costliest type of disaster. Additionally, storms constituted 30% of the events and were responsible for 15% of the total deaths. In 2018, Laos ranked fifth in the number of people affected per 100,910 inhabitants. Therefore, flooding is among the most recurrent natural disasters globally.



Source: The Centre for Research on the Epidemiology of Disasters (CRED)

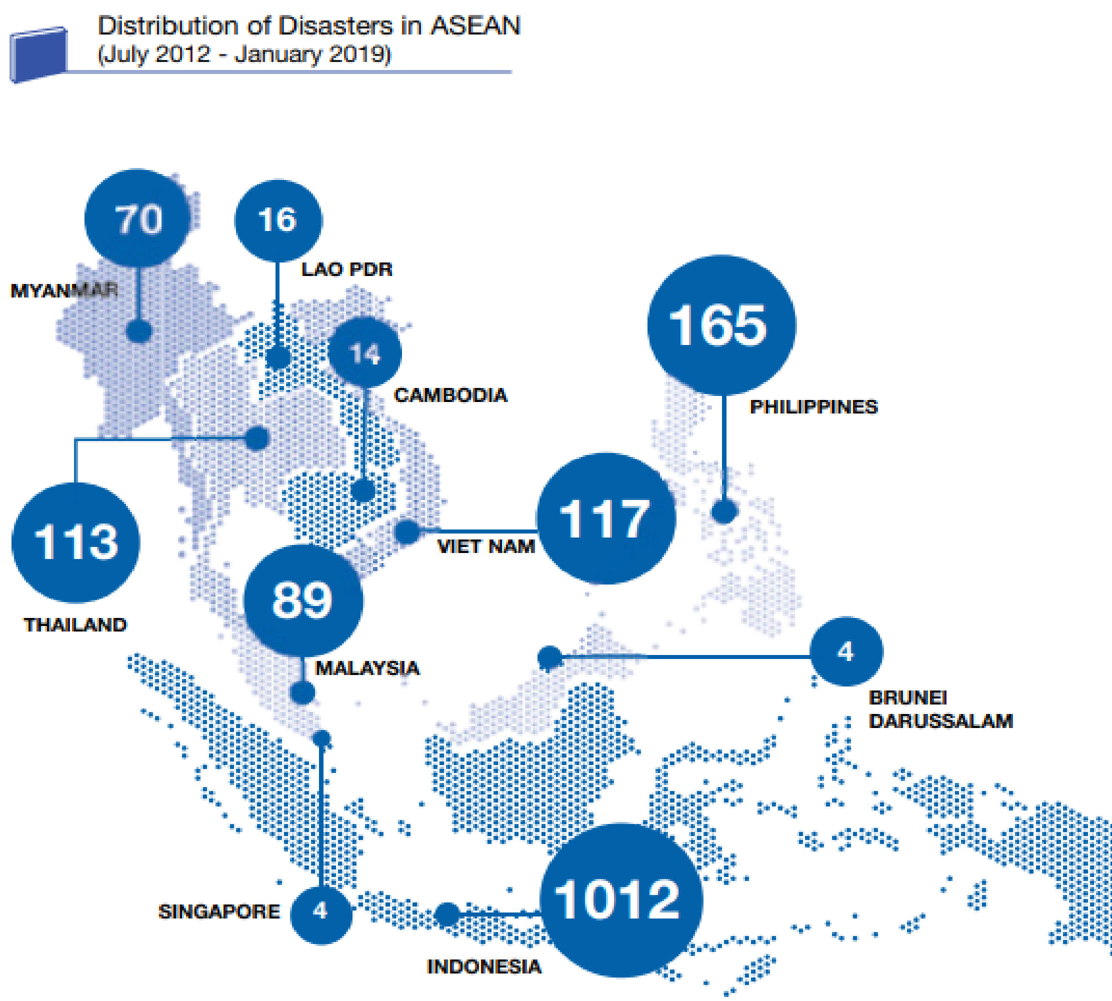
Figure 1. Depicts the global disaster type of events, death, and total affected

According to The ASEAN Risk Monitor and Disaster Management Review (ARMOR), ASEAN nations experienced a combined total of 1,604 disasters of varying severity between July 2012 and January 2019. Figure 1 shows that 85.17% of these were hydrological and meteorological disasters, including floods, strong winds, tropical storms, and droughts, occurring within that timeframe. In contrast, 14.83% were geophysical disasters, including landslides triggered by earthquakes, volcanic eruptions, and minor tsunamis. Furthermore, Figure 2 shows that Malaysia, Vietnam, Cambodia, Laos, Thailand, and Brunei Darussalam have the highest percentage of their population exposed to flooding.



Source: https://ahacentre.org/publication/armor/

Figure 2. Types of disasters occur in ASEAN



Source: https://ahacentre.org/publication/armor/

Figure 3. Percentage of Population Exposed to Floods (by country)

Floods in the Lao People's Democratic Republic (PDR) present a significant challenge, both environmentally and socioeconomically. Laos, a landlocked country in Southeast Asia, experiences recurrent flooding due to its tropical monsoon climate, characterized by heavy seasonal rainfall. The Mekong River, which flows through Laos, frequently overflows during the rainy season, causing widespread flooding.

These floods have severe impacts on the Lao population, agriculture, and economy. Annually, thousands of people are displaced, homes are destroyed, and agricultural lands are inundated, leading to substantial food insecurity. Agriculture, a major economic sector in Laos, is particularly vulnerable; floods often wash away crops, disrupt planting and harvesting schedules, and damage irrigation systems. Consequently, rural communities, which are heavily dependent on agriculture, face significant economic losses and hardships.

The Lao government, alongside international organizations, has been working to mitigate the impacts of floods through various measures. These include constructing dams and reservoirs, improving early warning systems, and implementing better land-use planning. Community-based approaches, such as educating locals on flood preparedness and response, are also crucial in reducing the adverse effects.

However, challenges remain due to limited resources and the increasing intensity of weather events linked to climate change. Continued efforts and international cooperation are essential to enhance resilience against floods and secure a safer future for the Lao PDR. To address this gap, this study proposes a simple and efficient data-driven machine-learning approach based on remote sensing data for effective flood susceptibility mapping in the NNRB. The findings of this study aid in informed decision-making for flood management and urban development. It aligns with the National Strategy on Disaster Risk Reduction (NSDRR) 2021–2030 (Government of Lao PDR and Asian Development Bank 2022), Sendai Framework for Disaster Risk Reduction, and sustainable development goals (SDGs-11 –sustainable cities and communities, and SDG-13 Climate change), aiming to address climate change and ensure safety in flood-prone areas.

1.2 Statement of the Problem

In recently, the rise in urbanization and land expansion has led to a notable increase in the frequency of floods. Lack of the inventory maps, absence of proper flood analyses, and interpretational difficulties are the main limitations in flood studies and subsequently urban planning (Sharma, Kumar, and Kumar 2024).

Detecting floods is the first step in flood susceptibility mapping, and it should be as rapid and accurate as possible. However, due to the presence of speckle noise in synthetic aperture radar (SAR) imageries (Anusha and Bharathi 2020), specular reflectance from other objects (Schlaffer et al. 2015) and spatial heterogeneity of urban areas, classification methods developed for optical images are often not adaptable for flood recognition and mapping (Shahabi et al. 2020). Visual interpretation is another method for flood detection which is based on expert's knowledge and can be biased (Mucsi and Bui 2023). Another method for mapping flood extents is the threshold segmentation algorithm, which is highly sensitive to images with low contrast and relies on expert judgment. The generated segments should be defined separately and individually for each imagery which makes this method not optimized for flood extent extraction (Hansana et al. 2023).

Flood susceptibility maps are the basis of further researches such as hazard and risk analysis (Pourghasemi et al. 2020). Governments allocate significant budgets to prevent flooding, yet the lack of precise flood forecasting and mapping persists. Based on the literature, most of the existing methods for flood analysis have few drawbacks which should be overcome (Liu & De Smedt, 2005). MCDA methods that have been used are the "Analytical hierarchy process". AHP relies on expert knowledge and contains many biases, which can be subjective (Das 2020). Hydrologic models like HEC-HMS and HEC-RAS are excellent for simulating flood scenarios (Zeleňáková et al. 2019), but they require accurate data and deep hydrology knowledge (Costache and Tien Bui 2019). Bivariate statistical methods such as frequency ratio (FR) and weight of evidence (WOE) neglect the impact of whole conditioning factor on flood occurrence(Muthu and Ramamoorthy 2024). On the other hand, multivariate statistical analysis methods such as logistic regression (LR) assess the influence of conditioning factors on flood occurrence while it neglects the impact of each class on flood (Tehrany and Kumar 2018; Tehrany, Pradhan, and Jebur 2015). Machine learning models, including Artificial Neural Networks (ANN) (Andaryani et al. 2021; Ighile, Shirakawa, and Tanikawa 2022; Priscillia, Schillaci, and Lipani 2021), Support Vector Machine (SVM) (Costache et al. 2020; Duwal, Liu, and Pradhan 2023; Tehrany, Kumar, and Shabani 2019; Tehrany et al. 2015), Decision Trees(Khosravi et al. 2018), K-nearest neighbors(Al-Aizari et al. 2022), Naïve Bayes (Hasanuzzaman et al. 2022), Adaptive Neuro-Fuzzy Inference Systems (Wang et al. 2019), and decision tree–based models like random forest (Hasanuzzaman et al. 2022; Kulithalai Shiyam Sundar and Kundapura 2023, 2023; Razavi-Termeh et al. 2023) CatBoost, LightGBM (Kulithalai Shiyam Sundar and Kundapura 2023; Saber et al. 2022), Extreme Gradient Boosting (Hasanuzzaman et al. 2022; Ma et al. 2021; Mirzaei et al. 2021; Razavi-Termeh et al. 2023) and gradient boosting machines (Felix and Sasipraba 2019; Saravanan et al. 2023) have been introduced to analyze large complex datasets for flood susceptibility investigation efficiently. As an illustration, it is viewed as a black box due to its complexity and the requirement for high-capacity computers. Hence, to resolve the current shortcomings in flood studies, it is crucial to devise more sophisticated and precise methods.

Different factors like altitude, slope, and aspect contribute to generating flood susceptibility maps. Each factor has its distinct influence on the analysis, although some factors may produce similar effects or negligible impact on the final outcomes. Therefore, identifying optimized conditioning factors is crucial to minimize the time and cost of data collection, thereby reducing computation time spent on analyzing less significant factors. This study aims to meet all necessary criteria for efficient flood susceptibility modeling by enhancing existing methods. The expected outcome is that improved susceptibility mapping will enhance previous study results.

1.3 Purpose of the Thesis

Natural disasters are becoming more frequent worldwide, emphasizing their crucial role in ensuring environmental and public safety. Increasing urbanization and climate change are expected to heighten the frequency of rainstorms and river flooding. Flood events in tropical nations, particularly in Lao P.D.R., illustrate the extremes in climate variability. Consequently, monitoring, mapping, modeling, and mitigating floods have become top priorities for governments.. (Pradhan, Tehrany, and Jebur 2016). These occurrences result from unpredictable alterations in natural conditions caused by natural forces. In general, these catastrophes are beyond human ability to predict or control. Large-scale natural disasters including floods, earthquakes, landslides, and subsidence have a significant effect on infrastructure, agriculture, human lives, and the environment. The outcomes of these natural hazards differ based on their magnitude and the specific geographic regions they impact

Floods are the most common natural disasters affecting humans and their environments. They are especially widespread in Asia and the Pacific regions, significantly impacting the social and economic stability of these nations. For example, Lao PDR has faced several devastating floods, notably in 2009, 2011, 2013, 2018, and 2019. These disasters have significantly impacted the country's socio-economic development. In 2018, flood damages amounted to about 2.1% of the nation's GDP, equivalent to around US$ 371 million(Anon 2022; UN, World Bank, GFDRR & EU 2018a) The country experiences damage from natural disasters every year, with floods in the plains and frequent landslides in hilly areas. These incidents endanger people's lives and property and significantly impact the economy and agriculture sectors. Over the past centuries, there has been an increased focus on improving flood management. Recent causes of recurring floods in certain areas are primarily attributed to unplanned urbanization, construction, and deforestation. Without adequate management, these factors can lead to disastrous outcomes such as dam failures, flooding damaged agricultural crops and livestock, thereby heightening flood risks. Despite these challenges, human intervention plays a crucial role in mitigating flood disasters through extensive use of technology. Technological applications can aid in preemptive actions against floods by identifying flood-prone areas and providing early warnings of impending catastrophes.

In the past, fieldwork was the primary method used to map and monitor floods, but it was limited by factors such as time and weather conditions. However, with the introduction of GIS and RS technologies, these constraints have been overcome, leading to continuous improvement in flood studies. These technologies have transformed hazard research, particularly in flood studies, resulting in more effective mitigation of this phenomenon. The adoption of GIS and RS technologies has indeed revolutionized the approach to mitigating flood disasters. Advancements in technology have made it easier to predict and mitigate flood damage, a feat that was previously unattainable. Despite various methods and techniques proposed and tested for mapping flood-prone areas and producing flood inventory maps, many of them have significant limitations that require attention. Conversely, certain methods, such as Rule-based machine learning, have yet to be evaluated in flood studies.

Flood detection analysis should be rapid (Hansana et al. 2023) because floods can subside quickly in an inundated area. Hence, researchers face time constraints when mapping all locations. Traditional methods like fieldwork are impractical for this task due to on-site challenges and lengthy procedures. Additionally, conventional hydrological techniques, such as gauge and discharge measurements, are inadequate for monitoring and mapping flood locations due to the temporal and spatial variations in extensive wetlands (Hamidi et al. 2023; Na and Li 2022). Another method, visual interpretation of satellite images, is time-consuming, prone to inaccuracies, and entails high costs. It is based on expert knowledge; therefore, it can be erroneous (Mucsi and Bui 2023).

The threshold segmentation algorithm or histogram thresholding is a simple but widely used and effective method to generate a binary image (Pulvirenti et al. 2023). Thresholding techniques in SAR sensors depend on distinguishing between flooded and non-flooded areas, making them effective for identifying floodplains. However, these techniques are sensitive to low-contrast images. Moreover, they are limited because they are tailored to specific satellite scenes, often relying on visual interpretation. Additionally, their manual and time-consuming nature further constrains their utility. (Pulvirenti et al. 2023). Active contour modeling can also be employed to map the extent of flooding in a region. However, this technique requires the researcher to possess prior knowledge of the statistical characteristics of images. Additionally, the method is impeded by local minima and becomes less accurate when the initial contour chosen is basic or distant from the object boundary. Synthetic aperture radar (SAR) interferometry as another available method, should produce a coherence map; however, this technique is often difficult to be (Fobert, Singhroy, and Spray 2021). The generation of a coherence map is also complex and disadvantageous; for instance, it requires ground data and two precisely co- registered SAR images (Zhang et al. 2022).

All of the optical images are unsuitable for flood detection applications (Jiang et al. 2021) because clouds usually cover the sky during a flood event, thereby limiting the observational capability of these optical sensors. However, SAR signals can penetrate vegetation and forest (Salem and Hashemi-Beni 2022). These sensors are capable of functioning during both day and night, illuminating various features of a terrain due to their single- or multi-polarized capabilities. Consequently, the objective of this research is to address the limitations of optical data by utilizing active Sentinel 1 Satellite Images. Regarding the susceptibility mapping, in some methods such as LR, the impact of classes of each conditioning factor on flood occurrence is not considered (Elmoulat and Ait Brahim 2018). Other statistical methods such as FR method, consider the relationship between flood occurrence and each conditioning factor separately, while not considering the relationships among all the conditioning factors themselves (Megahed et al. 2023). This thesis aims to introduce enhanced methods for mapping flood locations and identifying areas susceptible to flooding using machine learning approaches.

The primary aim of this research is to utilize the developed maps to prevent urbanization in flood-prone areas and promote environmental sustainability. Identifying at-risk areas is essential for reducing damage and casualties during floods. Developing flood inventory maps forms the basis for mapping flood susceptibility and pinpointing these vulnerable regions. Additionally, optimizing conditioning factors is a key focus. Governments and urban planners can use the study's findings to identify safe zones for residents, support emergency responders during crises, and update strategies for urban planning. This data can decrease the need for field surveys performed by organizations such as surveying departments.

1.4 Research Questions

To achieve the research goals of flood susceptibility mapping, this dissertation addresses the following research inquiries:

- Is it feasible to precisely detect flooded areas using satellite imagery?
- Is it feasible to improve flood susceptibility mapping through the application of

machine learning methods?

- Which conditioning elements are most effective in causing floods in each of the research areas?
- Which flood-conditioning variables are most important for mapping areas that are susceptible to flooding?
- Do the findings from diverse methods employed to map flood susceptibility and the real extent of flooded zones correspond effectively in the model validation phase?
- Which machine learning approach based on data is most efficient for pinpointing flood-prone areas?
- How can we evaluate the accuracy and reliability of temporal and spatial models?

1.5 Research Objectives

The primary objective of this research is to enhance flood mapping and modeling methods to create more dependable flood inventories and susceptibility maps.

- To provide a simple and precise RS method for mapping flood extent using Sentinel-1 SAR data in Google Earth
- To create detailed flood susceptibility maps using advanced machine learning approaches such as 1) Random Forest (RF), 2) Support vector machines (SVM), 3) Artificial Neural Networks (ANN), and 4) Long Short-Term Memory (LSTM).
- To find out the accurate flood susceptibility zones in Nam Ngum River Basin and classify the whole basin area into five classes: very low, low, moderate, high, and very high.
- To identify the most effective conditioning factors in flood susceptibility mapping through the utilization of the Area Under the Curve of Receiver Operating Characteristic (AUROC)

1.6 Research Scope

Floods are the most common and severe of all natural disasters. Each year, flood events result in substantial global losses and disruptions to societies. Flood management plans address all aspects of flood management focusing on prevention, protection, preparedness, including flood forecasts and early warning systems (Tariq, Farooq, and Van De Giesen 2020). During the pre-disaster stage of the flood management, many studies can be done such as flood detection, flood susceptibility, hazard, vulnerability and risk mapping (Kron 2005). This study's primary focus is on enhancing flood susceptibility mapping techniques. As it has been mentioned in problem statement, there are some weak points in existing techniques regarding the flood extent mapping and flood susceptibility mapping (AL-Areeq et al. 2023)

In the scope of flood detection, optical data and most of the available classifications techniques for them are not applicable (Sanyal and Lu 2004). The primary issue lies in the presence of cloud cover and the incapacity of optical sensors to penetrate through them. Traditional gauge and discharge measurements are based on very simple assumptions, and they have linear structure (Durand et al. 2023). However, flood and river structures are very complex and non-linear. Other available methods of visual interpretation and threshold segmentation algorithm are based on expert's knowledge which can be biased (Mucsi and Bui 2023). Change detection method using interferometric technique is very complex and it requires two precisely co-registered SAR images (Mastro et al. 2022). Therefore, this study seeks to address the existing drawbacks and challenges in flood detection by introducing an optimized technique utilizing Sentinel-1 (SAR) data in the Google Earth Engine. These data, along with the information from the Knowledge for Development (K4D) (https://apps.k4d.la/explorer) o online portal (for years 2018, 2019, and 2020) and the historical flooded area (from1985 to 2010) from Colorado Flood Observatory (https://floodobservatory.colorado.edu) were initially analyzed to locate flooded areas.

The research scope for flood susceptibility mapping in the Nam Ngum River Basin (NNRB) involves analyzing hydrological and climatological data, employing GIS and remote sensing technologies to create detailed topographical maps, and identifying flood-prone areas. It includes assessing the impact of land use, infrastructure, and human activities on flood dynamics. The study aims to develop accurate flood susceptibility maps, enhance community resilience through adaptive strategies, and provide policy recommendations for effective flood management. This research integrates scientific analysis, technological tools, and socio-economic considerations to mitigate flood risks and support sustainable development in the NNRB.

1.7 Dissertation Organization

The novelty of the present work is leveraging machine learning models and open-source remote sensing data in data-scarce areas, especially in Lao PDR. In addition, this study uses deep learning and freely available open-source datasets to assess flooding impacts on land cover, population, and critical facilities.

This thesis consists of six chapters, which are summarized as follows

Chapter 1: Introduction

This chapter discusses the background of the study, statement of the problem, purpose of the thesis, research question, research objective, research scope and dissertation organization.

Chapter 2: Literature Review

In the literature review, the chapter discusses the importance of flood disasters and their negative effects in Lao PDR. It covers valuable flood mapping techniques, flood susceptibility methods, conditioning factors, and the validation process for the model.

Chapter 3: Research Methodology

This chapter mainly mention about overall methodology, characteristic of the study area, flood detection method used in this research, flood susceptibility mapping method, data collection and preparation, preparation of the flood inventory data, optimization of conditioning factors.

Chapter 4: Result and discussion

This chapter primarily explores the use of machine learning methods to map flood susceptibility in the Nam Ngum River Basin, Lao PDR. It details the methodology for identifying and optimizing key conditioning factors that influence flood susceptibility. The chapter emphasizes the integration of open-source remote sensing data, Geographic Information Systems (GIS), and advanced machine learning algorithms to enhance the precision of flood-prone area predictions by using multicollinearity test and information gain ratio. The outcomes aim to provide valuable insights for local authorities and planners to implement effective flood mitigation strategies, reduce potential damages, and improve overall disaster preparedness and resilience in the region

Chapter 5: Flood Risk Assessment on Land-cover, Population and Critical Facilities in Nam Ngum River Basin, Lao PDR

This chapter provides an in-depth flood risk assessment of the Nam Ngum River Basin, Lao PDR, using machine learning techniques. It focuses on identifying and analyze the effect on the different land use/land cover classes and populations in in Nam Ngum River Basin (NNRB), Lao PDR, optimizing conditioning factors, and improving prediction accuracy to mitigate flood impacts and enhance regional flood management strategies.

Chapter 6: Conclusion and recommendation

This chapter summarized the conclusion of introduction, summary, limitation and recommendation for the future work.

CHAPTER 2

LITERATURE REVIEW

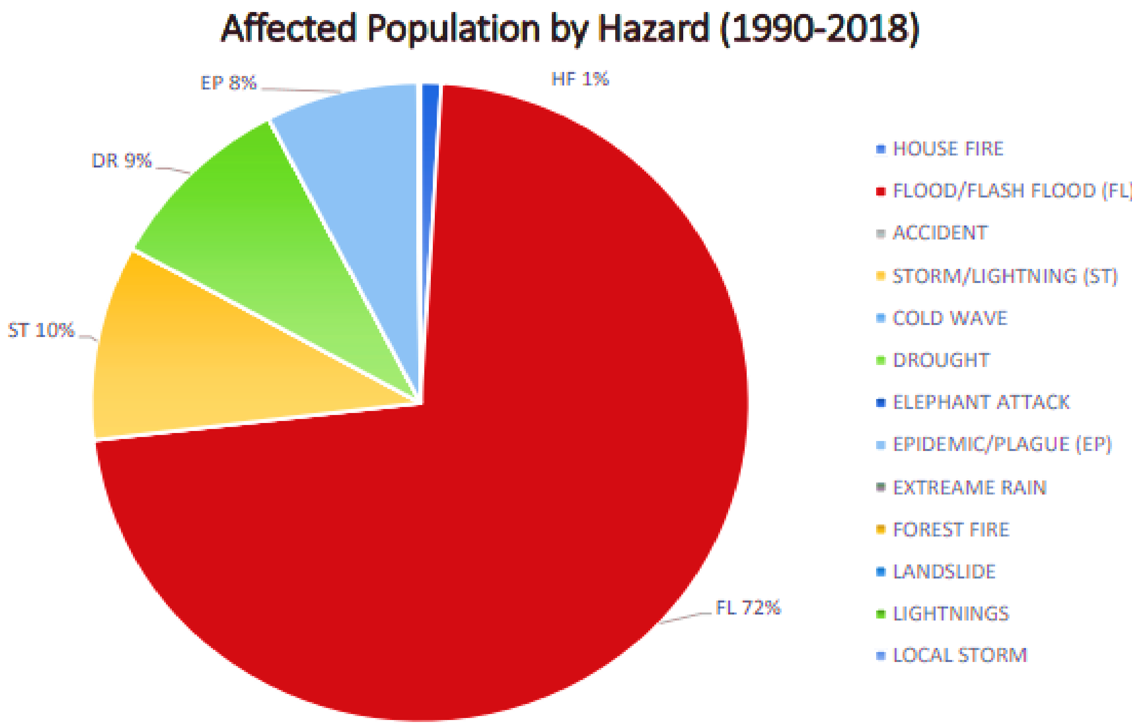
2.1 Overview

This chapter covers the broad impacts of flood disasters and examines both conventional and advanced methods and techniques for flood detection and modeling. It provides an overview of conditioning factors influencing the flood process. Additionally, it explores the uses and challenges of open-source remote sensing (RS) and geographic information systems (GIS) in flood modeling.

2.2 Flood Disaster and Its Impact

A flood is a remarkable high-water flow that overflows the banks of a river, causing water to spread to the floodplain due to severe rains (Cirella and Iyalomhe 2018).Flood is a major devastating natural disaster regarding the number of people affected and economic loss (Di Baldassarre et al. 2010; FitzGerald et al. 2010; Khalil and Khan 2017; Morrison, Westbrook, and Noble 2018; Rappaport 2014). Large and damaging floods are increasingly occurring every year around the world (Kundzewicz et al. 2014), particularly in low-economy countries (Imamura 2022; Li et al. 2012). For example, Lao PDR has faced several devastating floods, notably in 2009, 2011, 2013, 2018, and 2019. These disasters have significantly impacted the country's socio-economic development. In 2018, flood damages amounted to about 2.1% of the nation's GDP, equivalent to around US$ 371 million (UN, World Bank, GFDRR & EU 2018b). The country experiences damage from natural disasters every year, with floods in the plains and frequent landslides in hilly areas. These incidents endanger people's lives and property and significantly impact the economy and agriculture sectors.

Figure 4 During the period from 1990 to 2018, floods affected 72% of those impacted by disasters, while storms, droughts, and epidemics affected 10%, 9%, and 8% of the population, respectively.



Source: MLSW, Government of Lao PDR, 2020, using data from Lao-Di (1990-2018)

Figure 4. Affected Population by Hazard (1990-2018)

Table 1 Economical Damage and Losses by Floods in 2018

Sector	**Damage)billion LAK(**	**Loss)billion LAK(**	**Total)billion LAK(**
Social sector:			
Housing and settlement	21.12	0.57	21.69
Education	18.73	1.68	20.41
Health and Nutrition	8.58	3.32	11.89
Culture	10.11	0.25	10.36
Production sector:			
Agriculture: Crops, Livestock, Forestry, Irrigation	139.80	1,087.60	1,227.30
Industry and Commerce	0.80	2.99	3.78
Tourism	21.87	9.59	31.46
Infrastructure sector:			
Transport	822.02	785.80	1,607.82
Electricity	42.20	3.20	45.40
Water and sanitation	50.96	19.02	69.98
Waterway	116.90		116.90
Total)billion LAK(	**1,253**	**1,914.02**	**3,166.99**
Total) million USD(	**147**	**224.5**	**371.5**

Source: PDNA Report 2018

2.3 Flood Mapping Applications

Flood maps come in various types, each serving a unique purpose in identifying and understanding flood risk. These maps provide crucial information for flood risk management, emergency response, urban planning, and environmental conservation. (Bentivoglio et al. 2022) divided the flood maps into three categories as shown in Figure 5:

1. Flood inundation mapping delineates areas expected to flood, employing modeling techniques and data to aid in emergency response, risk assessment, and urban planning for flood-prone areas. (see Figure 5a).
2. Flood susceptibility mapping predicts areas vulnerable to flooding based on factors like terrain, hydrology, and land use, crucial for risk assessment and disaster preparedness planning. (see Figure 5b)
3. Flood hazard map measures the depth and coverage of water across an inundated area, aiding in assessing flood risks and planning mitigation strategies (see Figure 5c).

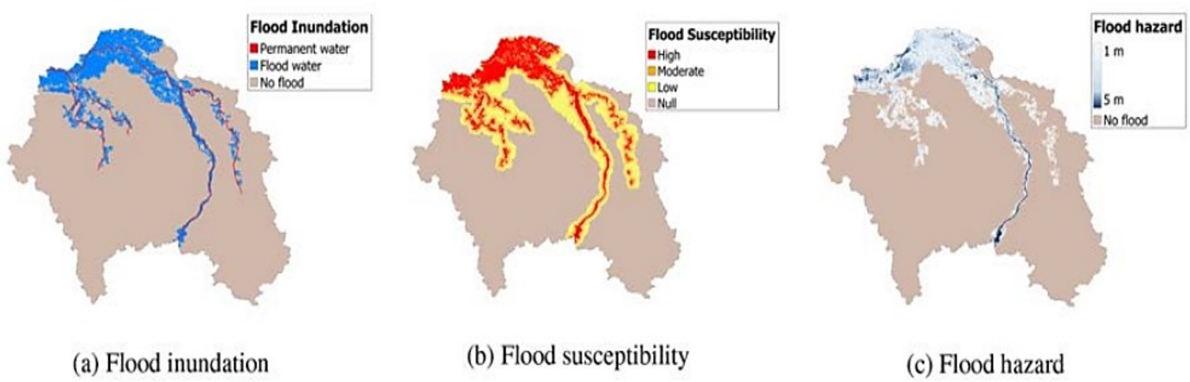


Figure 5. Various types of flood maps. (a) flood inundation map, (b) flood susceptibility map, and (c) flood hazard map.

These maps are widely used, each with its own limitations. For example, remote sensing data are not always able to capture the flood peak and can require manual refinement (Notti et al. 2018). Historical flood inventories are not always available to map flood susceptibility (Zhao et al. 2020). Finally, numerical hydrodynamic simulations are computationally expensive and time-consuming (Löwe et al. 2021).

2.4 Flood Inventory Map

Flooding is a natural phenomenon and flood inventory is defined as a map which illustrates the location of the flooded areas related to the specific time period from particular region (Yu et al. 2023). Flood inventory maps do not prevent floods, but they play a crucial role in identifying vulnerable areas to mitigate damage to property and save lives during flood events. Accurate and current flood inventory maps are essential for conducting additional analyses such as susceptibility, hazard, and risk mapping. Without mapping, potential flood-prone areas may go unnoticed, limiting planners' ability to guide development safely and minimize future flood damage. Flood inventory maps are effective and easily comprehensible tools for specialists like hydrologists, as well as for non-specialists including decision-makers, planners, and civil leaders

Flood inventories are crucial for comprehending river and landscape changes and for developing flood susceptibility and related maps. To create susceptibility maps using statistical methods, a reliable inventory is necessary, along with conditioning factor maps used as inputs (Shafapourtehrany et al. 2023). Flood inventory maps can be created according to research objectives and available resources. They are typically generated through analyzing aerial photographs, conducting field surveys, and reviewing reports and historical records. Sometimes, a combination of these methods is employed. According to (Reichenbach et al. 2018) stated that no standards are available for inventory mapping because of the lack of operational protocols in preparing and updating maps, which in turn reduces the credibility of inventory maps. The reliability of the analysis outcomes derived from flood inventory maps depends significantly on the quality and comprehensiveness of those maps. Validating and assessing the quality and precision of the flood inventories is, therefore, essential (Hitouri et al. 2024)

One of the most important difficulties related to the flood monitoring is a flood extent extraction, since it is almost impossible to recognize the flood inundation area via field survey. The flood inventory map plays a crucial role in calibrating and evaluating hydraulic models to reconstruct events during floods and determine the factors that influenced the water's path. Moreover, the flood inventory map can be used for damage assessment and risk management, and can assist to saviors during flooding (Masafu and Williams 2024). Monitoring flood disasters is crucial for assessing losses, issuing early warnings, conducting analyses, and reconstructing inundation areas affected by floods. Having knowledge about the affected areas may serve as a beneficial preliminary point for environmental planners and managements (Duwal et al. 2023). In every susceptibility analysis, inventory maps need to be divided into two categories for training and testing purposes. This step greatly influences the outcome of the final susceptibility map. No standard ratio or selection method exists for training and testing inventory (Zhu 2024).

Two primary considerations when evaluating the quantity and characteristics of data for testing and training are temporal and spatial robustness. Temporal robustness involves dividing the inventory map into two periods: the initial period used for training data and a subsequent period for validation data. Some researchers have investigated how different inventory maps impact the resulting outcomes. For instance, (Zhu 2024) compared susceptibility maps using decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. According to the literature, the inventory map is typically divided into 70% for training purposes and 30% for testing. The training subsets will be used for the construction of the model, and the validation subset was used for validation of the predictive power of the resulting models (Bhattarai et al. 2024).

2.5 Flood Detection

Flood detection is a crucial aspect of disaster management and environmental monitoring, aiming to mitigate the devastating impacts of flooding on communities, infrastructure, and ecosystems. As climate change and urbanization increase the frequency and severity of flood events, the need for advanced flood detection systems becomes ever more pressing. These systems utilize a combination of hydrological models, remote sensing technologies, and real-time data analytics to predict, monitor, and provide early warnings of potential flood occurrences. By leveraging innovations in satellite imagery, ground-based sensors, and machine learning algorithms, flood detection mechanisms can offer timely and accurate information, enabling authorities to implement proactive measures, evacuate at-risk populations, and minimize economic losses. Effective flood detection not only enhances the resilience of societies to natural disasters but also supports sustainable development by informing better planning and resource management strategies.

2.5.1 Traditional Methods

Traditional methods for flood detection have long been a cornerstone in safeguarding communities from the devastating impacts of flooding. These methods primarily rely on historical data, manual observations, and basic hydrological tools to predict and monitor flood events. Key techniques include the use of rain gauges to measure precipitation levels, river gauges to track water levels in rivers and streams, and empirical models based on past flood events to forecast potential floods. These conventional approaches, while fundamental, often involve significant manual effort and can be limited by the spatial and temporal resolution of the data collected. Despite these limitations, traditional flood detection methods have provided valuable insights and a foundation upon which modern technologies have been built. By combining local knowledge and historical records, these methods continue to play a vital role, especially in regions with limited access to advanced technological infrastructure. Understanding and enhancing these traditional techniques remains crucial for improving flood resilience and developing more comprehensive flood management strategies.

Field mapping is one of the traditional methods mostly used by researches in the past (Mollaei et al. 2018). The weak point of this method is related to the extent of the flood that is usually too large to be seen in a larger area. Furthermore, in some cases, flooded areas are under the vegetation and cannot be recognized easily. Traditional hydrological methods such as gauge and discharge measurements cannot be used to monitor and map the flood locations due to the temporal and spatial heterogeneity of large wetlands (Na and Li 2022).

2.5.2 Remote Sensing (RS) and Geographic Information Systems (GIS) Techniques

Remote Sensing (RS) and Geographic Information Systems (GIS) are transformative technologies that have reshaped the way we observe, analyze, and manage our natural and built environments. Remote sensing entails gathering data about the Earth's surface using satellites or aerial imagery, capturing data on various physical and environmental parameters without direct contact. This data provides critical insights into land use, vegetation cover, water bodies, and atmospheric conditions, among other factors

Generation of various numbers of satellites and sensors made revolution in monitoring, evaluating and predicting the natural disasters (Zhou et al. 2019). Over the last two decades, remotely sensed data have been used effectively for monitoring and analyzing hazards, and high-resolution imagery has revolutionized RS research (Baghermanesh, Jabari, and McGrath 2022). Considerable improvements in flood detection have been made due to increased data collection rates, higher sensor resolution, the development of change detection algorithms, and the incorporation of RS techniques (Albertini et al. 2022). Typically, studies of hazards require multi-temporal datasets in order to identify spatial changes and the process of hazards occurrence

(Tsyganskaya, Martinis, and Marzahn 2019). Remote sensing is valuable for such studies because it captures extensive areas of the Earth's surface regularly and at a comparatively low expense. When it is supplemented with non-RS data, it is easy and effective to assess the evolution of natural catastrophes(Lei et al. 2022). (Anusha and Bharathi 2020) have used GIS in their research in order to provide flood inventory map of Rapti and Ghaghara Rivers in India. They reported that using GIS tools and multi-temporal synthetic aperture radar and optical data provides additional benefits by enabling the determination of basin characteristics and facilitating the adjustment of river component scenarios to fit various river sizes. It also allows users to gain a more comprehensive understanding of watershed conditions during and after a flood. Two main types of data sources for flood analysis are optical and active sensors. More information and description about these data sources and their applications are explained below

2.5.2.1 **Optical Sensors**

Visual interpretation of aerial photos was one of the most famous methods to detect flood locations in many years ago. (Oberstadler, Hönsch, and Huth 1997) investigated the efficiency of the visual interpretation of aerial photos method and a European Remote Sensing (ERS-1) satellite data analysis with automatic classification techniques to derive the flood boundary. Their findings demonstrated that visual interpretation yielded more precise outcomes in comparison to satellite analysis. This discrepancy was attributed to the limited resolution of satellite data and the technological and computational limitations prevalent during that period. Recently, the accessibility to the wide range of software, very high-resolution satellite imageries, active and passive sensors, facilitated the data collection, flood analysis and mapping within few hours (Auynirundronkool et al. 2012). Nowadays, visual interpretation is seen by researchers as a method that consumes a lot of time, lacks accuracy, and is expensive. It relies on expert knowledge and consequently may contain errors. (Bovenga et al. 2018). Moreover, it requires field surveys simultaneously.

It is evident that optical imageries are not suitable for flood detection applications (Shastry et al. 2023). The reason is usually during the flood, the sky is covered by so many clouds which limit the observation for the optical sensors. These sensors are not capable of penetrating the cloud cover and they are highly affected by weather conditions. Using active sensors can overcome the difficulties and drawbacks of optical sensors which will be explained in the next sub-section.

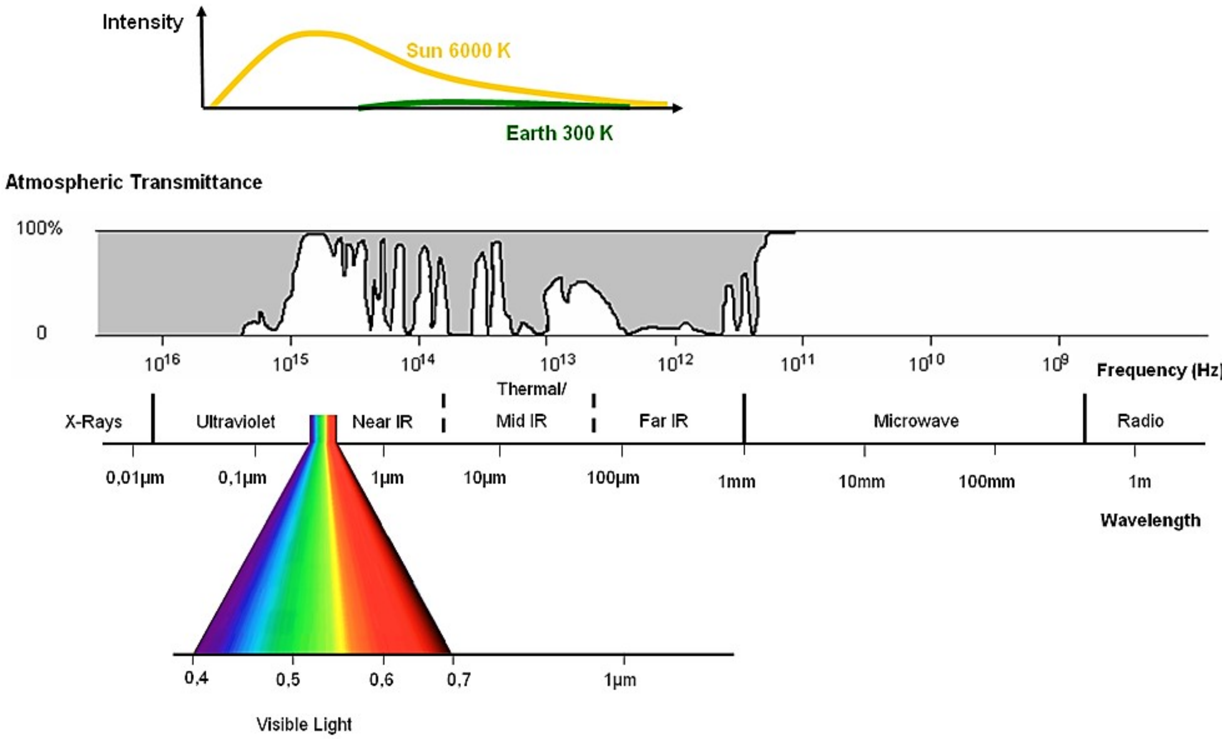


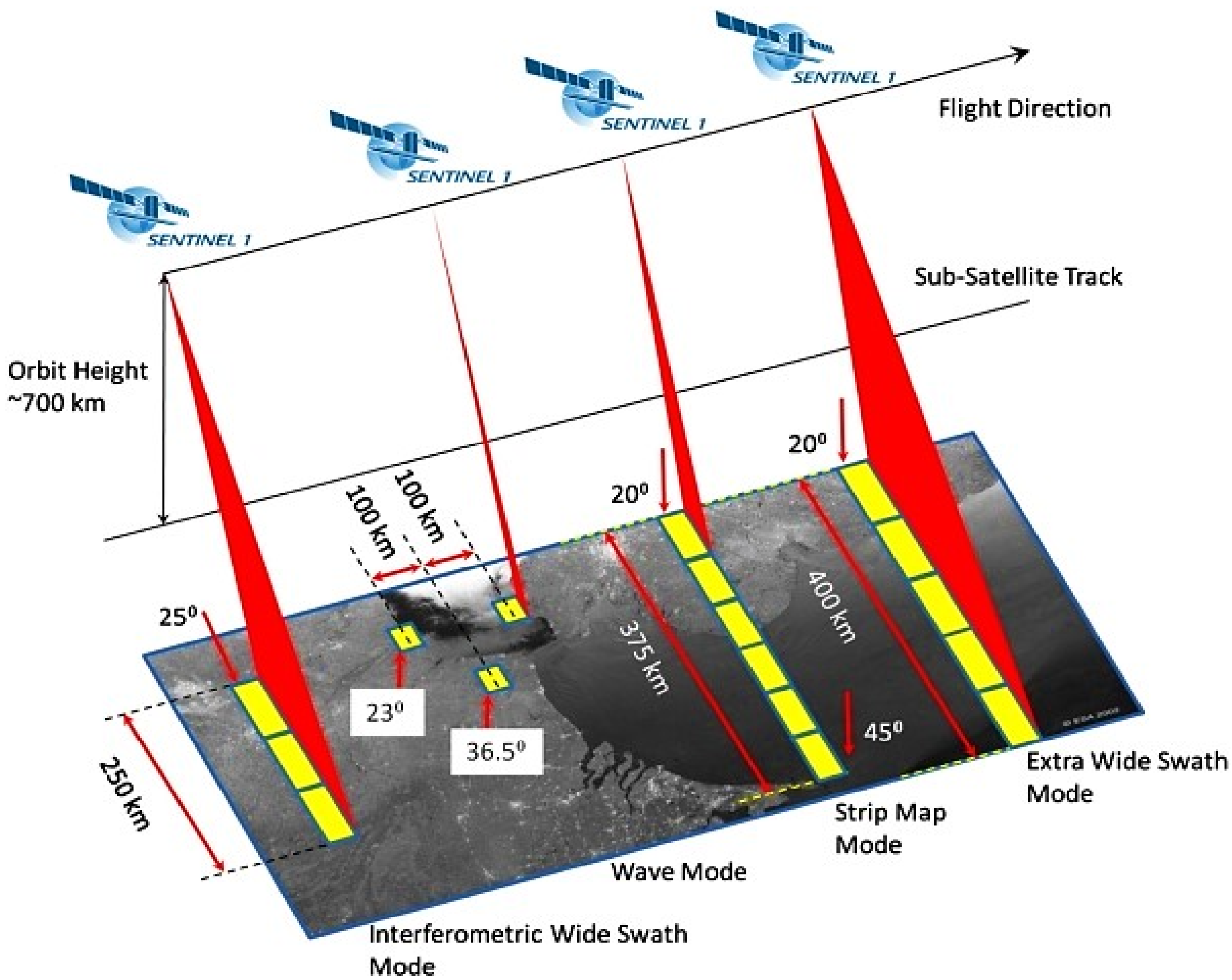
Figure 6. The electromagnetic spectrum and atmospheric transmittance.

2.5.2.2 **Active Sensors**

SAR is an active system that produces high-resolution imaging at radar frequencies, even from spaceborne platforms. The SAR system operates by emitting radar pulses to either side of the platform in sequences. It determines the cross-track position of objects based on the return time of the pulses, and their along-track position by recording amplitude and phase information. This information can be synthesized to simulate returns as if they came from a longer antenna. Finally, the amplitude and phase data obtained are processed to generate a digital image. That process might also contain some local averaging to decrease the image noise (Horritt 1999). Active sensors are not affected by sun illumination and atmospheric conditions (Horritt, Mason, and Luckman 2001). Moreover, they can penetrate the cloud cover and vegetation (Karjalainen et al. 2012). They have the capability to function day and night, and with either single- or multi-polarized modes, they can emphasize different features of the same landscape. Furthermore, the flooded area under the vegetation can be detected using SAR imageries (Horritt 2003)

Water bodies produce no return to the antenna in the microwave spectrum and will be shown as black in SAR imagery. Due to the specular backscattering characteristics of SAR pulses on plain water bodies and the resulting low signal return, the use of SAR data for high-resolution flood mapping is comparatively straightforward (Matgen et al. 2011). Therefore, it can be concluded that SAR imagery offers a huge potential for flood studies (Kim and Lee 2020). Not only in flood studies, SAR systems are rising in importance for environmental monitoring projects such as LULC mapping, oil-spill discovery, landslide and forestry (Singh and Vyas 2022).

For example, the resolution of Sentinel-1 ranges between 5 to 20 meters (in IW mode), and it revisits areas every 6 to 12 days (more frequently at lower latitudes). Signal processing enhances the image's clarity. Smooth surfaces reflect incident waves away from the satellite, while rough surfaces scatter waves back to it.



(image credit: Copernicus, ESA).

Figure 7. Sentinel-1 imaging modes

Sentinel-1 operates using two satellites that orbit the Earth in a polar trajectory, and each satellite is equipped with four different modes for acquiring data. I haven't completely comprehended the distinctions between these modes, and I believe such understanding may not be necessary. Through trial and error, I identified the optimal combination of modes that suits my requirements.

- Mode: Interferometric Wide Swath ("IW": analyzing the interference of two crossing waves)
- Polarization: VV (vertical transmission polarization, vertical reception)

2.6 Flood Conditioning Factor

Floods have various and often connected causes which are called conditioning factors (Al-Kindi and Alabri 2024). The relation and impact of each conditioning factor with flooding should be assessed in order to perform susceptibility analysis (Duwal et al. 2023). The conditioning factor needs to be quantifiable, and there should be a relationship between this factor and the dependent data (flood). Also, the conditioning factor should be collected from the whole study area while it should not represent uniform spatial information (Al-Juaidi, Nassar, and Al-Juaidi 2018). Moreover, its impact should not result in two different outcomes by the end of the process. These factors can be in nominal, ordinal, interval, or ratio scale format (Aldiansyah and Wardani 2023). In natural hazards research, huge databases are often needed. The key elements for producing the best flood maps include using accurate and comprehensive datasets, as well as employing robust modeling techniques.

Additionally, anthropogenic factors like urban areas, road networks, or land use should be considered when assessing flood susceptibility. These factors are related to flood events (Nguyen, Fukuda, and Nguyen 2024). New insights into hydrological research involve determining and mitigating flooding using GIS, digital soil-type maps, topography, and land use/cover data (Mangkhaseum et al. 2024). Different factors may impact flood occurrences in specific regions but might not affect other areas similarly. There is no consensus on which factors are essential for flood susceptibility assessments. However, many researchers commonly use certain factors, emphasizing their significance and critical role in flood research. Below are some of the primary conditioning factors widely utilized in flood studies.

Low-lying areas, subject to rapid drainage from high to low elevations, are particularly vulnerable to flooding (Choubin et al. 2019). Similarly, the gradient of the slope significantly impacts flood risk (Khosravi et al. 2016), with lower gradients posing challenges for effective drainage after flooding (Khosravi et al. 2018). Aspect, representing the orientations of the slope (Shafizadeh-Moghadam et al. 2018) affects soil moisture and weather conditions, influencing flood susceptibility (Rahmati, Pourghasemi, and Zeinivand 2016). SPI quantifies flow erosion power and runoff density (Florinsky 2017). Curvature, indicating surface shape, identifies regions prone to flooding(Khosravi et al. 2019; Tehrany et al. 2015) ,with negative values signifying convexity, positive values indicating concavity, and zero indicating flatness (Youssef, Pradhan, and Sefry 2016). Elevated TWI values highlight areas prone to inundation (Chen and Yu 2011; Sørensen, Zinko, and Seibert 2006). NDVI values represent vegetation vitality, with higher values indicative of denser vegetation cover (Askar et al. 2022), which is inversely related to flood susceptibility (Kumar and Acharya 2016). Rainfall is a primary driver of flooding (Kumar and Acharya 2016). The rainfall depth, intensity, and frequency majorly determine flooding. Annual average rainfall data was mapped using ERA5 data from 2010 to 2020. DTR influences flood probabilities, with closer proximity increasing the likelihood (Shahabi et al. 2020). Soil type affects water absorption and accumulation during floods (Rahmati et al. 2016), while LULC delineates areas at risk (Khosravi et al. 2019), vegetated areas increase travel time and slow the runoff, and bare lands and built-up areas facilitate the flow.

2.7 Flood Modeling

2.7.1 Traditional Hydrological and Hydrodynamic Methods

Over the years, many methods have been developed by the hydrologists for flood modeling (Bahremand et al. 2007; Jayakrishnan et al. 2005). The main aim for most of these flood modeling is to have clear and precise estimation of discharge conditions for the watersheds (Ingle Smith 1999). Many hydrological approaches have been used in the literature (De León Pérez et al. 2024; Fenicia et al. 2014). Traditional hydrological methods such as physical based rainfall-runoff modeling techniques and data-driven techniques are not capable for comprehensive analysis of rivers and inundation areas (Ingle Smith 1999). The reason being the hydrological methods follow one-dimensional procedure while the morphology of the river is not stable, and it has dynamic characteristic due to the high erosive potential (Refsgaard 1997). Moreover, these methods require fieldwork and huge budget for data collection (De León Pérez et al. 2024; Fenicia et al. 2014). Additionally, calibration and information about the internal procedure of basin are two other difficulties for some hydrological studies

In conceptual based models, the process of the flood spreading is typically defined by the Saint Venant equations which analytical solution cannot be achieved using these equations(Chau and Lee 1991). Although mechanism of hydrological process can be described using the conceptual based models, having specific data such as precipitation, river system, runoff and topographic characteristic is necessary for the purpose of calibration (Chau, Wu, and Li 2005). This is considered as a serious drawback of such approaches owing mostly to time consuming, costly and difficult to obtain. On the other hand, sophisticated physical models are also not suitable for real-time prediction as huge dataset is needed and the process requires significant time for model building (Beven et al. 1984). In recent years, numerous techniques have been developed and utilized for mapping areas susceptible to floods. These methods are in most cases combined with GIS and RS data (Chormanski et al. 2008). Examples of such models include; WetSpa (Azizi, Mohajerani, and Akhavan 2018), HYDROTEL (Fortin et al. 2001; Ibarra-Zavaleta et al. 2017) and SWAT (Jayakrishnan et al. 2005; Jimeno-Sáez et al. 2022), LISFLOOD (De Roo, Wesseling, and Van Deursen 2000; Van Der Knijff, Younis, and De Roo 2010), TOPMODEL (Januário, Filho, and Salviano 2022), and HEC-HMS (Bruno et al. 2022)

2.7.2 RS and GIS Based Methods

(Townsend and Walsh 1998) is one of the pioneer studies that proved the possibility of flood prediction through RS modeling in GIS environment. In the literature, many methods have been reported using these techniques such as (Al-Juaidi et al. 2018; García-Pintado et al. 2013; Swain, Singha, and Nayak 2020; Youssef, Pradhan, and Hassan 2011) and many more. Although some of them were able to produce acceptable results, still they contain some weak points that need to be improved (Matgen et al. 2007). For example, flood susceptibility can be analyzed by qualitative and quantitative techniques such as artificial neural networks (Ighile et al. 2022), FR (Megahed et al. 2023) , AHP (Selvam and Antony Jebamalai 2023), LR (Al-Juaidi et al. 2018), adaptive neuro-fuzzy interface system (ANFIS) (Razavi Termeh et al. 2018) and etc. The main concept of all those researches imply the analysis and transforming the input factors into a single output model using the different weighting, computing and interpolating techniques i.e. knowledge-based techniques, qualitative techniques and data mining techniques (Meyer, Scheuer, and Haase 2009). All those efforts came to a same result; that the flood phenomenon has non-linear structure, due to the complexity of the geological environment and existence of conditioning and triggering factors (Gupta et al. 2010).

2.7.3 Machine Learning and Deep learning Methods

The application of machine learning methods in flood modeling is proven by many researches (Bhattarai et al. 2024; Duwal et al. 2023; Mangkhaseum et al. 2024). Machine learning is the main source of techniques for the data-driven modeling (Chen et al. 2020). This field of computer science focuses on algorithms that learn from existing data to perform tasks like classification, prediction, or clustering. However, there is a lack of research on the effectiveness of methods such as Convolutional Neural Network (CNN), Deep Neural Network (DNN), Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Networks (ANNs), and Long Short-Term Memory (LSTM) specifically in flood studies.

Machine learning, a subset of artificial intelligence (AI), employs mathematical algorithms to detect data patterns efficiently, without the need for complex coding. For instance, machine learning is gaining popularity in analyzing non-linear systems and predicting floods. In contrast, traditional flood forecasting systems often involve multiple hydrologic and hydraulic models that simulate physical processes. While these models can help with system knowledge, they typically have high computational and data requirements, swift training and validation, less difficulty, and higher performance than physical models (Ighile et al. 2022; Kim et al. 2016; Wagenaar et al. 2020) In addition, machine learning models can be supervised, unsupervised, or reinforced.

a) Supervised learning algorithms teach themselves to perform functions that can generate predictions (Uddin et al. 2019). A typical example of a supervised learning algorithm is the linear regression model (Panigrahi, Kathala, and Sujatha 2023)
b) Unsupervised learning algorithms are a category of machine learning techniques where the model is trained on unlabeled data (Usama et al. 2019). These algorithms aim to find hidden patterns or structures in the input data without specific guidance or labeled outcomes.
c) Reinforcement learning, a field within machine learning, explores how intelligent agents can optimize their actions in different environments to maximize cumulative rewards. Unlike supervised learning, reinforcement learning operates without requiring labeled input-output pairs or explicit corrections for suboptimal behaviors. Instead, the focus is on striking a balance between exploration (new terrain) and utilizing existing information (Kaelbling, Littman, and Moore 1996). Typical applications of reinforcement learning include autonomous driving, robotics, game playing,

recommendation systems and finance and trading

2.7.3.1 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a computational model inspired by the human brain. It consists of interconnected nodes (neurons) organized in layers: input, hidden, and output. Neurons receive input, compute weighted sums, and pass outputs through activation functions to the next layer. ANNs learn by adjusting connection weights during training, using algorithms like backpropagation to minimize errors between predicted and actual outputs as shows in Figure 8, The idea of artificial neural networks. In flood forecasting, ANN has attracted the attention of researchers enduringly (Chai, Wong, and Goh 2016) as it has the great competence in nonlinear modelling and complex framework without clear physical justification. Hydrologist examines ANN in different flood scenarios such as flood inundation forecasting model (Chang et al. 2018), rainfall-runoff analysis (Vidyarthi, Jain, and Chourasiya 2020), stream flow forecasting (Kasiviswanathan et al. 2016). Recent researches have introduced hybridization approach within ML model or other models like physical based model in order to maximize accuracy rate (El-Telbany 2017; Šaur 2017). More details of ANN such as DL, BPNN and FFNN are further explained below.

1. The Back Propagation Neural Network (BPNN) is an algorithm utilized within ANN. During the processing of a feedforward neural network, if errors are detected, the backpropagation algorithm intervenes to adjust the weights iteratively until satisfactory results are achieved
2. The Feed Forward Neural Network (FFNN) is a neural network algorithm where information flows in a forward direction, from the input layer through hidden layers to the output layer, without adjusting weights during this process
3. Deep Learning (DL) is a relatively recent area of research in ML and Artificial Intelligence. Additionally, it is one of the most prevalent scientific study topics nowadays (Minar and Naher 2018; X. Wang, Zhao, and Pourpanah 2020). Applications such as cancer detection, precision medicine, self-driving cars, predictive forecasting, and voice recognition are just a few of the areas where it has already made a significant impact (Shrestha and Mahmood 2019). On top of that, Deep learning is considered the most effective approach for addressing challenges in image recognition, speech

recognition, forecasting, and natural language processing. It is also the optimal solution in time series data analysis to solve flood prediction problem. (Gude, Corns, and Long 2020) reveal in their study that flood prediction using DL model was more accurate than the physical and statistical models. Further description of different types of DL algorithms that can be applied for different problems and data types.

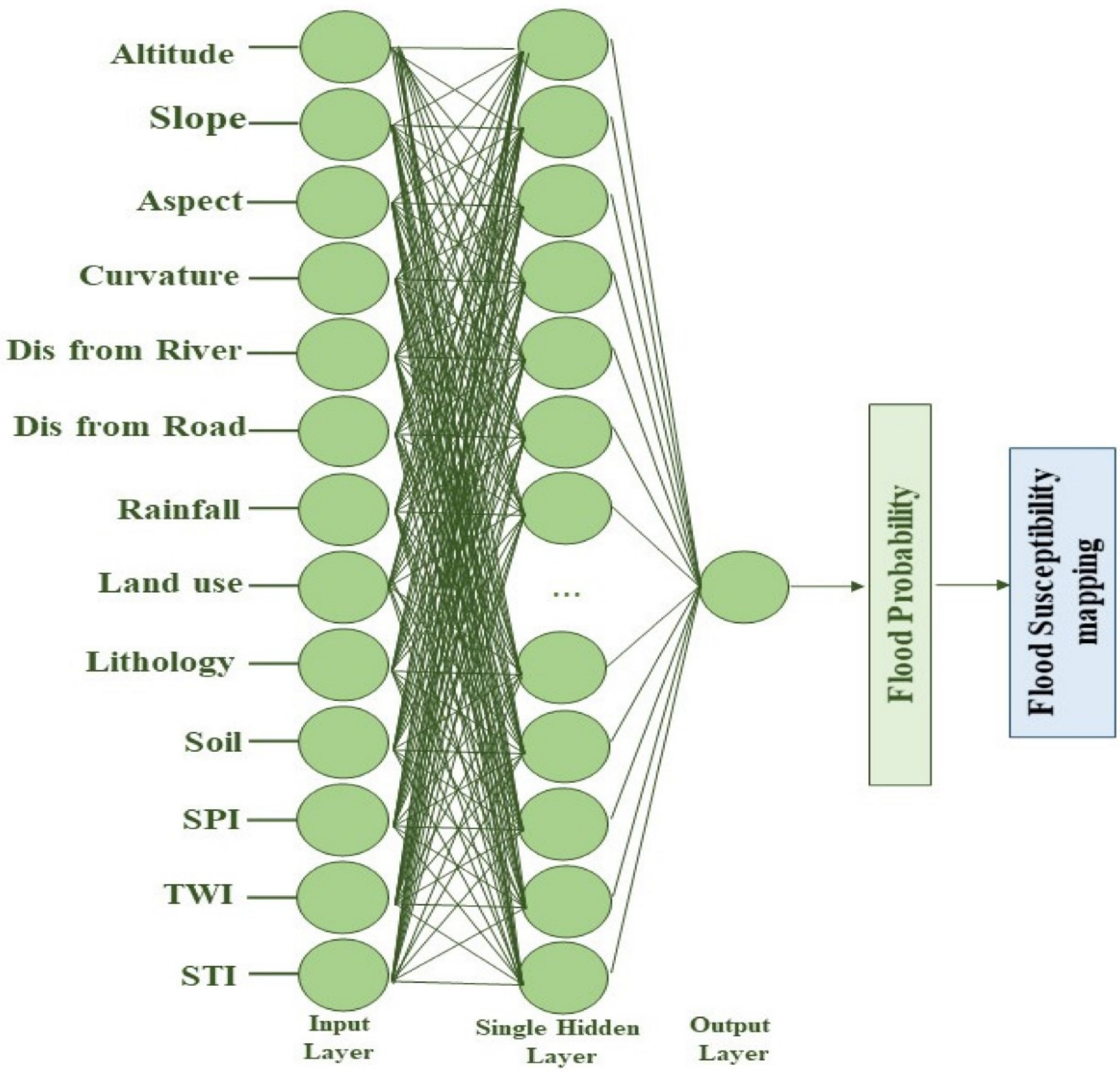


Figure 8. Concept of neural network, three-layer architecture

2.7.3.2 Random Forest (RF)

The Random Forest model was introduced by (Breiman 2001). It is a combined learning method that uses a large number of decision trees as weak classifiers to form a powerful classifier (Figure 9) (Nguyen 2022). It combines the methods of bagging and random subspace. The "bagging" method involves randomly selecting observations from the training data to create subsets, while the basic idea of the "random subspace" method is to randomly select derived independent variables (Li and Hong 2023). Each decision tree thus takes a random subset of observations with a subset of factors and produces a predicted outcome. The final prediction of the Random Forest model is the combination of all the outcomes, which is more accurate due to the applied randomness (Chen et al. 2020). This approach also reduces the common problem of overfitting the model to the training patterns, as decision trees that grow large enough tend to learn unusual patterns and overfit the results to the specific data. For classification problems, the final result is generated according to Breiman's original publication by the majority vote of the decision trees. In this paper, the Python scikit-learn library was used, where the final prediction is made probabilistically, considering the uncertainty of each tree (Kaiser, Günnemann, and Disse 2022)

The two main parameters of the model optimized are the number of decision trees, defined as n_estimators, and the number of independent variables in each tree, or max_features (Y. Wang et al. 2020). The max_depth parameter, as defined in the previous model, and ccp_alpha, used to perform "pruning" of the tree to avoid overfitting, were also chosen. Additionally, min_samples_leaf and min_samples_split was selected to further protect the model from overfitting.

The RF model is not sensitive to multivariate where one variable can be linearly predicted by others and can handle data unevenly and incompletely. Therefore, it is a widely used method for predicting areas of high flood susceptibility and can be used for both classification and regression problems (Abedi et al. 2022; Farhadi and Najafzadeh 2021; Tang et al. 2020)

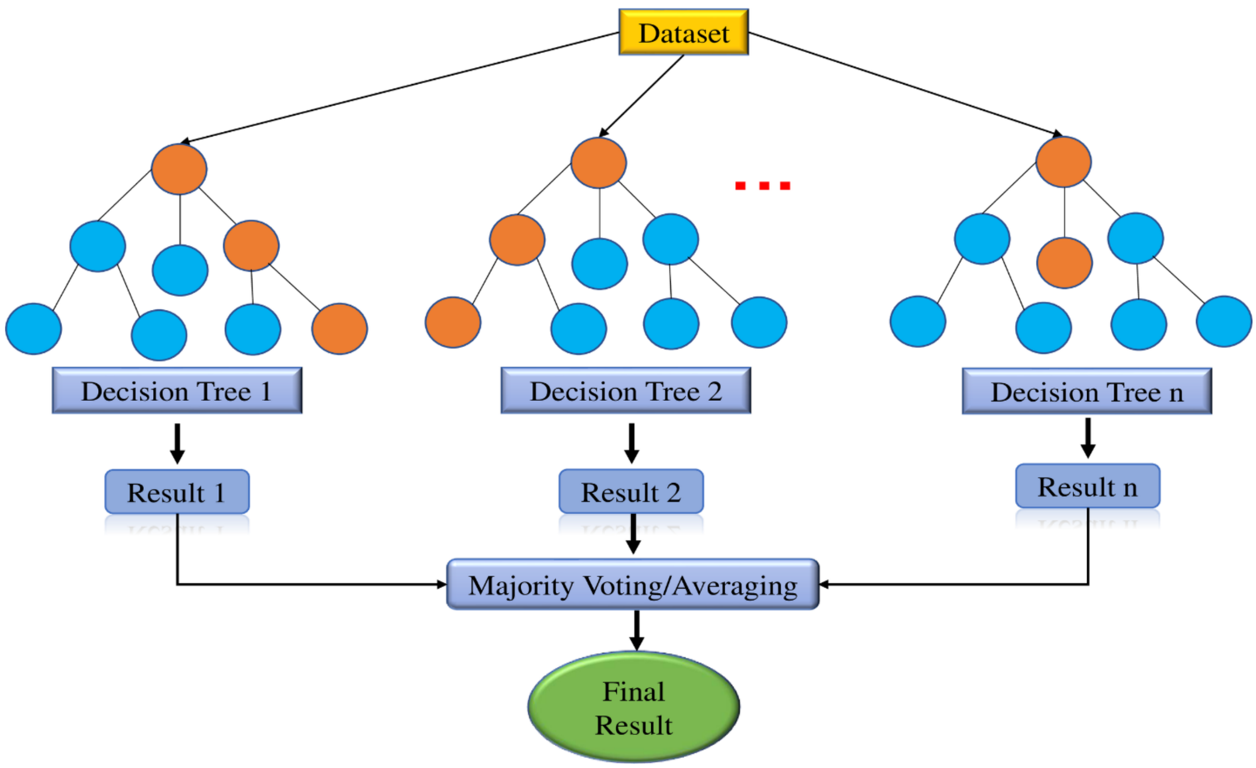


Figure 9. Training process of the Random Forest algorithm

2.7.3.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a binary classifier and machine learning algorithm and it is based on the structural risk minimization principle (Duwal et al. 2023; Yao, Tham, and Dai 2008). Creating a separating hyperplane from the training dataset forms the foundation of this approach. Separating hyper-plane is generated in the original space of n coordinates (xi parameters in vector x) between the points of two distinct classes (Marjanović et al. 2011; Tehrany et al. 2015). SVM finds a maximum margin of separation between the classes and builds a classification hyper-plane in the central of the maximum margin (Pradhan 2013). Points above the hyperplane are classified as +1, while those below are classified as -1. Support vectors are the training points that are closest to the optimal hyper-plane

The classification of new data can be done after the acquisition of the decision surface (Bui et al., 2012a). In the context of flood susceptibility, we have a training dataset consisting of pairs $(x_i - y_i)$ with $x_i \in R^n, y_i\{+1, -1\}$, and $i - 1, \ldots., m$. In the present circumstance of flood susceptibility, x is a vector of input space that contains altitude, curvature, river, SPI, rainfall, LULC, soil, TWI and slope. The two classes $\{+1, -1\}$ indicate flooded pixels and non-flooded pixels. Recognizing the optimal separating hyper-plane is the goal of SVM, which can separate the two classes into flood and non-flood $\{+1, -1\}$ from the training dataset (Tehrany et al. 2014). For the case of linear separable data, a separating hyper- plane can be defined as:

$$y_i(w.x_i + b) \geq 1\xi_i \qquad (2.1)$$

where w is a coefficient vector that defines the orientation of the hyper-plane in the feature space, b is the offset of the hyper-plane from the origin, and ξ_i is the positive slack variables (Cortes and Vapnik 1995). The following optimization problem using Lagrangian multipliers will be solved through the determination of an optimal hyper- plane (Samui 2008).

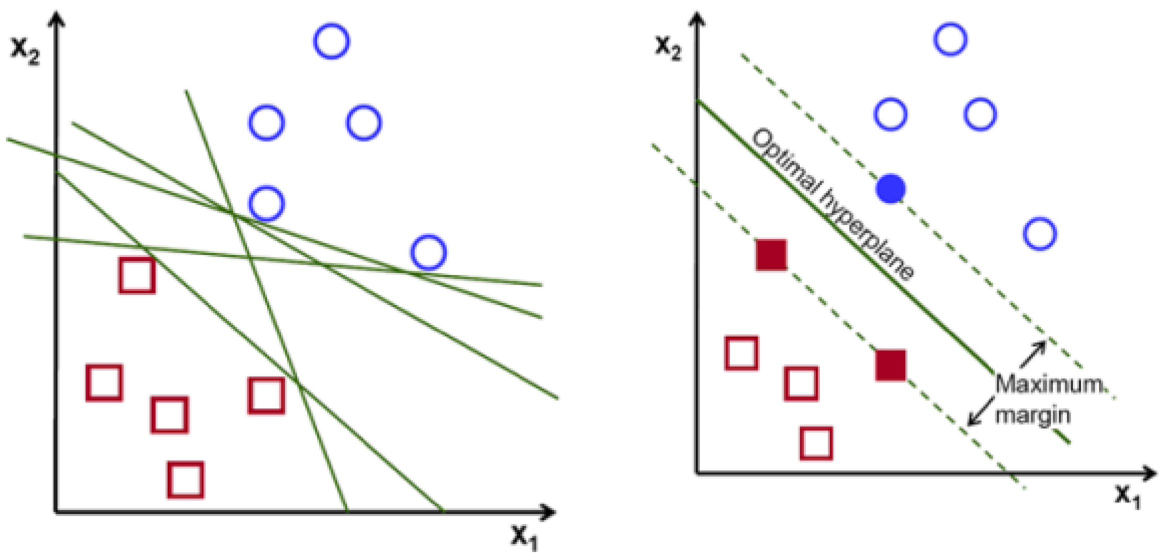


Figure 10. Hyperplanes and Margins

2.7.3.4 Long Short-term Memory (LSTM)

LSTM is a type of RNN that was designed to handle long-term serial data requirements (Gude et al. 2020). The LSTM network is comparable to the RNN in architecture, but it is capable of retaining long-term information. Although (Hochreiter Sepp and Schmidhuber Jürgen n.d.) developed the LSTM network in 1997, it was not employed for rainfall-runoff (RR) modelling until 2016. (Van Houdt, Mosquera, and Nápoles 2020) reported that since 2016 a few studies have employed LSTM for RR modelling and found positive results. Many of the experiments demonstrated that the LSTM network outperformed other RNN networks in capturing the dynamics of time-series for hydrologic applications (Bakhshi Ostadkalayeh et al. 2023; Gude et al. 2020; Van Houdt et al. 2020). Also, LSTM has become popular for applying flood susceptibility; The ability of LSTM models to capture temporal dependencies and process sequential data makes them superior to traditional methods (Fang et al. 2021; Zou et al. 2023)

Figure 11 illustrates an LSTM network cell that includes an input layer, hidden units, as well as an output unit. The input layer is defined as $X = [x_{t-1}, x_{t,} x_{t+1}, \dots x_n]$, and the output vector defined as $Y = [y_{t-1}, y_{t,} y_{t+1}, \dots y]$. Further unfold the hidden layer the gates can figure out which information in a series should be kept and which should be discarded. It can then send vital information through the lengthy chain of sequencing to create predictions (Gude et al., 2020; Li et al., 2021). The LSTM unit consists of a cell state ct, an input gate it, a forget gate ft, a cell gate gt, and an output gate o_t. For each time step t with the precipitation input vector x_t, previous hidden cell state h_{t-1}, and previous cell state c_{t-1},, the updated hidden state h_t is computed by (Xie, Randall, and Chau 2022) the following calculations using Equations (2.2) - (2.7).

$$i_t = \sigma(w_{ii}x_i + b_{ii} + w_{hi}h_{t-1} + b_{hi}) \qquad (2.2)$$

$$f_t = \sigma\left(w_{if}x_i + b_{if} + w_{hf}h_{t-1} + b_{hf}\right) \qquad (2.3)$$

$$g_t = \tan h(w_{ig}x_i + b_{ig} + w_{hg}h_{t-1} + b_{hg}) \qquad (2.4)$$

$$O_t = \sigma(w_{io}x_i + b_{io} + w_{ho}h_{t-1} + b_{ho}) \qquad (2.5)$$

$$C_t = f_t * C_{t-1+i_t} * g_t \qquad (2.6)$$

$$h_t = o_t * tanh_{(C_t)} \qquad (2.7)$$

Where σ is indicates sigmoid, W is weight matrices, b denotes bias. Since we are working on flood susceptibility mapping, future time steps should have no effect on previous time steps. As a result, a single directional LSTM network used in this research as shown in Figure 11

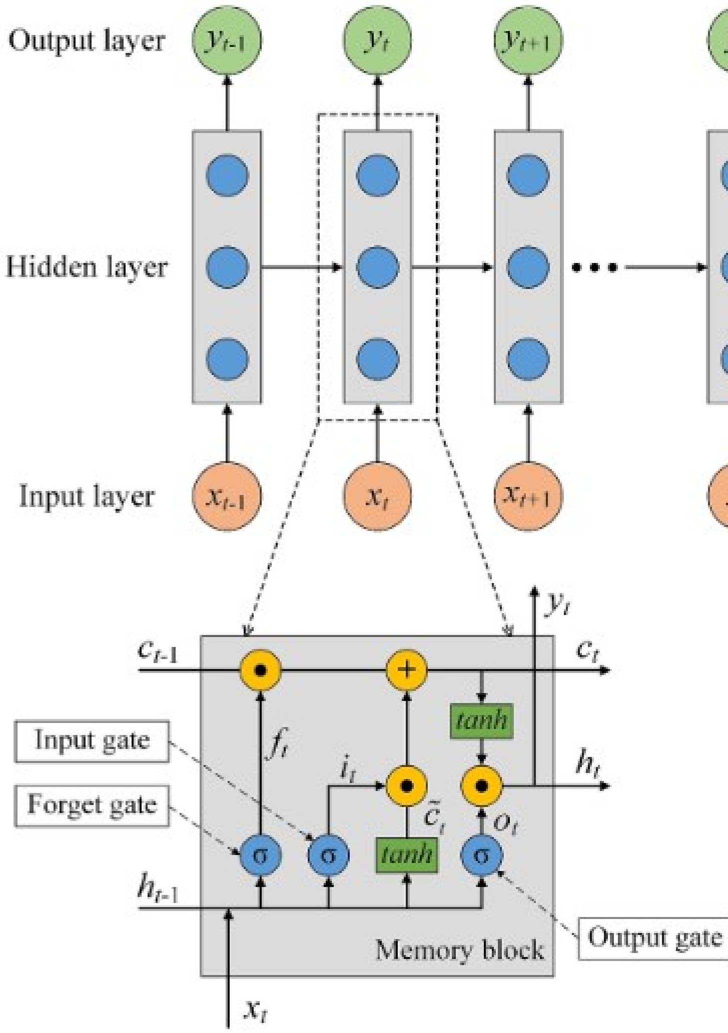


Figure 11. LSTM network architecture

2.7.3.5 Deep neural networks

DNN are generally categorized as ANN algorithms but with multiple (deep) hidden layers (based on the complexity of the features) applying a feed-forward network for the back-propagation training algorithm (Bui et al. 2020; Tien Bui et al. 2020). The use of numerous hidden layers empowers the algorithm to better describe the nonlinear and complex features such as flooding (Tien Bui et al. 2020). Herein, the hidden layer was set to three according to the previous studies and to obtain stronger feature learning (Bui et al. 2020; Tien Bui et al. 2020). DNN is a type of neural network with the Sigmoid function deployed within each neuron in the hidden layers to perform the backpropagation and weighting system. The sigmoid (Bui et al. 2020) activation function f_x is defined by $(1 + e^{-x})^{-1}$. Due to training via the gradient-based algorithm with backpropagation, $ReLU$ is used to avoid dispersing gradient (Tien Bui et al. 2020)

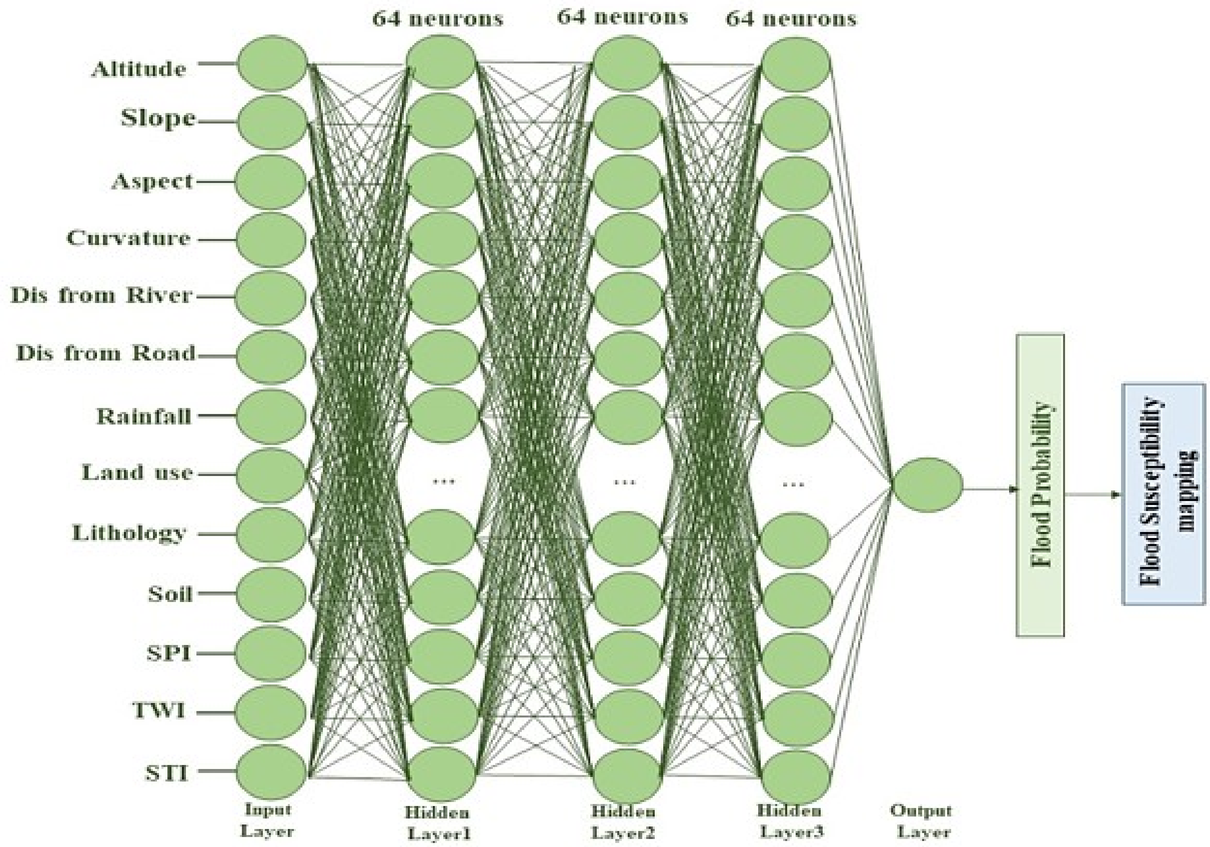


Figure 12. Structure of DNN model

2.8 Validation

2.8.1 Flood Location Map

The accuracy of the produced flood inventory map can be described based on the completeness of the map and precision of the information shown on the map. There are multiple approaches to assess the effectiveness and accuracy of the produced inventory map. However, one of the most popular methods is to create a confusion matrix (Duro, Franklin, and Dubé 2012; Palanisamy, Jain, and Bonafoni 2023; Wu et al. 2023) Using confusion matrix four types of accuracy assessments can be done which are related to different aspects. Various metrics, including Overall accuracy, producer's accuracy, user's accuracy, and kappa coefficients evaluate the quality of the inventory map. Overall accuracy is determined by dividing the number of correctly classified pixels by the total number of pixels in the test dataset. Producer's accuracy indicates how accurately a specific land use and land cover (LULC) type is identified. User's accuracy assesses the probability that a pixel identified as a specific LULC class on the map accurately represents that class. Kappa coefficient is the difference between the actual agreement between the reference data and an automated classifier and the chance agreement between the reference data and a random classifier (Mangkhaseum and Hanazawa 2021; Pande 2022)

2.8.2 Flood Susceptibility Maps

The AUC is considered as one of the most popular methods to assess the efficiency of the generated method which produces both success and prediction rates (Abedi et al. 2022). Prediction and success rates should be evaluated as an essential outcome of every program (Khosravi et al. 2018). Their popularity arises from their comprehensive, clear, and visually appealing method of assessing accuracy. The ranking of each prediction pattern is established by arranging the measured probability index from highest to lowest. As a result, the data points are categorized into 100 classes vertically (y-axis), with 1% increments horizontally (x-axis). Flood events in both training and testing sets within each category are analyzed, and success and prediction rates are subsequently calculated based on these evaluations.

Validation involves comparing the actual flood data with the flood probability map that has been generated. The range of the AUC varies from 0 to 1.0, as the value of 1.0 represents the highest accuracy showing that the model was completely satisfied to predict the disaster occurrence without any biased effect (Duwal et al. 2023; Tehrany, Kumar, et al. 2019). Therefore, the model is considered more accurate and reliable as the AUC value approaches 1.0. The success rate is based on the training dataset, which contains 70% of the flood inventory. The training flood layer is used to create the model and therefore cannot be used to validate its actual performance. Therefore, the model's predictive performance cannot be fully evaluated using only the training data. The prediction rate shows how well the model can predict the flooding in an area (Bhattarai et al. 2024; Saber et al. 2023; Shafapourtehrany et al. 2023; Tehrany and Kumar 2018). As a result, the prediction rate is measured using the testing dataset to assess the model's generalization capability.

CHAPTER 3

STUDY AREA

3.1 **Overview**

This chapter focuses on the characteristics of the NNRB, which is the study area of this research. The background of the NNRB will be presented first to establish an understanding of the basin. The following parts of this chapter cover the flood Inventory, conditioning factor, and the methodology involved in the flood susceptibility modeling.

3.2 **Nam Ngum River Basin (NNRB), Lao PDR**

Lao PDR, situated in Southeast Asia, is rich in water resources, with the Mekong River Basin covering 90% of its territory. The Nam Ngum River, a key watercourse, extends about 420 km from the Xiengkhouang plateau to the Vientiane Plain (Meema et al. 2021). The NNRB, the country's fourth-largest basin, spans 16,931 km^2 and is located between longitudes 102º 25' E and 103º 30' E and latitudes 18º 30' N and 19º 30' N (Adams, Blakers, and Someth n.d.; Dhungana et al. 2023; Meema et al. 2021). It is characterized by hilly terrain with elevations ranging from 2,569 to 114 meters above sea level (MASL) at the Mekong River confluence (Figure 13). The basin, which includes 19 districts across six provinces, contributes 4% to the Mekong's mean annual flow. It has a tropical climate with clear wet and dry seasons, influenced by East Asian and Indian monsoons. From June to October, the rainy season sees heavy rainfall of 1,500 to 3,000 mm annually, exacerbated by Southwest monsoons and Pacific Ocean typhoons causing floods almost every year

Table 2 The area size of province within the NNRB

No	Province	Area size (km^2)	Area size within basin (km^2)	Area size within basin $(\%)^2$
1	Vientiane Capital	3,920	1,928	11.4
2	Luangphabang	16,875	696	4.1
3	Xiengkhuang	14,751	2,858	16.9
4	Vientiane	15,610	6,957	41.1
5	Bolikhamxay	14,863	63	0.4
6	Xaysomboon	8,551	4,429	26.2
	Total	**74,570**	**16,931**	**100**

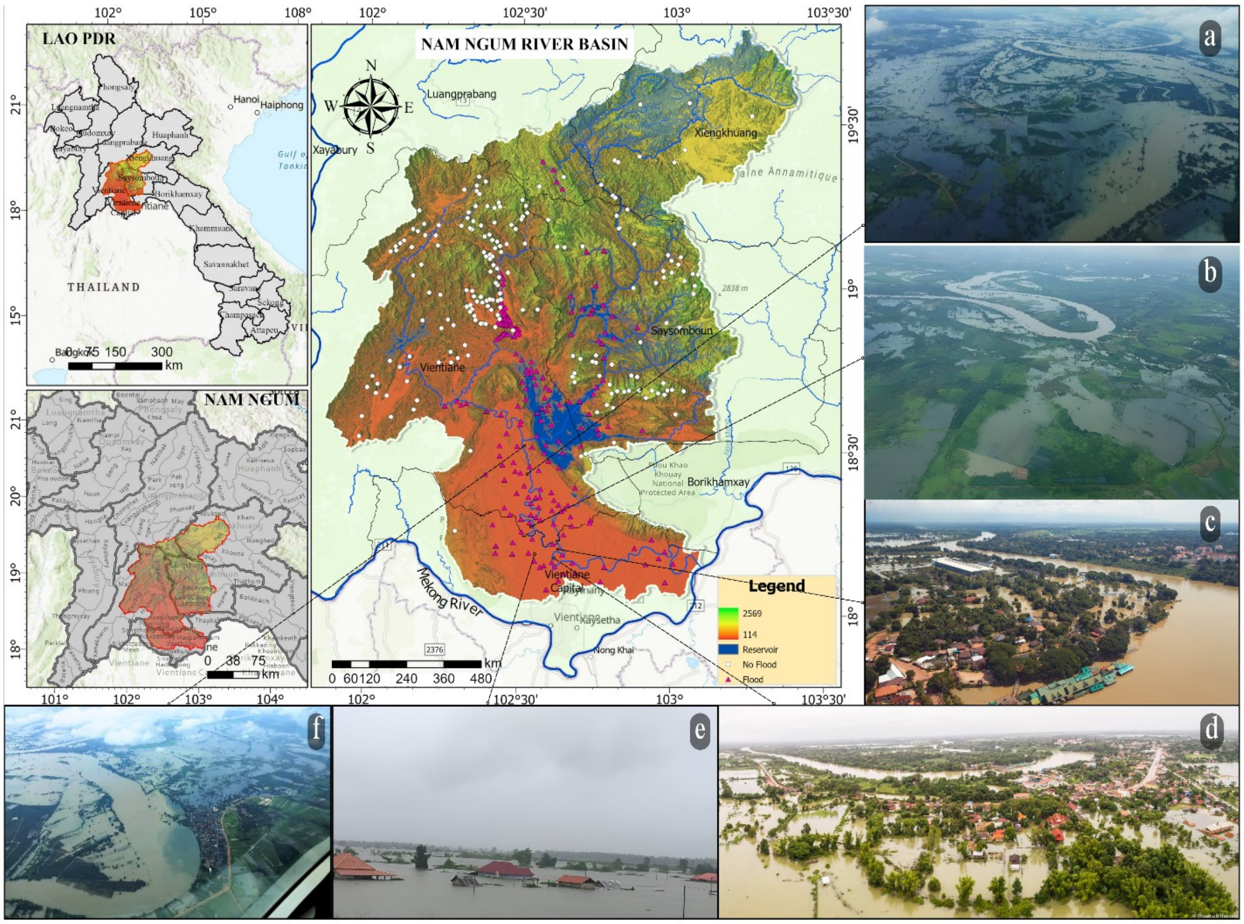


Figure 13. The location of the Nam Ngum River Basin and photographs of flooded areas

3.3 Methodology

The proposed methodology comprises (a) preparation of flood inventory, (b) preparation of flood conditioning factors, (c) Selection of the suitable conditioning factors using multicollinearity test, (d) flood susceptibility modeling using machine learning approaches, (e) comparison and validation of the approaches, and (f) flood susceptibility mapping. The detailed methodological framework is shown in Figure 14.

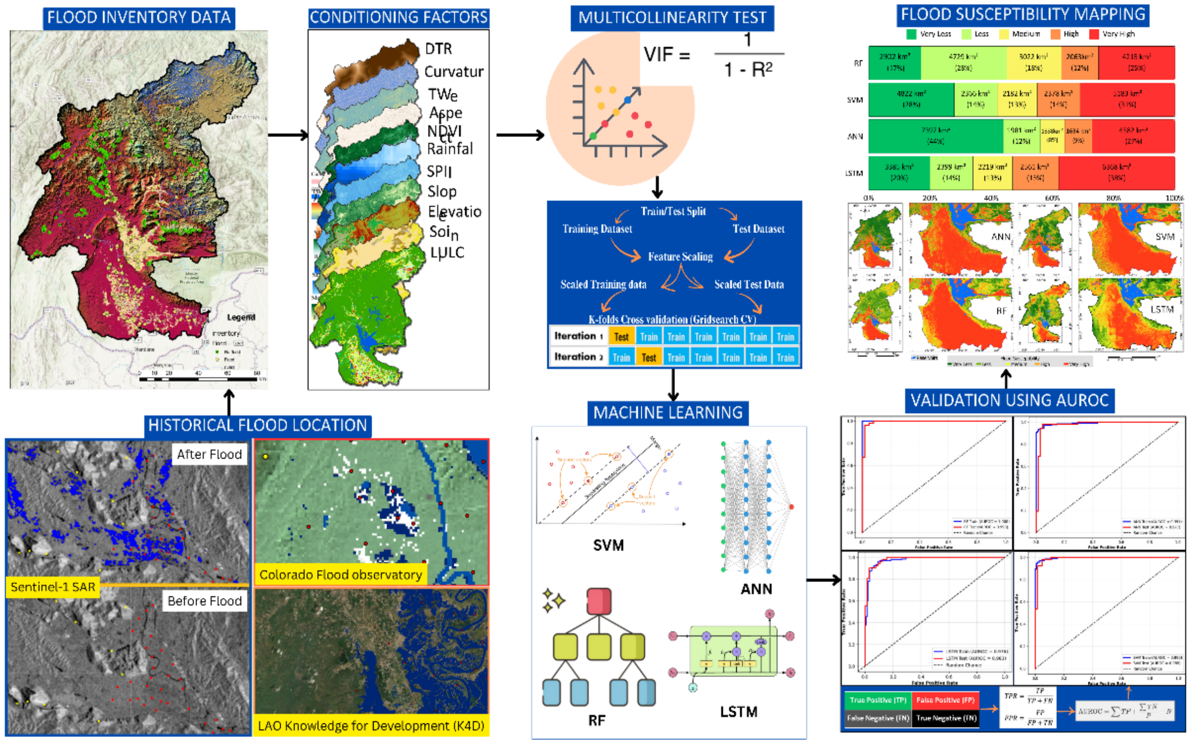


Figure 14. The methodology involved in flood susceptibility modeling.

3.3.1 Data collection and preparation

The flood susceptibility mapping was initiated from data collection. One of the study's main objectives was to use open-source remote sensing data for flood susceptibility mapping. Since data scarcity is the major hindrance in developing countries like Lao PDR, publically available satellite-based remote sensing data are vital (Saha et al. 2021). Historical flood records were collected from online news portals for the tentative location of the flooded areas. Likewise, sentinel-1 SAR images were analyzed in the Google Earth Engine for flood area detection. These data, along with the information from the Knowledge for Development (K4D) (https://apps.k4d.la/explorer) online portal (for years 2018, 2019, and 2020) and the historical flooded area (from 1985 to 2010) from Colorado Flood Observatory (https://floodobservatory.colorado.edu) were initially analyzed to locate flooded areas. For further processing, we acquired ALOS-PALSAR DEM (https://asf.alaska.edu/) to prepare conditioning factors. We extracted factors, such as slope, aspect, elevation, curvature, TWI, and SPI from ALOS-PALSAR DEM. The average rainfall data from 2010 to 2020 were acquired from ERA5 (https://cds.climate.copernicus.eu/), while NDVI was derived from Landsat 8 images (https://earthexplorer.usgs.gov/) using the Google Earth Engine. Land use/landcover (LULC) data were obtained from 10m Sentinel 2 images based on Table 4.

3.3.2 Preparation of flood inventory data

Flood inventory is the foremost and essential step in flood susceptibility modeling (Ahmed et al. 2022; Hansana et al. 2023; Khosravi et al. 2018). The historically flooded areas were used to prepare the flood inventory. To locate the flood points, random points were generated in the flooded area detected using Sentinel-1 SAR images in the Google Earth Engine. Sentinel-1 SAR satellite data of 10m resolution excels in capturing images regardless of time and weather (Hamidi et al. 2023; Martinis, Plank, and Ćwik 2018; Twele et al. 2016). We used Level 1 GRD data in IW Swath mode (Table 1), which has a 250km swath width (Askar et al. 2022; Nagler et al. 2015). After visual inspection of the generated flood points and assessment of the change in accuracy of the modeling results, we selected only 390 past flood points. The non-flood locations were selected visually, where the probability of flooding is none—for example, the hilltops and ridges of the mountains. Equal numbers of flood and non-flood locations were used for the inventory for increased accuracy, as suggested by (Buitinck et al. 2013; Tang et al. 2020; Towfiqul Islam et al. 2021)Values of 1 as flood and 0 as non-flood points were assigned for model training and testing, using 70% of the data for training and 30% for testing.

Table 3 Description of Sentinel-1 data

Sensors	Sensor Mode	Polarization	Pass Direction	Dates of Acquisition
Sentinel-1A	Interferometry wide swath (IW)	VV, VH	Ascending	August 03 to September 07, 2018 July 25 to September 07, 2019 August 05 to October 22, 2020

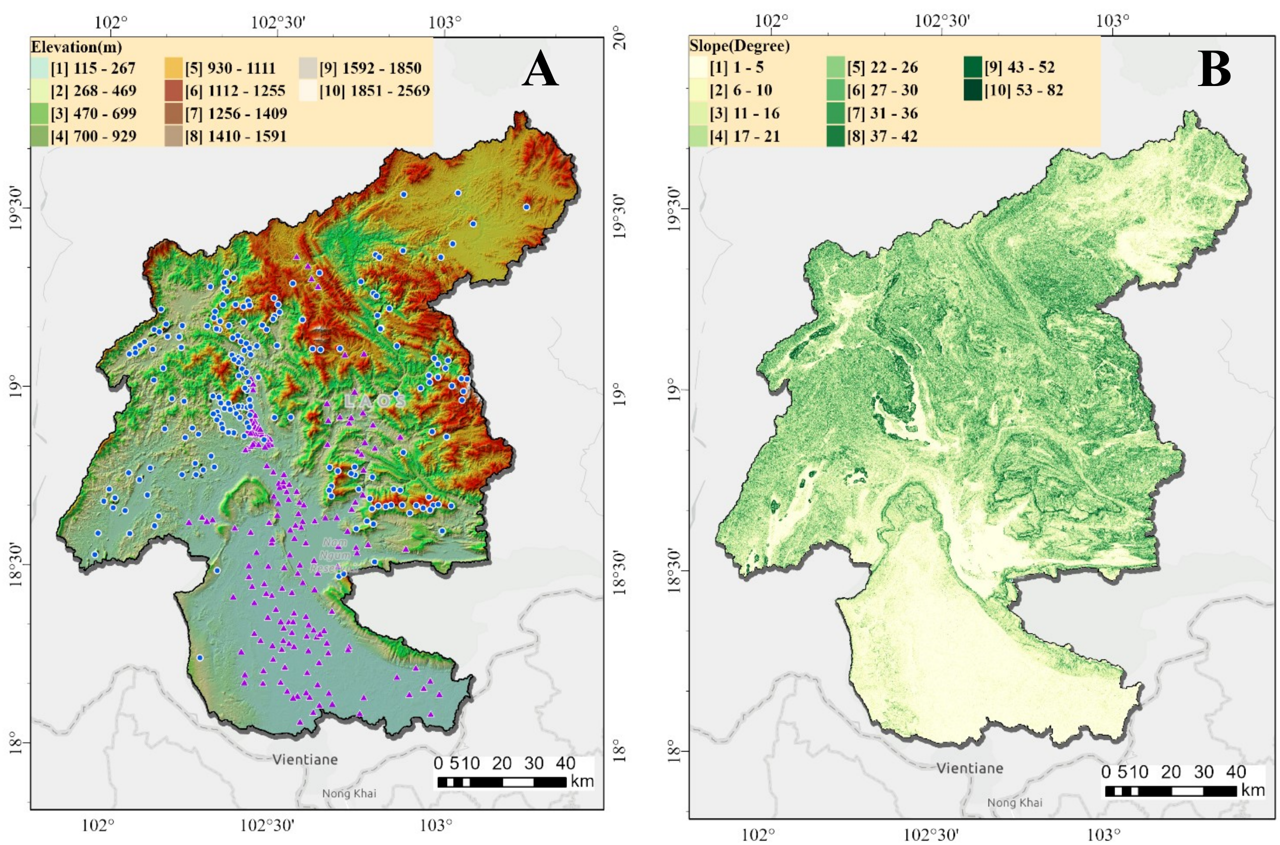
3.3.3 Flood conditioning factors

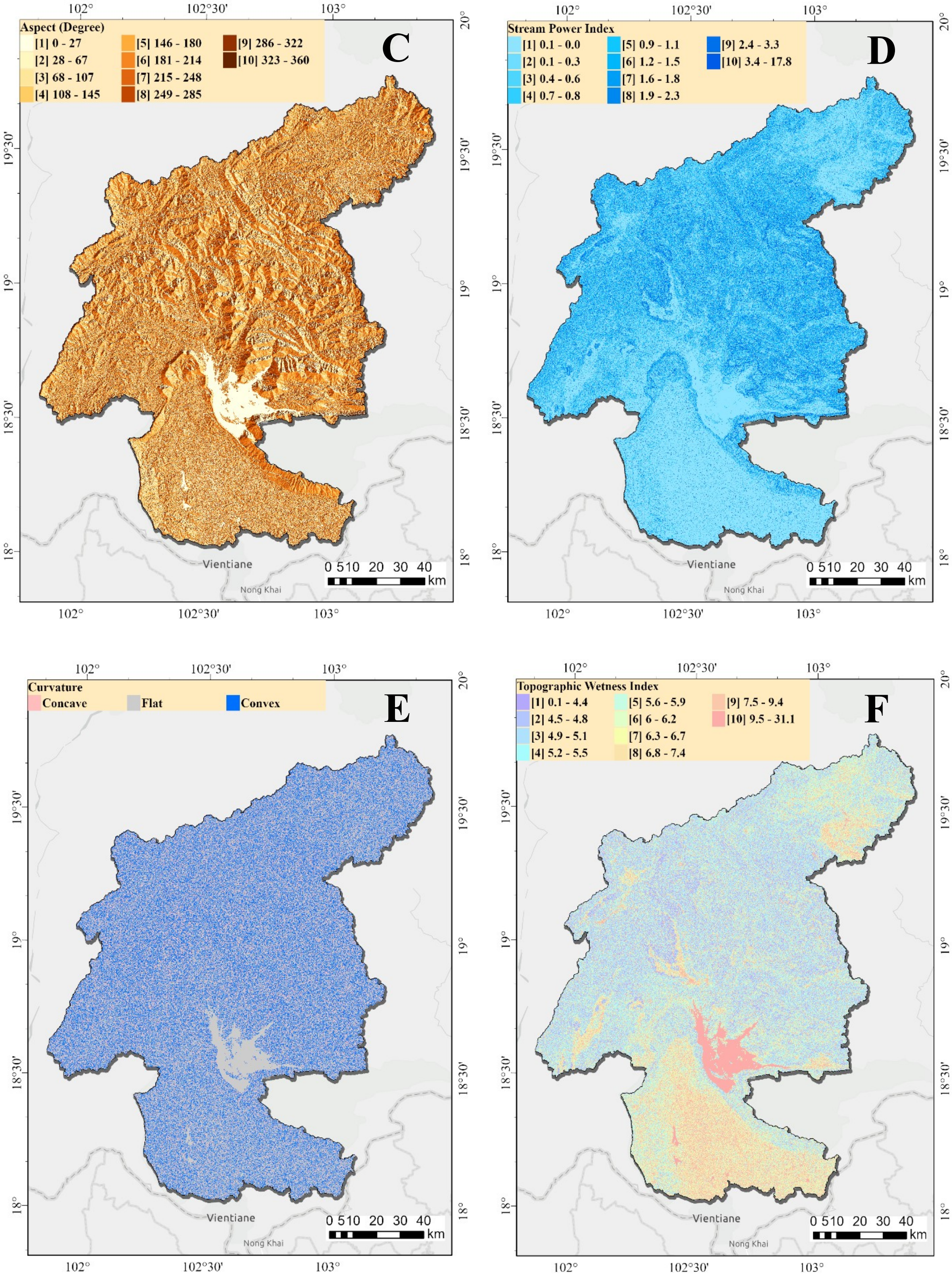
The flood conditioning factors are crucial in mapping flood susceptibility (Mojaddadi et al. 2017). Identification of the factors that play a vital role in flood susceptibility mapping. However, the selection of appropriate conditioning factors depends on the nature of the particular region (Amiri et al. 2024). Eleven flood conditioning including elevation, slope, aspect, curvature, topographic wetness index (TWI), stream power index (SPI), distance to the river (DTR), normalized difference vegetation index (NDVI), land use/land cover (LULC), rainfall, and soil type were selected based on literature (Dodangeh et al. 2020; Khosravi et al. 2019; Shafizadeh-Moghadam et al. 2018). Parameters like slope, curvature, aspect, DTR, drainage density, SPI, and TWI were derived from the ALOS-PALSAR DEM of 12.5m resolution, and other data were acquired from different sources, as shown in Table 4.

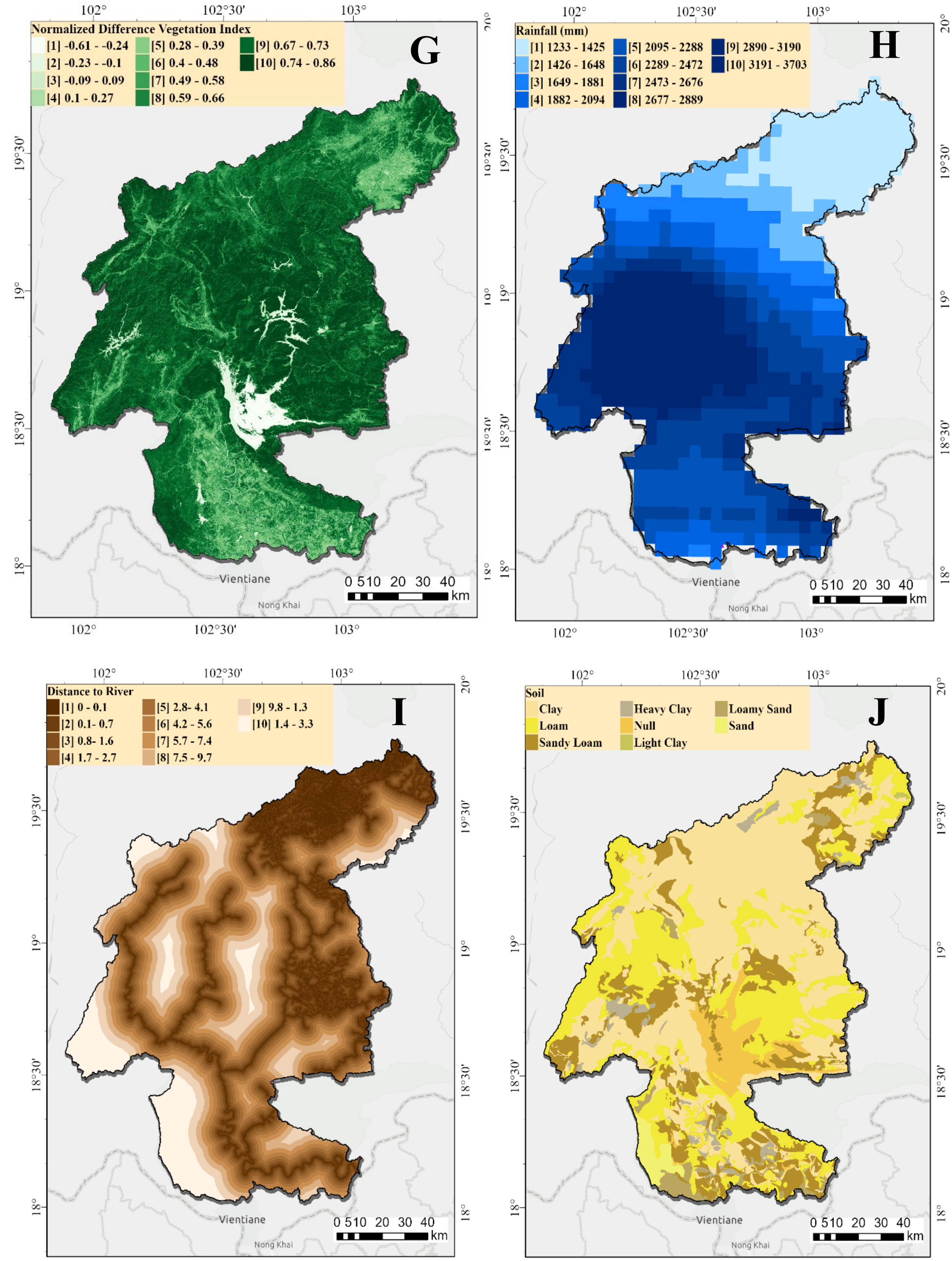
The spatial and temporal resolution of the conditioning factors affects the precision of the underlying results (Saha et al. 2021). However, the study by (Avand et al. 2022) states that spatial resolution alone does not affect the model's prediction significantly, but the type of model and local condition are affected remarkably. For this study, we focused on the publically available data with better spatial and temporal resolution for the analysis. Since most of the conditioning factors are derived from DEM data, other conditioning factors were resampled to the resolution of the DEM data, i.e. 12.5m, to obtain the final results. The conditioning factors were reclassified to create class data. Based on previous studies (Bhattarai et al. 2024; Chapi et al. 2017; Duwal et al. 2023; Mangkhaseum et al. 2024; Shahabi et al. 2020; Tien Bui et al. 2018) the natural break was used to reclassify elevation, slope, aspect, and rainfall. NDVI. TWI, SPI, and DTR were reclassified using quantile division, and LULC, soil, and curvature were reclassified manually.

Table 4 Description of data used in the study

Primary Data	Original Data Format	Data Source	Spatial Resolution	Derived Data
DEM	Raster	ALOS PALSAR DEM (https://search.asf.alaska.edu/)	12.5m x 12.5 m	Elevation, Slope, Aspect, Curvature, TWI, SPI, DTR
Landsat 8	Raster	USGS	30m x 30 m	NDVI
Sentinel 2	Raster	ESA	10m x 10 m	LULC
ERA5	Raster	www.cds.climate.copernicus.eu	30km x 30km	Rainfall
Environmental map	Vector	FAO	1:500,000	Soil Map




Aspect (Degree)
[1] 0 - 27
[2] 28 - 67
[3] 68 - 107
[4] 108 - 145
[5] 146 - 180
[6] 181 - 214
[7] 215 - 248
[8] 249 - 285
[9] 286 - 322
[10] 323 - 360
C
Stream Power Index
[1] 0.1 - 0.0
[2] 0.1 - 0.3
[3] 0.4 - 0.6
[4] 0.7 - 0.8
[5] 0.9 - 1.1
[6] 1.2 - 1.5
[7] 1.6 - 1.8
[8] 1.9 - 2.3
[9] 2.4 - 3.3
[10] 3.4 - 17.8
D
Curvature
Concave
Flat
Convex
E
Topographic Wetness Index
[1] 0.1 - 4.4
[2] 4.5 - 4.8
[3] 4.9 - 5.1
[4] 5.2 - 5.5
[5] 5.6 - 5.9
[6] 6 - 6.2
[7] 6.3 - 6.7
[8] 6.8 - 7.4
[9] 7.5 - 9.4
[10] 9.5 - 31.1
F
102°
102°30'
103°
20°
19°30'
19°
18°30'
18°
Vientiane
Nong Khai
0 510 20 30 40
km


G
Normalized Difference Vegetation Index
[1] -0.61 - -0.24
[2] -0.23 - -0.1
[3] -0.09 - 0.09
[4] 0.1 - 0.27
[5] 0.28 - 0.39
[6] 0.4 - 0.48
[7] 0.49 - 0.58
[8] 0.59 - 0.66
[9] 0.67 - 0.73
[10] 0.74 - 0.86
H
Rainfall (mm)
[1] 1233 - 1425
[2] 1426 - 1648
[3] 1649 - 1881
[4] 1882 - 2094
[5] 2095 - 2288
[6] 2289 - 2472
[7] 2473 - 2676
[8] 2677 - 2889
[9] 2890 - 3190
[10] 3191 - 3703
I
Distance to River
[1] 0 - 0.1
[2] 0.1- 0.7
[3] 0.8- 1.6
[4] 1.7 - 2.7
[5] 2.8- 4.1
[6] 4.2 - 5.6
[7] 5.7 - 7.4
[8] 7.5 - 9.7
[9] 9.8 - 1.3
[10] 1.4 - 3.3
J
Soil
Clay
Loam
Sandy Loam
Heavy Clay
Null
Light Clay
Loamy Sand
Sand
Vientiane
Nong Khai
0 510 20 30 40
km
102°
102°30'
103°
18°
18°30'
19°
19°30'
20°

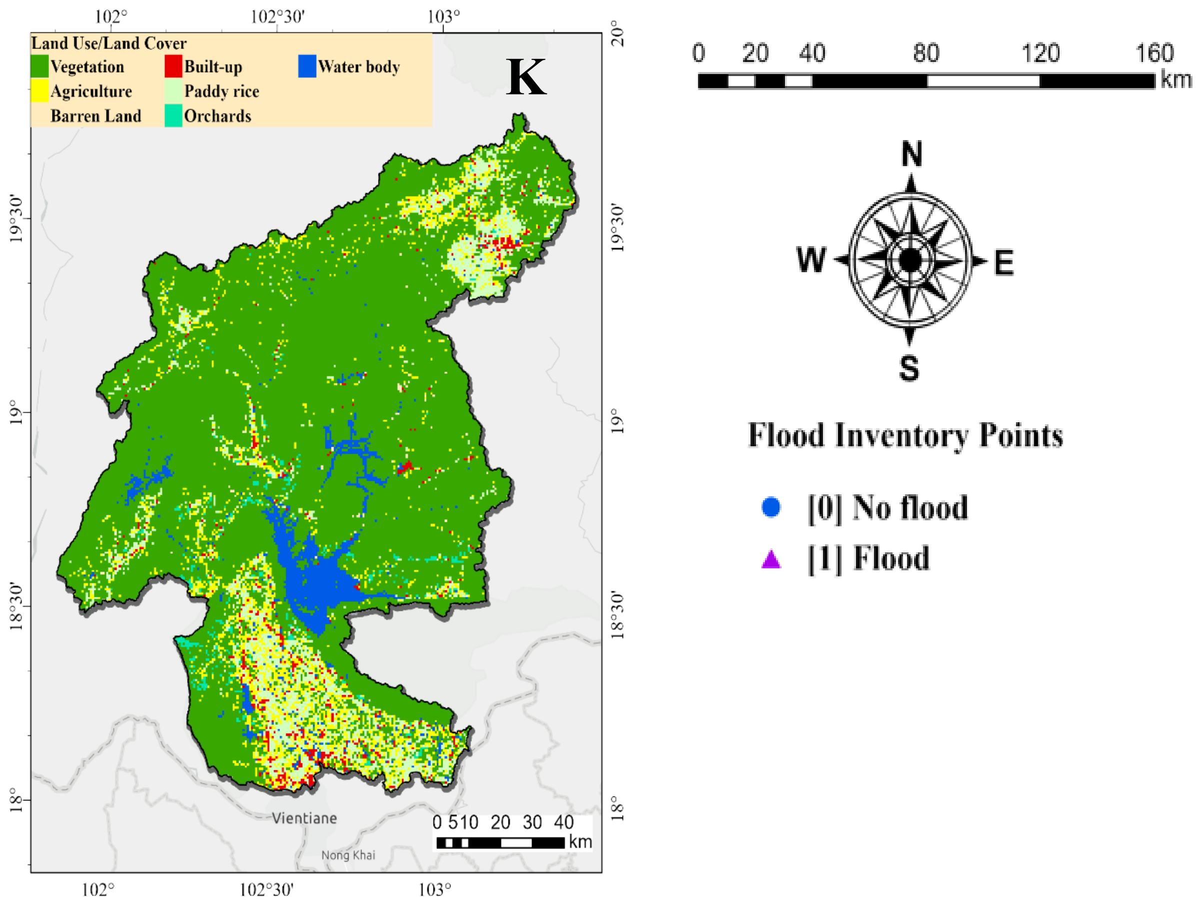


Figure 15. Flood conditioning factors

a. Elevation is defined as the difference in height from sea level, which is related to climatic conditions (Shafizadeh-Moghadam et al. 2018). According to (Tehrany et al. 2015), it is one of the most important factors in flood-related studies. More specifically, higher elevation implies a lower probability of a flood event, as the runoff is directed downstream of the basin (Chapi et al. 2017; Khosravi et al. 2018). The elevation ranges between 115 and 2569 meters. It was manually classified into ten categories: 115-267m, 268-469m, 470-700m, 700-929m, 930-111m, 1112-1255m, 1256-1409, 1410-1591m, 1592-1850m, and 1851-2569m. Al-Juaidi et al., 2018; Khosravi et al., 2016
b. The slope angle of the terrain was shown by (Arabameri et al. 2019) and (Costache and Tien Bui 2019) to be the most important factor affecting the flooding process. It controls the velocity of water on the soil surface and affects the processes of runoff and infiltration inversely (Al-Juaidi et al. 2018; Khosravi et al. 2016). Areas of high susceptibility to

flooding are typically flat areas (Tehrany, Pradhan, and Jebur 2013). It was divided into five categories using the Natural Breaks method which are 1°, 1°-5°, 6°-10°, 11°-16°, 17°-21°, 22°-26°, 27°-30°, 31°-36°, 37°-42°, 43°-52° and 53°-82°

c. Aspect refers to the horizontal direction in which the slope faces (Mojaddadi et al. 2017). Areas that are oriented towards the sun's rays are flooded with less frequency. Area that receives less solar radiation will have more soil moisture therefore more runoff directed downstream increasing the risk of flooding (Rezaie et al. 2022; Yariyan et al. 2020). Classified into the categories which are 0°, 0°-27°, 28°-67°, 68°-107°, 108°-145°, 146°-180°, 181°-214°, 215°-248°, 249°-285°, 286°-322° and 323°-360°
d. Stream Power Index (SPI) expresses the erosive capacity of a river (Poudyal et al. 2010). A high SPI value implies that the flow has high power, and areas with low SPI values receive the flow, increasing the probability of flooding (Chapi et al. 2017). According to (Moore, Grayson, and Ladson 1991), the index is calculated using Equation 3.1

$$SPI = (\alpha * \tan \beta) \qquad (3.1)$$

where, α is the specific basin area and β is the slope angle in degrees at the point. The Natural breaks method created 10 classes which are >0.1, 0.1-0.0, 0.1-0.3, 0.4-0.6,0.7-0.8, 0.9-1.1, 1.2-1.5, 1.6-1.8, 1.9-2.3, 2.4-3.3, and 3.4-17.8 Khosravi et al., 2019; Tehrany, Pradhan, & Jebur, 2015)
e. The curvature represents the ground surface shape, such as flat, convex, and concave areas, which gives useful information that flat areas are more prone to flooding, as noted earlier on the slope (Khosravi et al. 2019; Tehrany et al. 2015). The curvature map, derived from the digital elevation model (DEM), includes three classes, as depicted in (Fig. 11e). In these classes, negative (-) values represent convex, positive(+) values represent concave, and values corresponding to 0 represent flat (Youssef et al. 2016)
f. The Topographic Moisture Index (TWI), proposed by (BEVEN and KIRKBY 1979), expresses the amount of water accumulated per unit area (pixel) in a catchment, taking into account the downward flow due to gravitational force. A higher value of the index implies a higher susceptibility to flooding in that area (Khosravi et al. 2019). It is calculated using Equation 3.2

$$TWI = \ln\left(\frac{\alpha}{\varepsilon\varphi\beta}\right) \qquad (3.2)$$

where, α is the cumulative area upstream of the point draining to it and β is the slope angle in degrees. It was grouped into 10 classes using the Natural Breaks method >0. 1, 0.1.-4.4, 4.5-4.8, 4.9-5.1, 5.2-5.5, 5.6-5.9, 6-6.2, 6.3-6.7, 6.8-7.4, 7.5-9.4, 9.5-31.1

g. The Normalized Difference Vegetation Index (NDVI) plays a crucial role in studies on flood susceptibility. Vegetation acts as a protective factor against flooding, and this index is the best method for estimating its cover and density (Chowdhuri et al., 2020). The NDVI was computed from Landsat 8 OLI/TIRS Collection 2 Level-2 satellite images sourced from the USGS Earth Explorer, utilizing Google Earth Engine. The equation that calculates the index is (3.3):

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \qquad (3.3)$$

where NIR and RED are the values of channels NIR and RED respectively. It takes values between -1 and +1 (TUCKER and SELLERS 1986). Higher value equals healthier and denser vegetation, therefore able to retain more water (Tang et al. 2020). Using Natural breaks method, it was divided into 10 classes which are >-0.61, (-0.61)- (-0.24), (-0.23)- (-0.2), (-0.09)- (-0.09), 0.1-0.27, 0.28-0.39, 0.4 -0.58, 0.59-0.66, 0.67-0.73 and 0.74-0.86

h. Rainfall is a prerequisite for a flood event to occur (Saber et al. 2022, 2023). The characteristics of rainfall vary depending on the climatic and topographic features of the region (Gudiyangada Nachappa et al. 2020). Data were obtained from ERA5, monthly for the years 2010-2020. In a GIS environment, the mean annual rainfall for each year was calculated and then the mean rainfall for all years was calculated. The spatial resolution of the data is 20sq km (Lavers et al. 2022). The classification was done in 10 classes using the Natural breaks method: 1233-1425, 1426-1648, 1649-1881, 1882-2094, 2095-2288, 2289-2472, 2473-2676, 2677-2889, 2890-3190 and 3191-3703

i. Distance to river is one of the most important factors for the occurrence of a flood, according to the results of several studies (El-Haddad et al. 2021; Fang et al. 2021; Shafizadeh-Moghadam et al. 2018) . The hydrographic network, with its rivers and streams, serves as the main pathway for flood discharge, making areas close to it more susceptible to flooding (Opperman et al. 2009). The hydrographic network was extracted from a DEM, and the distance map was constructed using the Proximity (raster distance) tool of GDAL in ArcGIS. Manual classification into the following six classes was performed: >0, 0-0.1, 0,1-0.7, 0.8-1.6, 1.7-2.7, 2,8-4.1, 4.2-5.6, 5,7-7.4, 7.5-9.7, 9.8-1.3, and 1.4 -3.3

j. Soil type affects water absorption, akin to the concept previously discussed regarding proximity to roads. For example, the type of soil can influence the amount of water transported during a flood, just as concrete roads, which do not absorb water, contribute to flood formation (Rahmati et al. 2016). According to (Huang, Wu, and Zhao 2013), the type of soil surface plays a crucial role in water infiltration. Apart from the soil's ability to absorb water, the volume of water it retains also affects flood formation. Thus, a soil type map (Fig. 13j) was employed as a factor influencing flood susceptibility mapping. This map was derived from vector data obtained from the Food and Agriculture Organization of the United Nations (FAO) for the Nam Ngum River Basin (NNRB) study area in Lao PDR.

k. Land use and land cover (LULC) directly or indirectly influence processes such as subsidence, runoff, and evapotranspiration, making them critical factors in flooding(Avand, Moradi, and Lasboyee 2021; El-Haddad et al. 2021; Samanta, Pal, and Palsamanta 2018; Talukdar et al. 2020). Areas with vegetation experience reduced surface runoff compared to built-up areas, which prevent water from infiltrating into the soil (Hammami et al. 2019). The land use map was created using Sentinel2 Image from the European Space Agency. It is categorized into the following types: Built-up Areas, Vegetation, Agriculture, Barren land, Paddy rice, Orchards and water body (Figure 15k)

In flood susceptibility mapping topological, geological, and hydrological factors can be used. These factors play a role in causing flooding in a particular area. There is no particular rule exists to specify how many conditioning factors are enough in flood susceptibility analysis (Kaya and Derin 2023). Moreover, no framework is available for the choice of conditioning factors. These factors are usually selected based on the expert's knowledge. Some researchers believed that as the number of conditioning factors increases, the accuracy of the produced susceptibility map increases (Tehrany, Jones, and Shabani 2019). On the other hand, other case studies represent that a small number of conditioning factors are adequate to provide precise susceptibility maps (Kavzoglu, Kutlug Sahin, and Colkesen 2015). If all the pertinent influencing factors are examined and incorporated into the model, it will be comprehensive but may require extensive data, which can be challenging to obtain. It is thus important to determine which variables can be eliminated without harming the model's precision (Sanyal and Lu 2004).

3.4 Python Programming Language – Libraries

The programming language chosen for this work is Python. Programming languages are artificial languages used to write program commands. A programming language has strictly defined syntax and semantics. Correct syntax is necessary for a program to run, while semantics determine the computational procedures the program will perform.

Python was developed by the Dutch programmer Guido von Rossum in the early 1990s and is considered the successor to the ABC programming language, as it was the main source of inspiration for its creation. Its great advantage is the ability to expand easily, allowing elements to be added according to the needs of each task. It is used internationally for both educational and commercial purposes. Python is a high-level, easy-to-learn, open-source, and general-purpose language, making it useful for both beginners and experienced developers. It is compatible with most commercial operating systems, including Windows, Unix/Linux, and Mac OS.

3.4.1 Libraries

A library in Python is a piece of programming code that can be reused across different programs. This means there is no need to write the same code repeatedly, making programming in Python simpler and more practical. When we link a library to our program and run it, the linker searches for the library on the system and extracts its functions. Some libraries are essential in the field of machine learning.

The libraries used in this work are the following:

- Pandas is a powerful and easy-to-use tool for data analysis and management. It can import various types of data, such as CSV files and tables from SQL. It has been open source since 2009 (https://pandas.pydata.org/).
- NumPy is an open-source library that enables scientific calculations in Python. It provides functions that work with tables, primarily using "arrays." Unlike lists, arrays are stored in a contiguous sequence in computer memory, allowing programs to manipulate them more quickly. It was created in 2005 (https://numpy.org/doc/stable/index.html).
- Scikit-learn is a free machine learning library for Python. It includes classification, regression, and clustering algorithms. It was developed to work with the two libraries mentioned above. It was created in 2007 (https://scikit-learn.org/).
- Matplotlib was created in 2003 for making static, interactive, and live charts and graphs (https://matplotlib.org/).
- Mealpy is the largest library of nature-inspired metaheuristic algorithms for optimization problems. It has a simple structure, allowing researchers to access optimization algorithms easily and quickly (https://pypi.org/project/mealpy/).

- Mapclassify is a Python library for thematic map classification. It helps classify and visualize spatial data by determining class numbers and mapping observations, supporting various classification methods for effective data visualization. (https://pypi.org/project/mapclassify/).

3.5 Optimization

3.5.1 General-purpose optimization

The concept of optimization is present in everyday life and is widely understood. Optimization refers to any problem or process where improvements can be made. It can involve simple decisions, such as choosing between a straight path or a turn in a route, or more complex scenarios, such as determining the optimal conditions for machine operation (French 2018). More generally, optimization underpins the decision-making process.

The function that should be maximized or minimized is called the objective function. Cost minimization is a common optimization problem, so the methods used aim to minimize the objective function. The function is expressed as:

$$minf(x), \quad x = (x_1, x_2, x_3, \ldots., x_n)$$

where f is the objective function and x are the variables or parameters of the problem. For maximization problems such as that of the work as it was chosen to maximize the measure of overall accuracy (accuracy), the opposite or inverse value of the objective function is minimized (Manolis, 2021). The variables (design variable) are the quantities that can be changed to change the objective function, but without being able to take any value. The limits (upper and lower value) within which they can take a certain value are called constraints. In machine learning, these are the hyperparameters of algorithms. The variables are located in the so-called design space.

The design space is multidimensional, with one dimension for each variable and an additional one for the objective function. The solution to the problem involves searching for the point in the design space that corresponds to the maximum or minimum value of the objective function while satisfying the constraints. If the design space is two or three-dimensional, indicating one or two variables, it is often possible to construct a graph of the design space (French 2018).

3.5.2 Classification of optimization problems

Optimization problems are primarily classified into linear (linear programming, LP) and non-linear (non-linear programming, NLP) categories. Linear programming methods are used to solve problems where both the objective function and all constraints are linear functions of the variables. The term "programming" in this context does not refer to writing computer programs but to the process of finding the optimal solution. If a problem can be efficiently expressed using linear equations, these methods can determine the optimal solution. However, real-world problems are often complex and cannot always be accurately represented by mathematical linear equations

The general linear programming problem can be formulated in its simplest form as finding values $x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0$ such that min_z or max_z is achieved (Dantzig and Thapa 2013).

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = z(\text{min})$$
$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b_1$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_n$$

The most well-known linear programming method is the Simplex method, created in 1947 by George Dantzig. He addressed the problem of planning the allocation and distribution of military resources, which is why the concept of "planning" is central to the method

Often, the assumption that the objective function and the variable functions are linear cannot be applied. In such cases, the problem is addressed as a non-linear one. The general form of a non-linear problem involves searching for $x = (x_1 + x_2 + \ldots + x_3)$ to maximize:

$f_{(x)}$ under the constraints $g_i(x) \leq b_i, \gamma\alpha\ i = 1,2, \ldots, m$

and $x \geq 0$

where, $f_{(x)}$ and $g_{(x)}$ given functions of the variables

There are many different non-linear programming problems, depending on the characteristics of the functions of the variables, and thus a variety of algorithms are used. Some problems with simple function forms can be solved easily, while others, even smaller ones, can be particularly challenging (Hillier and Lieberman 2015)

Optimization problems can be classified in additional ways, as mentioned by (Fang and Wang 2021):

By Variables: Each variable can take a different number of values. Continuous variables can take any real value, while binary variables can only take two values. Various types of variables result in different optimization problems (e.g., continuous optimization, binary optimization).

By Objective Function: The objective function can be either a scalar or a vector, leading to classifications of "single-objective optimization" and "multi-objective optimization" problems.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 **Overview**

This chapter presents the findings obtained from multicollinearity test and information gain ratio, statistical evaluation method, the area under the Receiver Operating Characteristics (AUROC), analysis of flood susceptibility using machine learning methods and evaluation and comparison of models. Also, provides a detailed analysis of flood risk assessment in Nam Ngum River Basin, Lao PDR

4.2 **Selection and evaluation flood conditioning factors using multicollinearity test and information gain ratio**

In this study, we first identified correlations in the data using a multicollinearity test (Alin 2010). A high correlation between independent variables can lead to errors in machine learning models and affect the accuracy of the final flood susceptibility map. We used the variance inflation factor (VIF) and tolerance values to detect and eliminate multicollinearity. Factors indicating multicollinearity (tolerance less than 0.10 and VIF greater than 10) (Arabameri et al. 2019; Baig et al. 2022; Mehravar et al. 2023) were should be removed. The VIF is calculated using equation (4.1) as

$$VIF_i = \frac{1}{1-R_i^2} \qquad (4.1)$$

where R_i^2 is the coefficient of determination obtained by regressing the factor i on all other factors in the analysis (Miles 2014).

The Information Gain Ratio (IGR) (Ghorbanzadeh et al. 2019; Talukdar et al. 2020; Towfiqul Islam et al. 2021) test, a widely used feature selection technique, evaluates the significance of factors in flood events. A higher IGR value indicates a more decisive influence of the factor on the target variable (Al-Abadi 2018; Bhattarai et al. 2024; Mangkhaseum et al. 2024; Saber et al. 2023)

4.3 Evaluation metrics for flood susceptibility models

Generating a flood susceptibility map with a machine-learning algorithm involves a binary classification technique. In this approach, a chosen pixel from the study area is categorized as either flood pixels (P) as 1 or non-flood (N) as (Duwal et al. 2023; Tehrany et al. 2015; Towfiqul Islam et al. 2021). The selected machine-learning algorithm may not always yield accurate predictions during the classification process. The model's performance is assessed using evaluation metrics, such as the area under the Receiver Operating Characteristics (AUROC) (Equation 4.5), Kappa Score, and Accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FM} \tag{4.2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.3}$$

$$Specificity = \frac{TP}{FP + TN} \tag{4.4}$$

$$AUROC = \sum TP + \frac{\sum TN}{P} + N \tag{4.5}$$

$$MSE = \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{n} \tag{4.6}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{n}} \tag{4.7}$$

where n denotes the sum of the data points, and X_i and Y_i for the observed and predicted values, respectively. In every situation, the smallest MAE, MSE, and RMSE value signifies greater model fitness. Lower MAE, MSE, and RMSE values always indicate greater model fitness.

In flood susceptibility mapping, correctly classified flood pixels and non-flood pixels are termed True Positives (TP) and True Negatives (TN). Conversely, inaccurately identified flood pixels and non-flood pixels are labeled as False Positives (FP) and False Negatives (FN) (Chapi et al. 2017; Costache et al. 2020; Duwal et al. 2023; Janizadeh et al. 2019; Mangkhaseum et al. 2024) AUROC is a major evaluation criterion for classification model performance (Tien Bui et al. 2018). It represents the degree or measure of separability (Davis and Goadrich 2006; Towfiqul Islam et al. 2021). The AUROC is a tool for evaluating model performance, with the y-axis representing the true positive rate (sensitivity) (Equation 4.3) and the x-axis representing the false positive rate

(1 − specificity) (Equation 4.4) (Hanley 1989). It is a quantitative statistic to assess the model's performance; a value closer to 1 indicates superior model performance, and 0.5 represents an inaccurate model.

4.4 Feature selection and influence of conditioning factors on flood

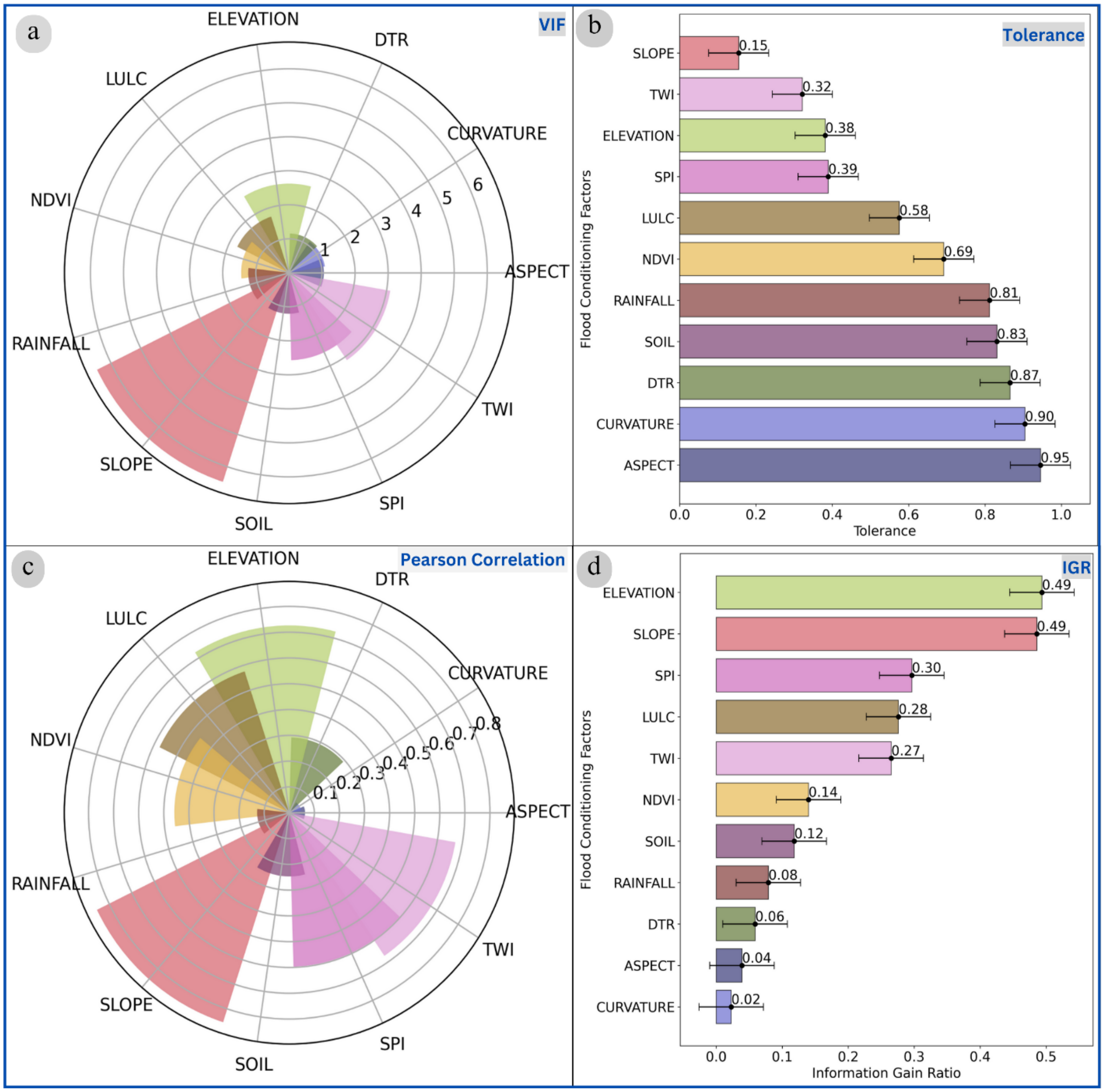


Figure 16. Assessment of flood conditioning factors based on (a) Variance inflation factor, (b) Tolerance, (c) Pearson correlation, and (d) Information gain ratio

For the feature selection multicollinearity and Information gain ratio values are considered (Arora et al. 2021; Bhattarai et al. 2024; Duwal et al. 2023). In this study, VIF and Tolerance were used to evaluate multicollinearity. Likewise, Pearson's correlation test results provide insight into the correlation between flooding and conditioning factors. The high correlation among the factors affects the prediction result in flood susceptibility mapping. The VIF values were <10, and tolerance values were >0.1 for all the selected 11 factors, so no multicollinearity exists. For further analysis, none of the conditioning factors were removed (Figures 16). The highest VIF was obtained for slope (6.46), followed by TWI (3.11), elevation (2.62), and SPI (2.57). The tolerance values were highest for Aspect (0.95), followed by curvature (0.90), DTR (0.87), Soil (0.83), and Rainfall (0.81), and lowest values for Slope (0.15), TWI (0.32) and elevation (0.38). These results indicated that slope and elevation are the critical factors in flooding, followed by TWI, SPI, LULC, NDVI, DTR, Soil, Rainfall, Aspect, and Curvature. Based on IGR values, slope (0.49) and elevation (0.49) are highly influencing factors in the NNRB compared to other factors. Likewise, SPI (0.30), LULC (0.28), TWI (0.27), NDVI (0.14), and soil (0.12) demonstrated a slightly strong influence on the flood prediction compared to rainfall (0.08), DTR (0.06), Aspect (0.04), and Curvature (0.02) showed less influence on flood susceptibility. The Pearson correlation was highest for slope (0.85), then elevation (0.73), TWI (0.67), and SPI (0.60) signifying that these factors are more correlated to flooding in NNRB (Figure 16 (d)). Based on the values of VIF, tolerance, Pearson Correlation, and IGR, it can be observed that the topographic factors like elevation, slope SPI, and TWI followed by landcover factors like LULC and NDVI are influential in flooding in NNRB compared to meteorological, geological, and location factors, such as rainfall, DTR, soil, and aspect.

4.5 Flood susceptibility mapping

The results obtained from the four machine learning models are presented in Figures 17 and 8. The study revealed that for the LSTM model, 38% (6368 km^2) of the area lies in a very high flood susceptibility zone, followed by 15% (2561 km^2) in highly susceptible, 13% (2219 km^2) in medium, 14% (2399 km^2) in less and 20% (3385 km^2) in very less flood susceptible zone. Similarly, SVM shows that 31% (5183 km2) of the area lies under very high susceptibility area, 14% (2378) in high, 13% (2182 km^2) in medium, 14% (2366 km^2) in less, and 28% (4822 km^2) in very less flood susceptible area. The ANN model shows 27% (4582 km2) in very high, 9% (1634 km^2) in high, 8% (1338 km^2) medium, 12% (1981 km^2) less, and 44% (7397 km^2) in a very less susceptible area. Lastly, RF shows that 25% (4215 km²) lies in very high, 12% (2063 km2) in high, 18% (3022 km²) in medium, 28% (4729 km2) in less, and 17% (2902 km²) in very less susceptible areas. The spatial patterns of flood susceptibility of the RF and ANN are similar. For the downstream part of the basin, the very and high flood susceptibility patterns are similar for RF, ANN, and SVM compared to LSTM. The pattern for the upstream part of the basin is similar for SVM and RF, whereas ANN showed less area in the very high and high susceptible areas. In contrast, LSTM showed more areas under the very high and high susceptibility zone (Figure 17). Most of the vulnerable zones were in the lower portion of the basin (areas denoted as 'very high' in Figures 17 and 18).

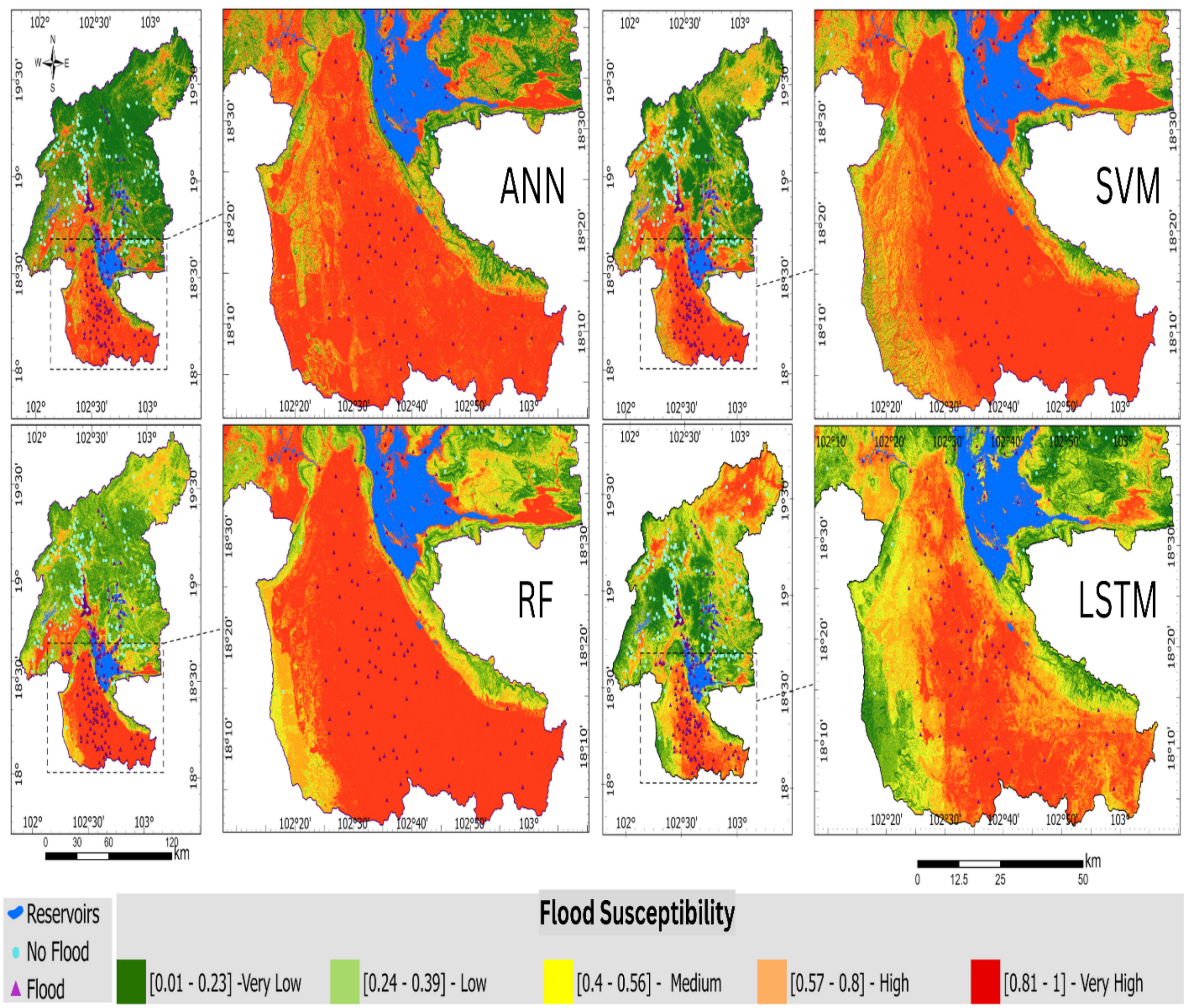


Figure 17. Flood susceptibility maps generated from machine learning models

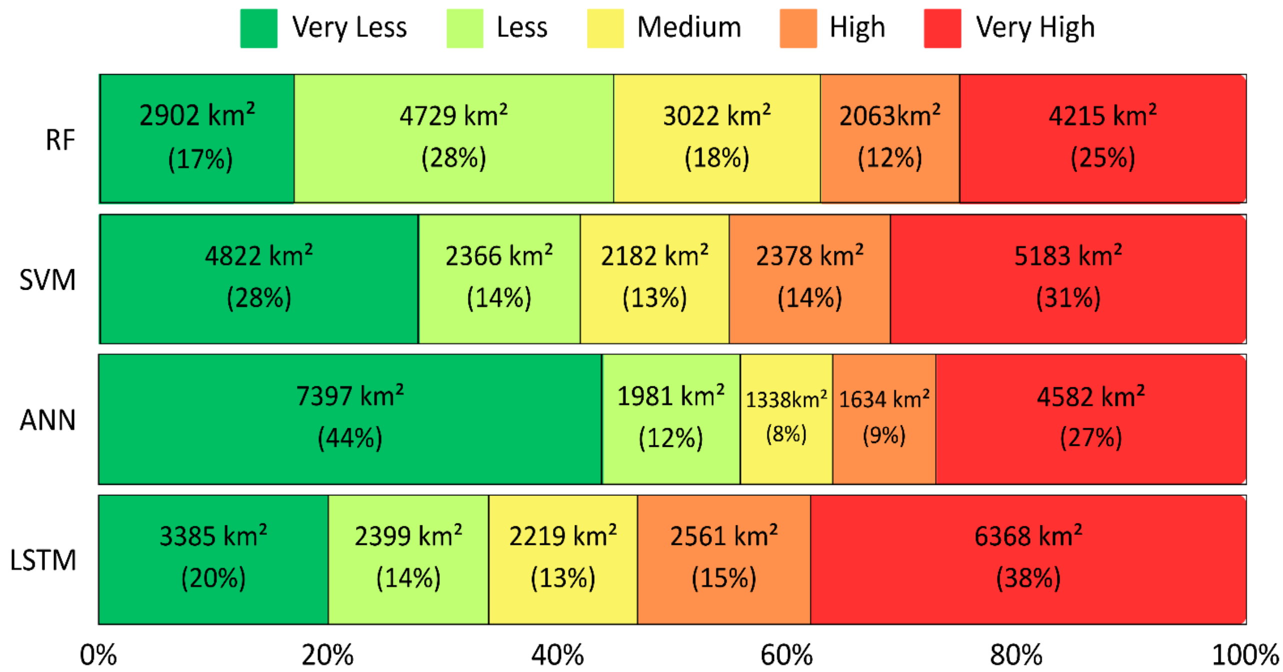


Figure 18. Flood susceptible areas in percentage and sq. km

4.6 Evaluation of model and validation

The machine learning models in this study were developed and validated using training and test datasets. Previous studies (Bera, Das, and Mazumder 2022; Bhattarai et al. 2024; Duwal et al. 2023; Fang et al. 2021; Y. Wang et al. 2020; Wubalem et al. 2021) have suggested AUROC for model validation. The AUROC value indicates the performance of the model in detecting flood-prone areas. A higher AUROC value suggests better model performance. Sensitivity measures how accurately the model predicts positive instances, while specificity indicates the accuracy of predicting negative instances. The results revealed that all the models performed with higher precision with values >0.90. Based on AUROC, the RF performed the best (Figure 19), evidenced by the train and test AUROC 1.00 and 0.993, respectively, followed by SVM (0.996 and 0.989), ANN (0.991 and 0.977), and LSTM (0.97 and 0.983). F1-score, precision, recall, kappa, MSE, and RMSE were employed for detailed evaluation accuracy. The F1-Score offers an equilibrium between precision and recall, where a high F1-Score reflects effective identification of flood-prone regions while minimizing false positives (Zhang et al. 2022). A high level of sensitivity ensures that most flood-prone areas are correctly identified (Chapi et al. 2017). RF exhibits superior performance for other parameters also; with the highest Sensitivity (0.969), Accuracy (0.957), F1 Score (0.962), and Recall (0.969), indicating its robustness in correctly classifying flood-prone areas. In the case of accuracy, RF and SVM have the highest value (0.957) and Kappa scores (0.913

and 0.914, respectively), suggesting a strong agreement between the predicted and observed classifications. Specificity (True Negative Rate) is essential to avoid false alarms in non-flood-prone areas (Pourghasemi et al. 2020). The SVM model stands out with the highest specificity (0.962) and precision (0.969), highlighting its capability to correctly predict negative cases, i.e. non-flood points, and the proportion of correct positive predictions, i.e. flood points. LSTM lags in all metrics, with notably lower Accuracy (0.915) and higher MSE (0.292) and RMSE (0.085). Overall, RF emerges as the most reliable model for this application, effectively balancing accuracy and error metrics (Figures 19 and 20). However, it should be noted that all the models performed well, and their results should be considered for flood susceptibility mapping.

Table 5 The optimum values of the tuning parameters of different models

Method	Hyperparameter	Optimum Value
RF	'bootstrap'	True
	'max_depth'	30
	'max_features'	2
	'min_samples_leaf'	4
	'min_samples_split'	10
	'n_estimators'	200
	'oob_score'	True
SVM	'C'	[0.1,1,100,1000]
	'degree'	[1,2,3,4,5,6]
	'kernel'	['rbf','poly','sigmoid','linear']
ANN	'batch_size'	[8, 32, 64]
	'epochs'	[5,10]
	'Optimizer_Trial'	['adam', 'rmsprop']
	'Neurons_Trial'	[10,15]
	Hidden layer unit	[1]
DNN	'batch_size'	[64]
	'epochs'	[500]
	'Optimizer_Trial'	['rmsprop']
	'Neurons_Trial'	[64]
	Hidden layer unit	[3]
	Activation	‘Sigmoid”
	Function	‘relu’
LSTM	'batch_size'	[10, 32, 64]
	'epochs'	[10, 20]
	'Optimizer_Trial'	['rmsprop']
	'Neurons_Trial'	[40]
	Activation	‘Sigmoid”

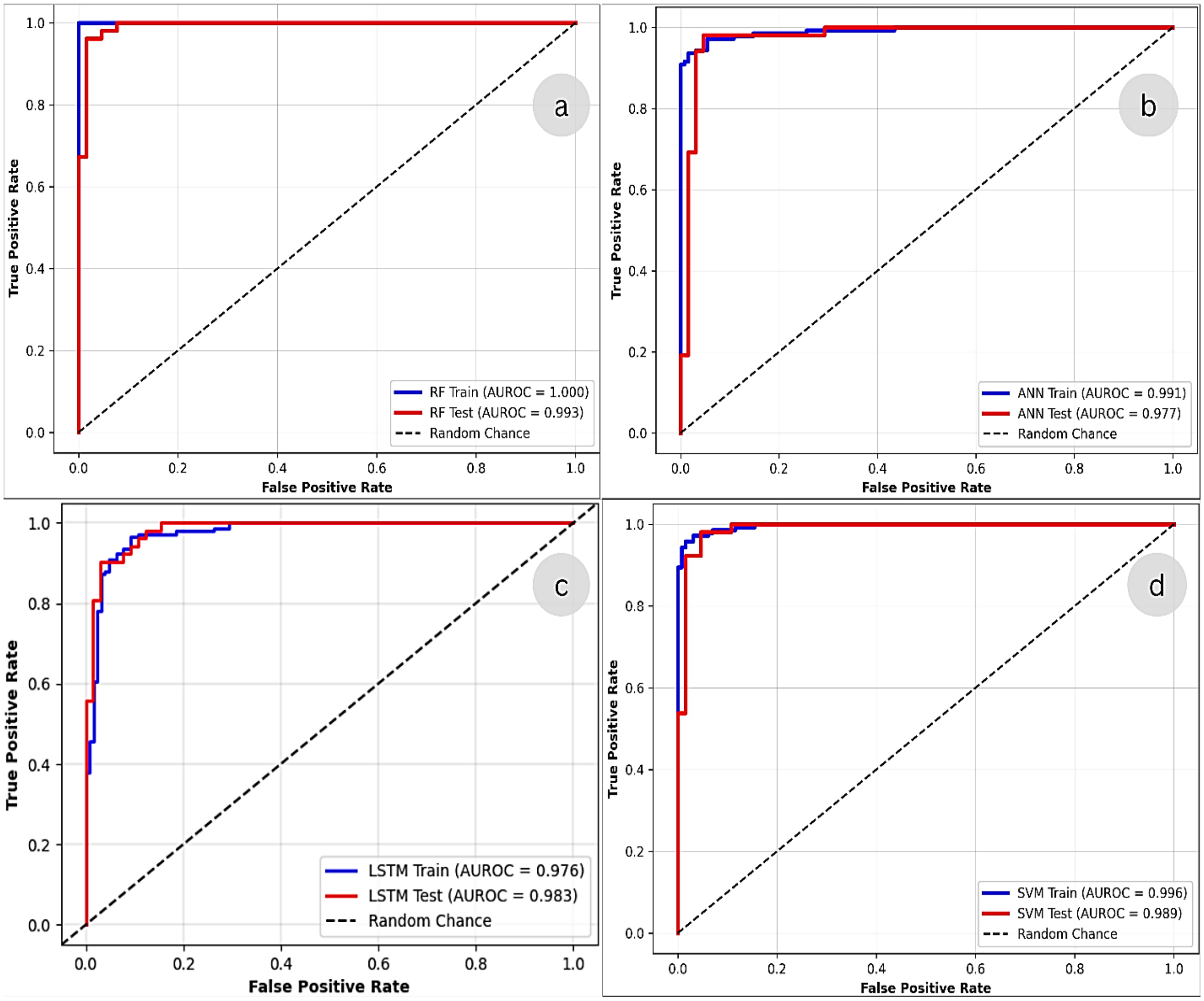


Figure 19. AUROCs For all models, (a) random forest, for artificial neural network, (c) long short-term memory, and (d) support vector machine

- The results revealed that all the models performed with higher precision with values >0.90.
- Based on AUROC, the RF performed the best (evidenced by the train and test AUROC 1.00 and 0.993, respectively, followed by SVM (0.996 and 0.989), ANN (0.991 and 0.977), and LSTM (0.97 and 0.983).

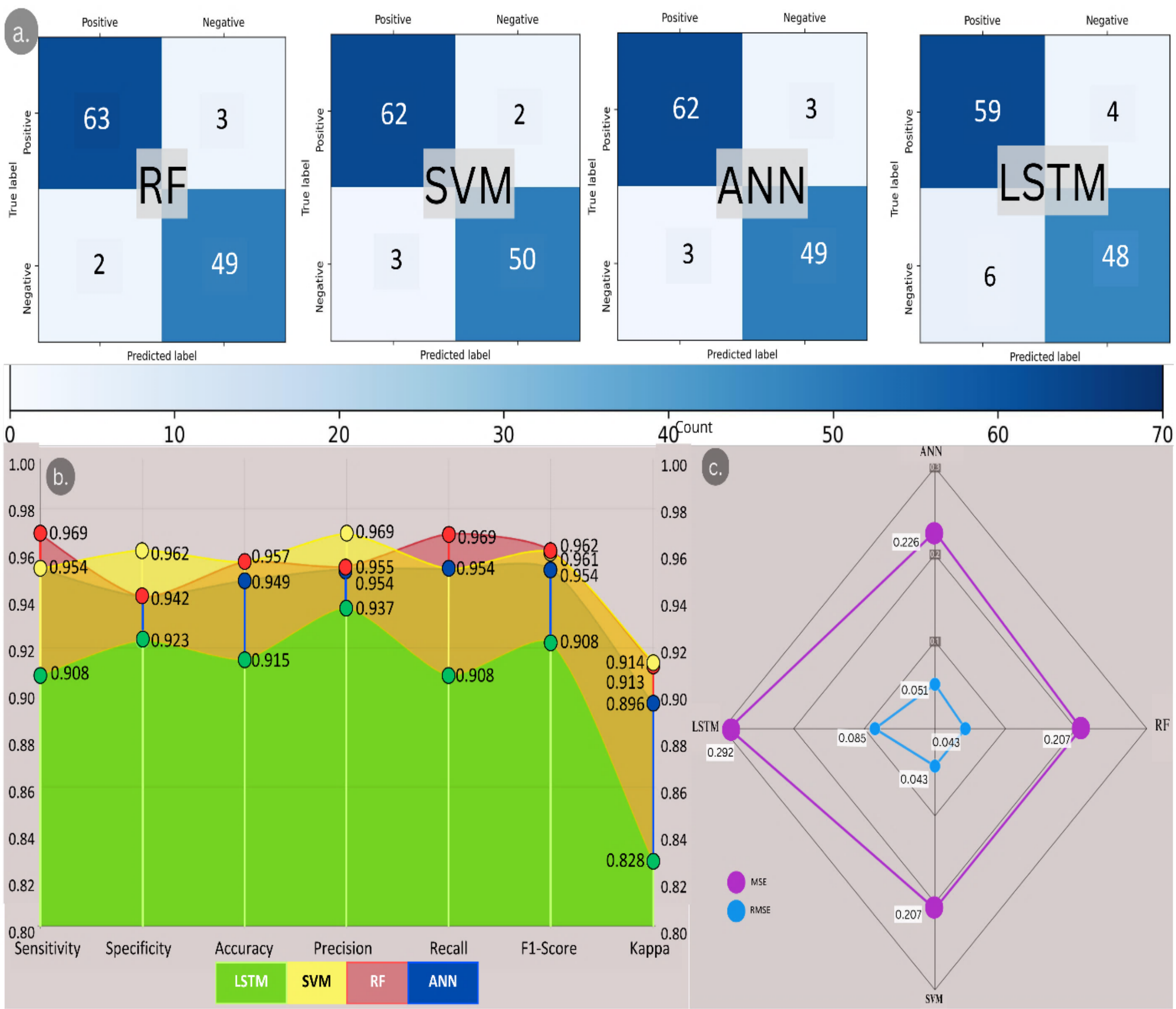


Figure 20. Performance of the models: (a) confusion matrix, (b) precision and accuracy assessment parameters, and (c) MSE and RMSE

- RF exhibits superior performance for other parameters with the highest Sensitivity (0.969), Accuracy (0.957), F1 Score (0.962), and Recall (0.969), indicating its robustness in correctly classifying flood-prone areas.
- SVM has the highest specificity (0.962), precision value (0.96), and Kappa scores (0.914), respectively
- LSTM lags in all metrics, with notably lower Accuracy (0.915) and higher MSE (0.292) and RMSE (0.085)

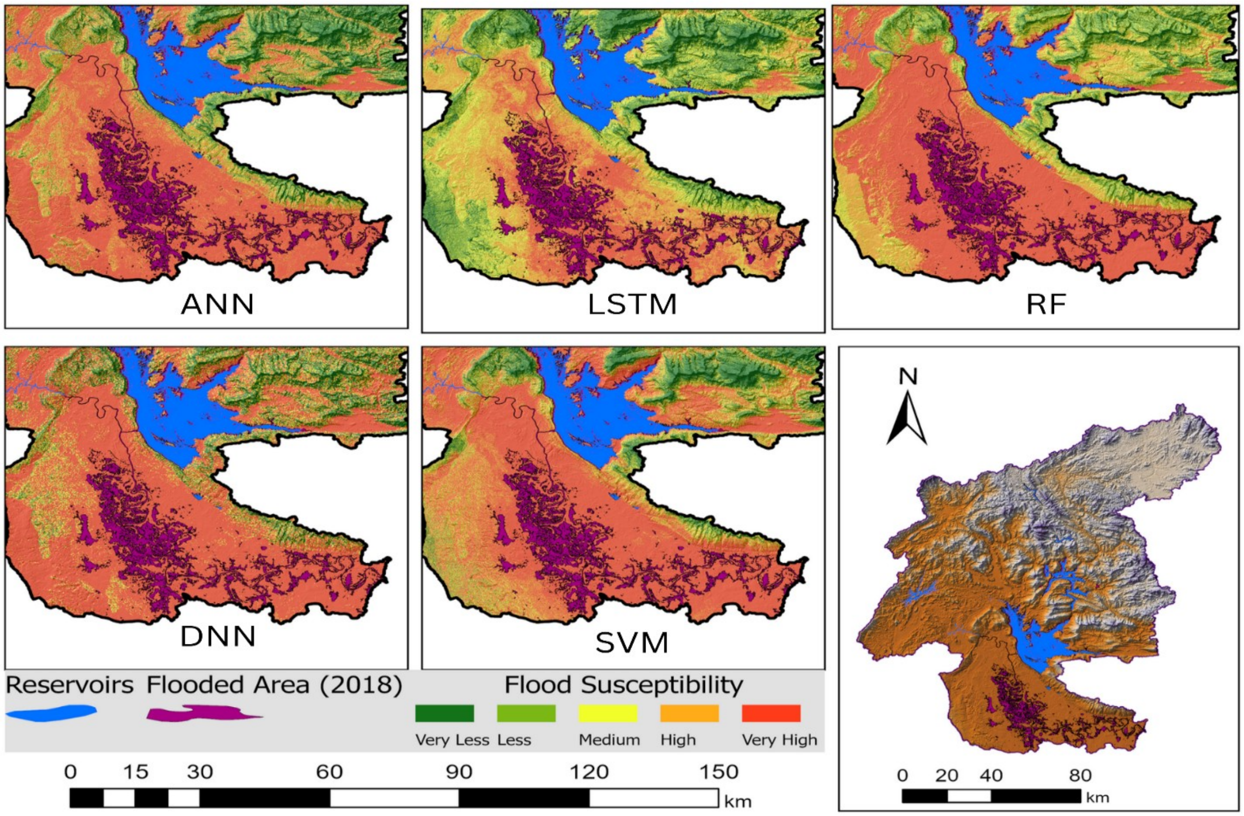


Figure 21 validation of ground truth data based on K4D map from SAR image 2018

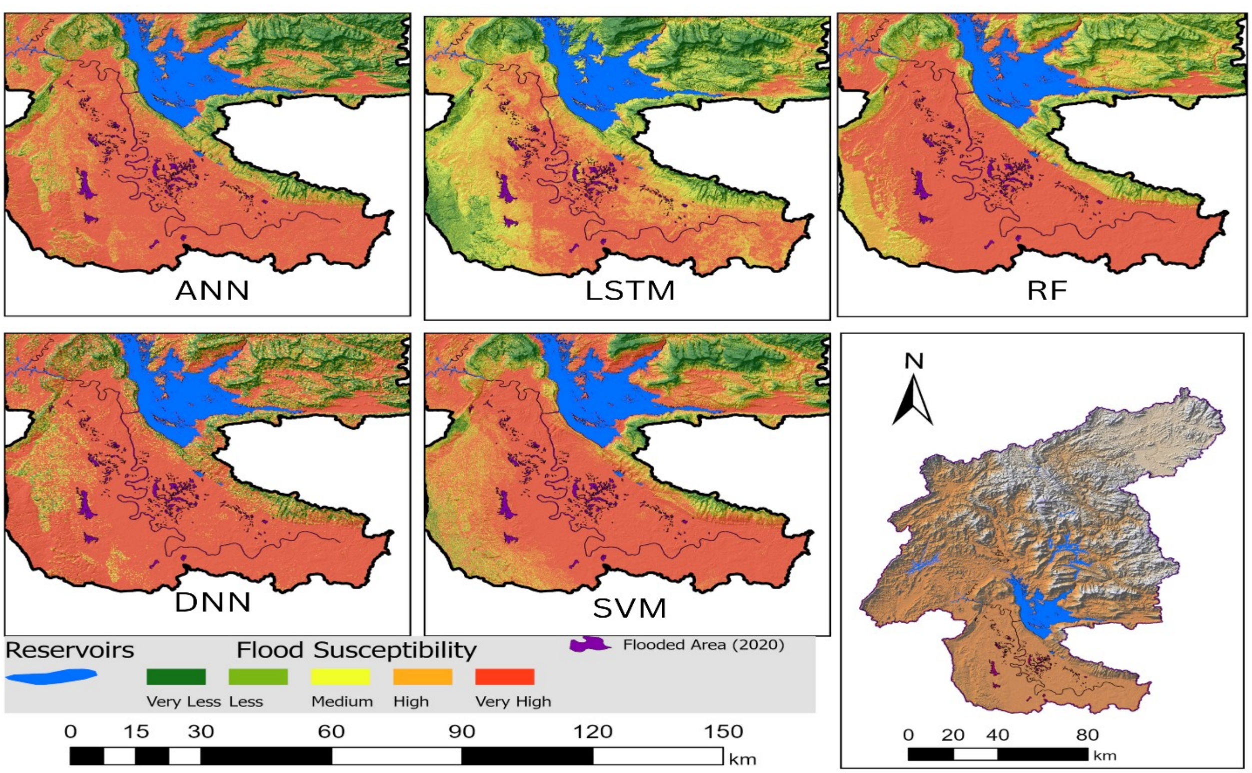
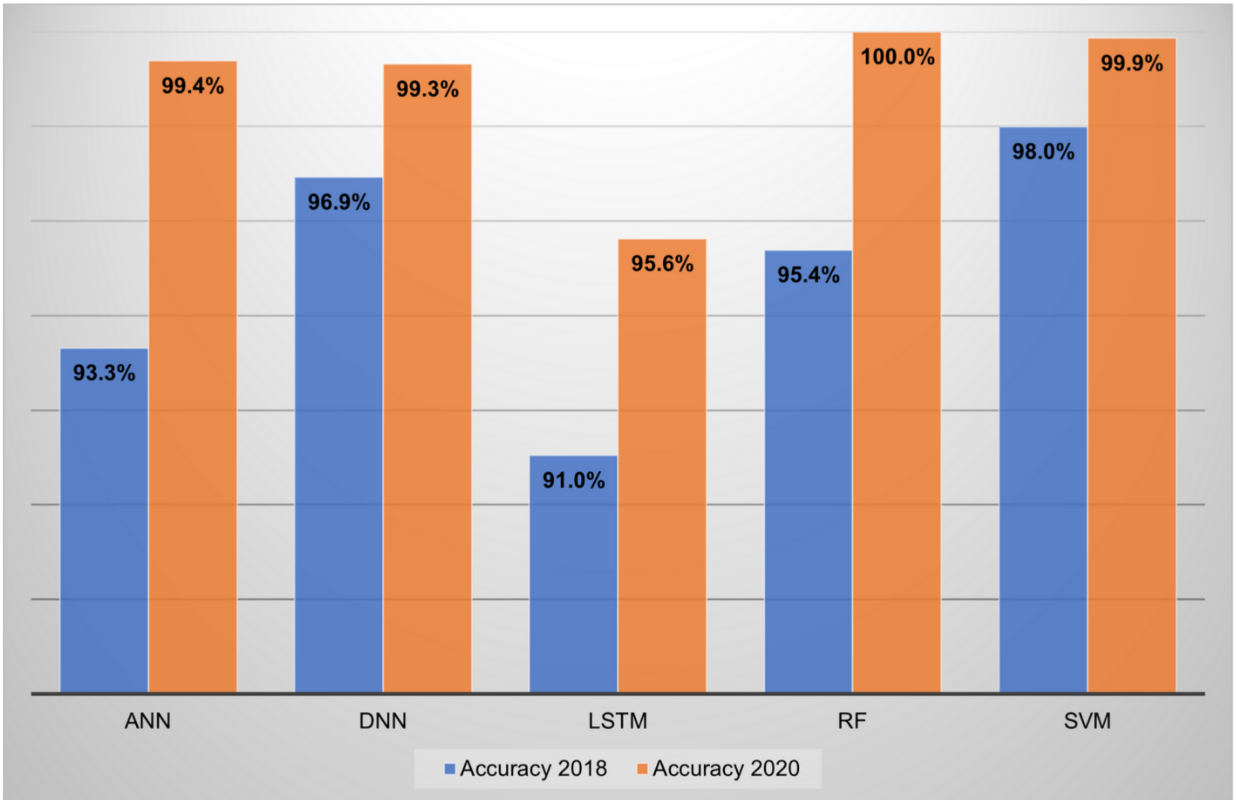


Figure 22 validation of ground truth data based on K4D map from SAR image 2020



	ANN	DNN	LSTM	RF	SVM		ANN	DNN	LSTM	RF	SVM
Very Less	39	22	19	7	6		0	3	0	0	0
Less	28	16	47	23	10		3	3	10	0	0
Medium	33	8	68	39	14		6	4	54	0	2
High	61	20	213	83	46		26	6	321	13	11
Very High	1335	1430	1149	1344	1420		1428	1447	1078	1450	1450

Figure 23 Accuracy assessment data for flood area in 2018 and 2020

The sampling was random, with each point separated by a minimum distance of 100 meters. A total of 1,500 random points were selected from flooded area data in ArcGIS for each year, and their corresponding susceptibility values were taken from the flood susceptibility map shown above in Figure 23. If a point lies in high or very high susceptibility areas, it is considered a true location; otherwise, it is regarded as false. The accuracy assessment data for 2018 and 2020 is over 90%, indicating that the ground data and advanced machine learning achieved high accuracy in identifying high and very high flood susceptibility zones in low-lying areas of the Nam Ngum River Basin (NNRB), Lao PDR.

CHAPTER 5

FLOOD RISK ASSESSMENT ON LAND COVER, POPULATION, AND CRITICAL FACILITIES IN NAM NGUM RIVER BASIN, LAO PDR

5.1 Overview

This chapter provides a comprehensive examination of flood risk assessment on land cover, Population, and Critical Facilities within the Nam Ngum River Basin, located in Lao People's Democratic Republic PDR. The objectives and methodology employed deep learning in conducting the flood risk assessment, detailing the data sources, modeling techniques, and analysis. In addition, it will mention the effect of flooding on different LULC Classes and Populations in NNRB, using eleven flood-influencing factors and three DL algorithms: ANN, DNN, and LSTM.

5.2 Methodology

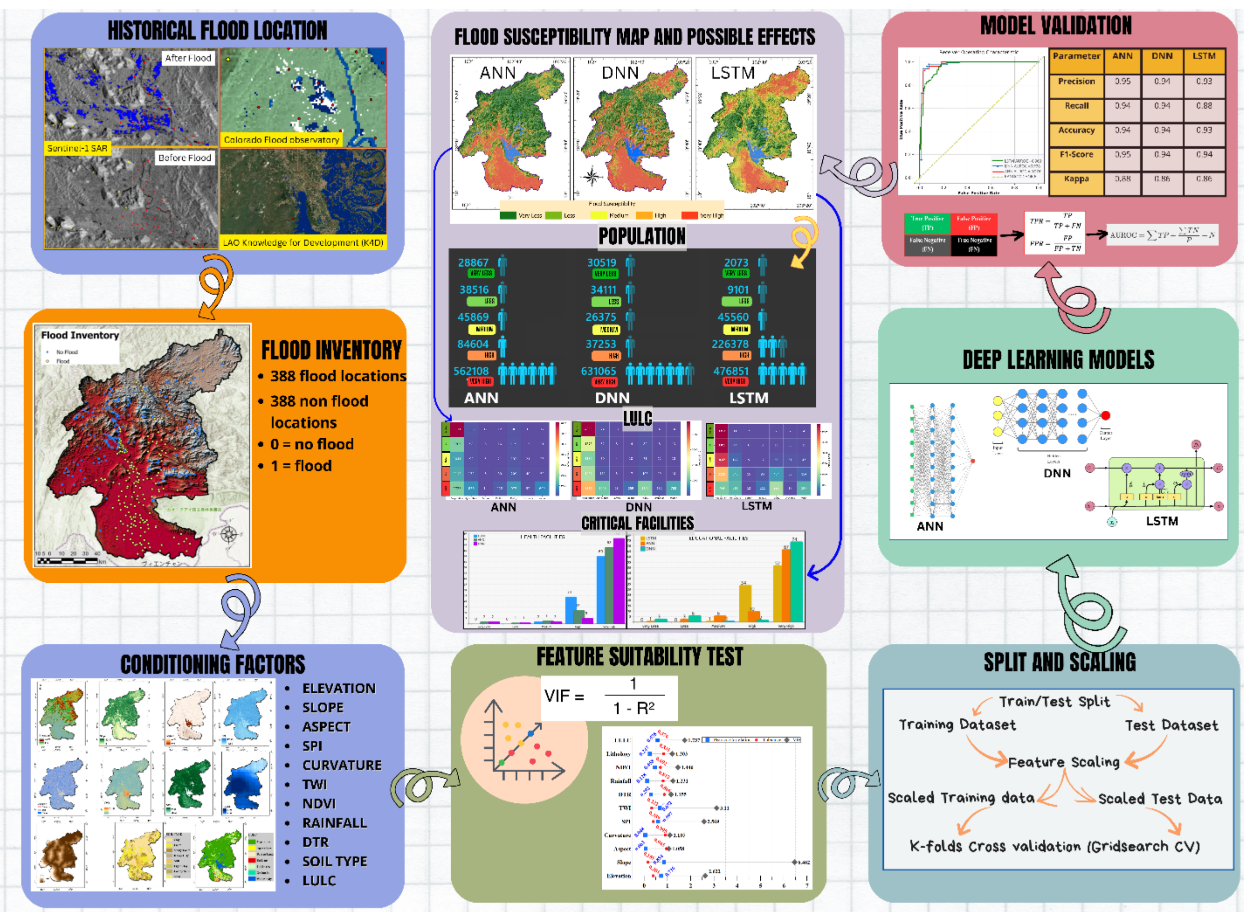


Figure 24. Schematic diagram of the flood susceptibility mapping in Nam Ngum River Basin

The study was initiated by collecting geospatial data from various sources, including ALOS-PALSAR DEM for factors like slope, aspect, elevation, curvature, TWI, and SPI, as well as rainfall data from ERA5 and NDVI from Landsat 8 images. Land use/land cover (LULC) data were obtained from Sentinel-2 images, and flood locations were identified using Sentinel-1 SAR images, the Colorado Flood Observatory (https://floodobservatory.colorado.edu/), and the Knowledge for Development portal(https://en.data.k4d.la/).

The flood inventory map was generated using the flood location data prepared in google earth engine https: //code.earthengine.google.com/). A total of 776 past flood locations were selected, along with an equal number of non-flood locations, for model training and testing. Eleven flood conditioning factors were selected based on existing literature, including elevation, slope, aspect, curvature, Topographic Wetness Index (TWI), Stream Power Index (SPI), distance to river (DTR), Normalized Difference Vegetation Index (NDVI), rainfall, lithology, and land use/land cover (LULC). Deep learning methods, including Deep Neural Network (ANN), Artificial Neural Network (ANN), and Long Short-Term Memory (LSTM), were employed for flood susceptibility mapping. ANN models replicate the brain's interconnected neurons, processing sensory inputs through layers of artificial nodes. Each connection between layers carries weighted information, influencing the final output. Researchers value ANNs for their nonlinear modeling abilities and adaptability to complex systems. LSTM is a type of recurrent neural network (RNN) designed to handle the issue of vanishing gradients, enabling it to effectively capture long-term dependencies in sequential data by selectively remembering or forgetting information over time. DNN is an artificial neural network with multiple layers between the input and output layers. These networks consist of an input layer, several hidden layers, and an output layer. Each layer comprises nodes (neurons) that process and transform the input data. The 'deep' aspect refers to the network's depth, i.e., the number of hidden layers, which enables it to model complex patterns and relationships in data. Multicollinearity among the conditioning factors was assessed using the variance inflation factor (VIF) and tolerance values. Factors with VIF values above 10 or tolerance values below 0.10 were considered multicollinearity issues and were not used for the analysis. Information gain ratio (IGR) values were used for the feature selection and evaluate the impact of the conditioning factor on the susceptibility modeling. Evaluation metrics such as area under the curve for Receiver Operating Characteristics (AUROC), Kappa Score, Accuracy, Sensitivity, and Specificity were used to assess the performance of the machine learning models. These metrics help determine the model's ability to correctly classify flood and non-flood pixels and provide insights into its overall performance. The validated model is then used to look at the scenario of effects on different land use classes and the number of people vulnerable to flooding in NNRB using LULC data, population density data for year 2020, Health and Educational Facilities data from HUMANITARIAN DATA EXCHANGE v1.76.0 PY3 (https://data.humdata. org) is shown in (28).

5.3 Feature Selection and Evaluation

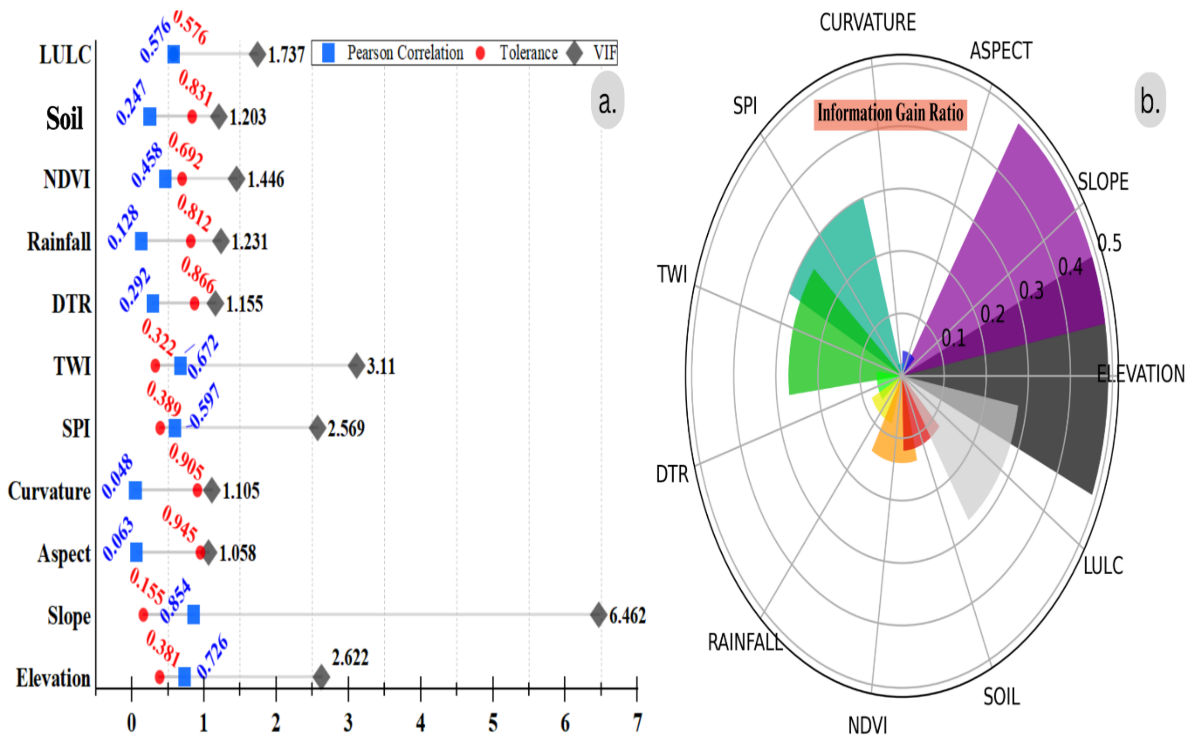
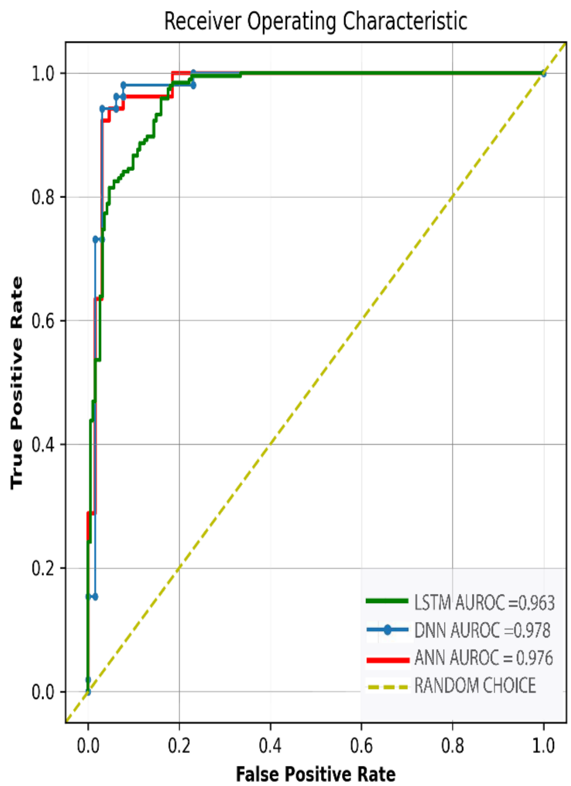


Figure 25. Feature Selection and Evaluation using a) Variance Inflation Factors, Tolerance & Pearson Correlation and b) information gain ratio

Information gain ratio (IGR) is one of the most popular techniques for assessing important variables influencing flooding and enhancing the models' performance and prediction accuracy (Al-Abadi 2018; Saber et al. 2023). The information gain ratio analysis depicts significant insights into flood conditioning factors in the Nam Ngum River Basin. Slope and elevation with an IGR value of 0.49 emerged as highly influential, while SPI (0.30), LULC (0.28), TWI (0.27), NDVI (0.14) and Soil (0.12) also exhibited notable impact on flood prediction. Factors like Rainfall (0.08), DTR (0.06), and curvature (0.04), aspect (0.02) being the least influential factors (Figure 25). Similarly, multicollinearity tests show that slope has the highest VIF (6.462), followed by TWI (3.11). Notably, no significant collinearity was observed among the selected flood conditioning factors. Moreover, Pearson's correlation coefficients clarify the relationship between flood occurrence and conditioning factors. Slope and elevation emerged as pivotal, followed by TWI, SPI, Landcover, NDVI, DTR, Lithology, Rainfall, Aspect, and Curvature, respectively.

5.4 Evaluation of Model Parameters



Parameter	ANN	DNN	LSTM
Precision	0.95	0.94	0.93
Recall	0.94	0.94	0.88
Accuracy	0.94	0.94	0.93
F1-Score	0.95	0.94	0.94
Kappa	0.88	0.86	0.86

Figure 26. validation of prediction rate and success rate curves for (a) ANN, (b) DNN, and (c) LSTM

Performance appraisal values for flood prediction were generated using the training dataset to assess each of the three machine models. Area Under Curve for Receiver Operating Characteristics (AUROC), a critical evaluation criterion in model performance, was utilized to evaluate model performance. AUROC serves as a probability curve, where the AUROC value indicates the degree of separability. AUROC values for the DNN model stood out, with the highest value of 0.978, showing its excellent performance in predicting flood-prone areas. The ANN and LSTM models were followed by AUROC of 0.976 and 0.963 (Figure 26). However, ANN shows the best overall performance with the highest precision, recall, accuracy, F1-score, and Kappa. It also has a slightly lower AUROC than DNN but is better. DNN performs very closely to ANN with slight variations but has the highest AUROC, indicating its excellent ability to distinguish between classes. LSTM has the lowest performance among the three models, with lower true positives and recall, which affects its overall accuracy and Kappa. However, it still maintains a good AUROC. In summary, ANN is the most balanced and high-performing model, followed closely by DNN. LSTM, while still effective, falls slightly behind in several key performance metrics.

5.5 Flood Susceptibility Mapping

The results of the three deep learning models are presented in (Figure 27). The study revealed that for the ANN model, 27% (4582 km^2) of the area lies in a very high flood susceptibility zone, followed by 9% (1634 km^2) in highly susceptible, 8% (1338 km^2) in medium, 12% (1981km^2) in less and 44% (7397 km^2) in very less flood susceptible zones. Similarly, DNN shows that 29% (4832 km^2) of the area lies under very high susceptibility area, 19% (3189) in high, 11% (1919 km^2) in medium, 14% (2395 km^2) in less, and 27% (4596 km^2) in very less flood susceptible areas. Lastly, LSTM shows that 38% (6369 km^2) lies in very high, 15% (2561 km^2) in high, 13% (2219 km^2) in medium, 14% (2399 km^2) in less, and 20% (3385 km^2) in very less susceptible areas. It can be observed that the variations in parameter values and flood susceptible areas are less for ANN and DNN compared to LSTM. In this scenario, it can be concluded that the results of ANN and DNN are more reliable than those of LSTM. However, the result from LSTM should also be considered for flood susceptibility mapping since the parameters are in the higher range. In this scenario, it should be noted that 36% to 56% of the area of NNRB is highly susceptible to flood (Figure 27). Overall, while there are variations in the identification of flood susceptibility areas among the models, they are similar in the extent of high and Very High susceptibility areas, highlighting the importance of flood management and mitigation efforts in NNRB. This alarming condition depicts a need for serious concern in flood disaster prevention and management strategies

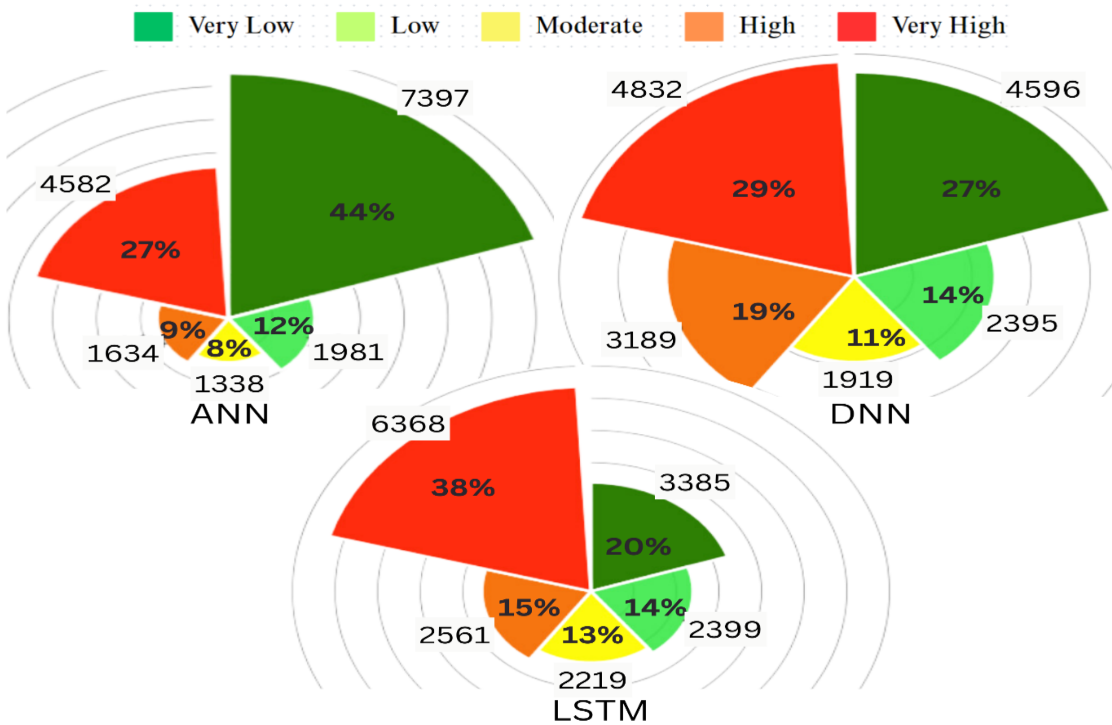


Figure 27 Flood susceptible area (in km^2) comparison according to the models

5.6 Probable effect of flooding on different LULC Classes and Population

5.6.1 Effects on the population

The analysis of flood susceptibility in the population using three deep learning methods (ANN, DNN, and LSTM) reveals significant variations in the number of people affected across different susceptibility levels. For the ”Very High” susceptibility category, the DNN method indicates the highest impact, with 631,065 individuals affected, followed by ANN with 562,108, and LSTM with 476,851. In the ”High” category, LSTM reports a notably high number of 226,378 affected individuals, compared to 84,604 for ANN and 37,253 for DNN. The ”Medium” susceptibility shows similar impacts for ANN (45,869) and LSTM (45,560), both higher than DNN (26,375). The ”Less” susceptibility category affects 38,516 individuals for ANN, 34,111 for DNN, and significantly fewer for LSTM (9,101). Finally, the ”Very Less” category sees DNN and ANN affecting 30,519 and 28,867 individuals, respectively, while LSTM affects only 2,073. The increase in the susceptibility and the increase in the number of people vulnerable to flood are in line (Figure 28). The analysis shows that the number of people susceptible to flooding ranges from around 6.5 million to 7 million, which is 85% to 93% of the total population

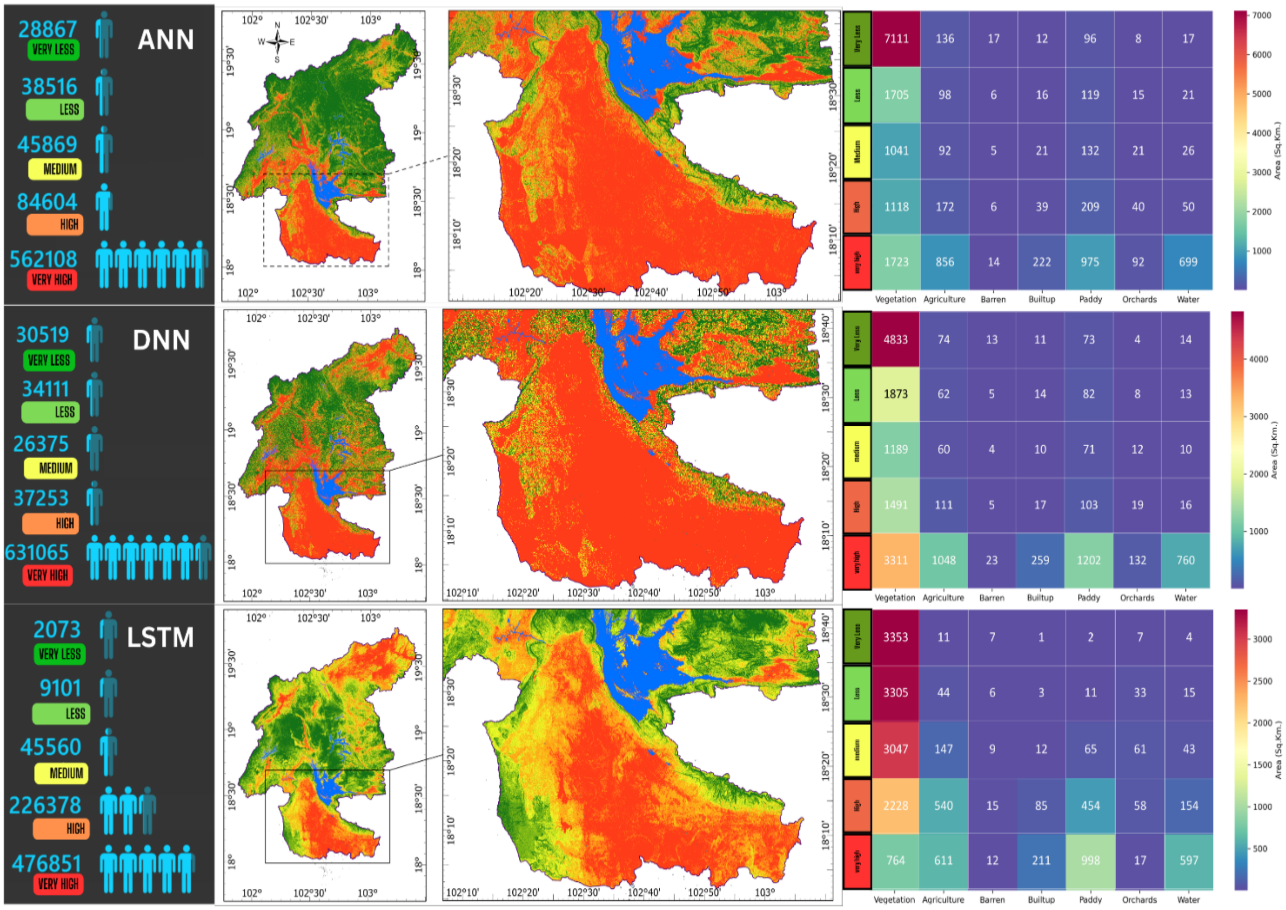


Figure 28. Scenarios of the effect of flood susceptibility on Population and Land use / Land Cover

5.6.2 Effects on LULC

According to Figure (28), flooding has serious consequences to the anthropogenic activities in NNRB. The very high and highly susceptible areas are the agriculture, paddy, orchards, and built-up areas, clearly showing that flooding has a high probability of affecting the activities of people there. There is a probability that 260 to 280 km^2 of built-up, 1050 to 1120 km^2 of agriculture area, 1180 to 1300 km^2 of paddy, 130 to 140 km^2 of orchards, and 2700 to 2800 km^2 of vegetation are highly susceptible to flooding (Figure 28). The majority of the vegetated area lies in the very less and less susceptible areas. This result signifies the importance of vegetation for flood control.

According to this study, in NNRB, the majority of the built-up areas or the highly populated areas lie in the very high and high flood susceptible areas. The densely populated areas lying at the downstream of NNRB with major settlements are highly vulnerable to flooding. Similarly, the populated built-up area at the upstream side of NNRB are also vulnerable to flooding. For example, the Naxaithong, Xaythany, and Thaulakhom districts (Figure 29 a-1 & b-1), The area upstream of the Nam Ngum Reservoir in Vangveing District (Figure 29a-2 & b-2), and the populated area in the upstream area at Pek District (Figure 29a-3 & b-3).

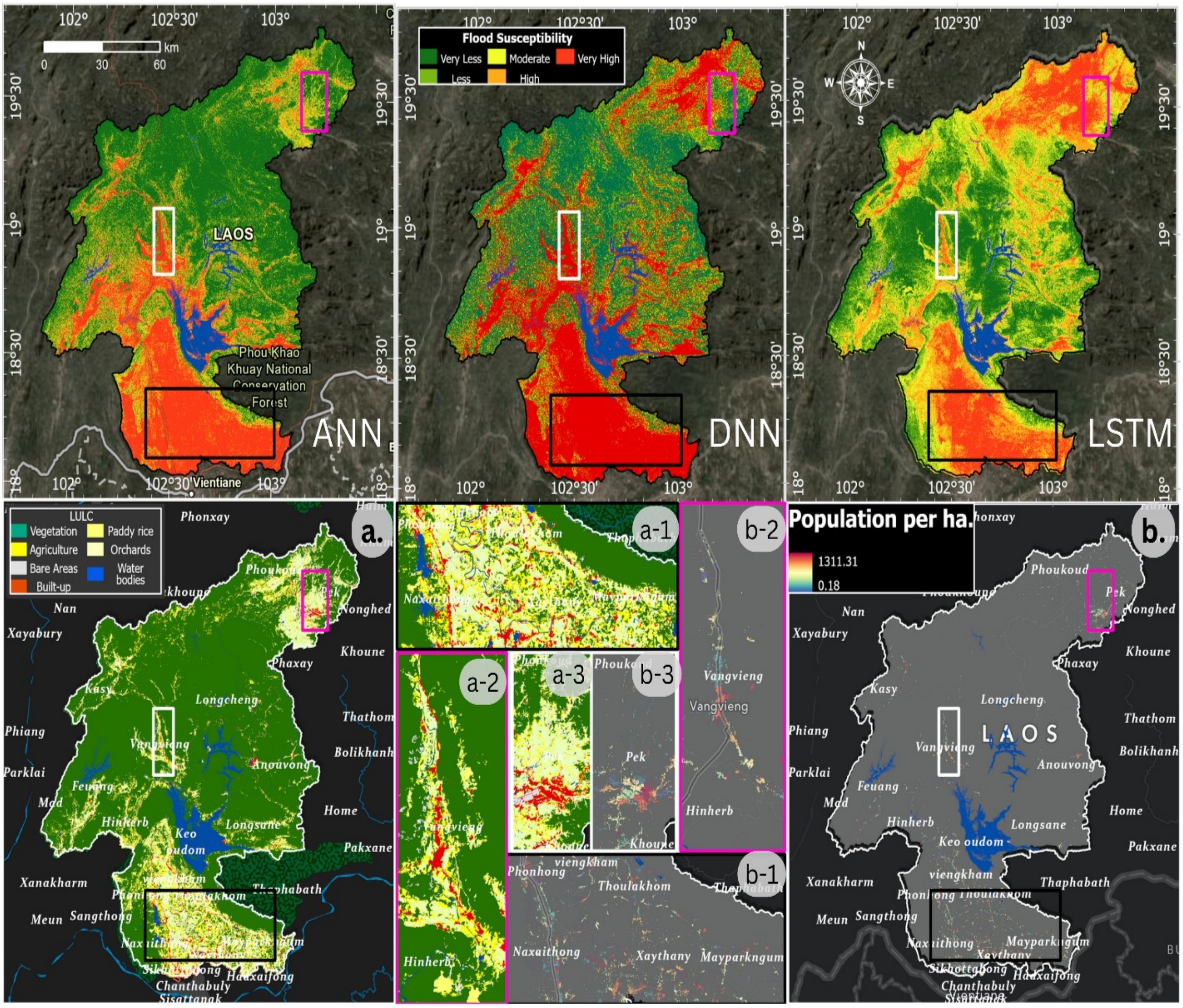


Figure 29. Flood Susceptibility to a. Land use/Land cover, and b. Population (a-1. & b-1. Downstream of NNRB (Naxaithong, Xaythany and Thaulakhom area), a-2. & b-2. Vangvieng Area, a-3 & b-3 Upstream of NNRB (Pek Area))

5.3 Effects on Critical Facilities: Health and Educational Facilities

For health facilities in (Figure 30), DNN, ANN, and LSTM predict 76, 68, and 60 facilities lying in ”Very High” susceptible areas. Similarly, LSTM predicts 24, followed by ANN with 12, and DNN with five facilities in highly susceptible areas. For educational facilities, DNN predicts 74, ANN predicts 67, and LSTM predicts 52, lying in very highly susceptible areas. In the highly susceptible areas, LSTM predicts the most affected facilities with 34, followed by ANN with 10, and DNN with 2 in (Figure 30). This data highlights the critical areas that need attention for flood risk management and mitigation, particularly in the Very High and High susceptibility categories, where the numbers of affected health and educational facilities are significantly higher for all the models.

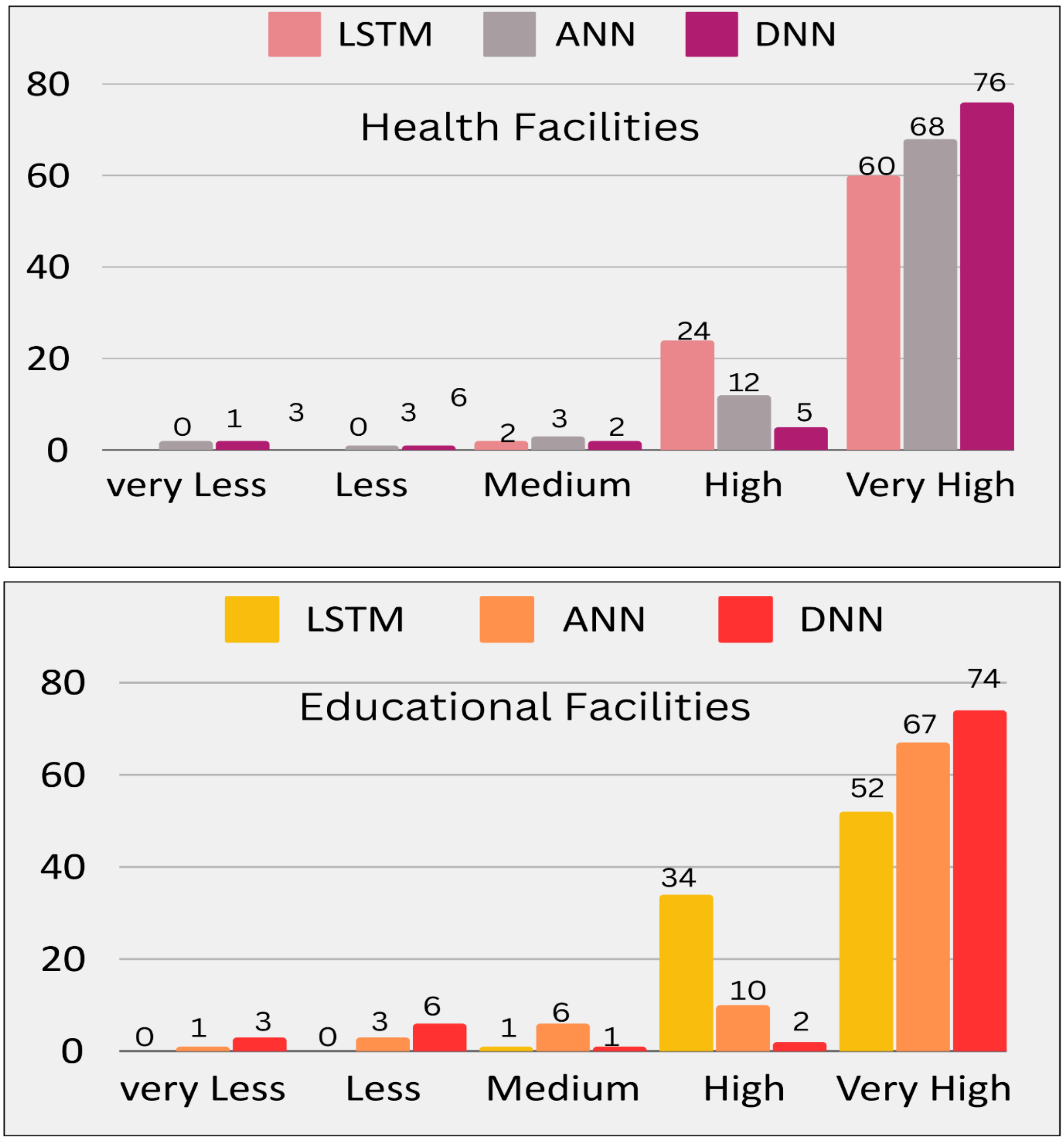


Figure 30. Possible Effects of Flood on the Health and Educational Facilities in NNRB

CHAPTER 6

CONCLUSION AND RECOMMENDATION

6.1 Introduction

In general, this research used open-source remote-sensing data and machine learning approaches to perform flood detection, modeling and optimization of conditioning factors. The study areas were utilized in this research which is Nam Ngum River Basin (NNRB), Lao PDR. Key concerns for governments and planners include ensuring safety and maintaining sustainable management of a region. Achieving this objective requires thorough analysis to facilitate appropriate actions during hazard events to minimize damage. This study introduces a relatively precise and comprehensible flood prediction method for mapping flood-prone areas. Random Forest, Support Vector Machine, Artificial Neural Networks, Deep Neural Networks and Long Short-Term Memory were applied to test its efficiency in flood susceptibility mapping.

The most critical flood conditioning factors were identified through scientific analysis, along with the detection of non-significant factors. The models developed and implemented in this study can be adapted for use in different regions to identify flood-prone locations, aiding in the mitigation of future flooding events. The analysis generated several figures and tables to facilitate comprehension of the procedure and methodology. Validation affirmed the effectiveness and reliability of the methods introduced. Additionally, this research highlighted the efficacy of open-source remote sensing (RS) and geographic information systems (GIS) in hydrological studies. Specifically, GIS and modeling methods can be employed to examine the potential impacts of land use and land cover (LULC) changes on flood generation. Our study has identified critical factors influencing flood occurrences in NNRB. This data is essential for developing precise flood forecasting models, implementing advanced flood warning systems, and guiding infrastructure and urban planning projects. According to the results, policymakers could concisely reconsider disaster risk reduction and flood management strategies. Land-use planners could determine the settlement zones, dams, and other structures in highly susceptible areas. Furthermore, the results could be utilized to inform the residents about evacuation actions, flood prevention, and preparedness.

6.2 Summary

Flood susceptibility maps are crucial in floodplain management relating to disasters, and integrating various models offers critical insights for flood risk management. This study used four machine-learning models (SVM, RF, ANN, DNN and LSTM) to predict flood susceptibility in NNRB, an area annually affected by typhoons and heavy rainfall. The flooded areas were delineated using Sentinel 1 SAR imagery in the GEE platform by applying the threshold for the pre-and post-flood scenarios. The conditioning factors were quantified using the information gain ratio (IGR), Pearson correlation coefficient, and multicollinearity test. The most influential factors affecting floods were identified as topographic factors, particularly elevation, slope, and stream power index, which are major drivers of flooding in NNRB. The models were trained and tested using eleven flood conditioning variables and 390 locations. The RF model outperformed the others, as evidenced by its superior AUROC precision, F1-score, accuracy, and kappa values. Additional metrics, including MSE and RMSE, were employed for a comprehensive evaluation. This indicates that 36% to 53% of NNRB, particularly in the downstream southern region, is highly vulnerable to flooding. Despite constructing numerous reservoirs for hydropower and flood protection, downstream areas continue to flood yearly (Kimmany, Ruangrassamee, and Visessri 2020) challenging flood management (Keophila, Promwungkwa, and Ngamsanroaj 2019; Vilandone Keophila 2018). Effective floodplain management is needed to protect lives and property. The study's results offer valuable insights for flood risk assessment and developing effective flood control plans. As floods threaten infrastructure, agriculture, and the economy worldwide, these machine-learning insights using remote sensing data can aid local authorities, planners, policymakers, and stakeholders in disaster management and climate change mitigation for sustainable development.

6.3 Limitations

This research can be improved further by considering the following recommendations:

- Although remote sensing (RS) and geographic information systems (GIS) hold significant potential for flood studies, their use has been somewhat limited. One major issue is the presence of cloud cover during flood events, which hinders the effectiveness of optical sensors. Synthetic Aperture Radar (SAR) data, with its radar pulses that have greater penetration power, can overcome the cloud cover problem.
- Additional challenges in utilizing remote sensing (RS) include limited availability of images across different timeframes and locations, seasonal variations, technical limitations, and insufficient temporal resolution. In terms of temporal resolution, radar images are typically taken either before or after a flood event, often failing to capture the peak of the

flooding event.

- Further challenges in employing remote sensing (RS) and geographic information systems (GIS) in flood studies include expensive digitization and data gathering, predictive method complexities, inadequate spatial data availability, hardware limitations for managing extensive datasets, and GIS constraints in handling historical data essential for certain natural hazard assessments.
- The type of underlying ground can influence the changes observed in SAR imagery of flooded areas. Flooded barren land shows lower backscattering compared to adjacent non-flooded regions because a smooth water surface provides specular reflection. Wind can increase backscattering from flooded areas, thereby reducing the contrast between flooded and non-flooded regions. Flooded forests create a robust radar response because of a mechanism where the signal reflects twice between the water surface and the tree trunks.

6.4 Recommendation for Future Work

1. Even though flood susceptibility mapping has been trialed across different geographic areas, further in-depth research is necessary to enhance and optimize the models. The methods suggested in this study should be utilized in different regions to collect additional data on their efficacy. Implementing these methods in different study areas will be valuable for assessing their performance.
2. Incorporating additional datasets could enhance the effectiveness of the proposed change detection method for identifying flood-prone areas
3. Developing an automatic flood detection technique is highly recommended due to the limited time available for mapping flooding in an area. Real-time analysis is typically necessary in hazard studies. Therefore, future research should focus on detecting flood locations in real-time to expedite susceptibility analysis
4. Future research works can further explore a hybrid combining hydrodynamic-machine learning-based modeling to predict flood susceptible area, velocity, and depth of flooding, enabling more comprehensive modeling of flood vulnerability in Nam Ngum River Basin (NNRB), Lao PDR

References

Abedi, Rahebeh, Romulus Costache, Hossein Shafizadeh-Moghadam, and Quoc Bao Pham. 2022. "Flash-Flood Susceptibility Mapping Based on XGBoost, Random Forest and Boosted Regression Trees." *Geocarto International* 37(19):5479–96. doi: 10.1080/10106049.2021.1920636.

Adams, Geoff, Rachel S. Blakers, and Paradis Someth. n.d. "Rapid Initial Assessment of the State of Water Resources in Lao PDR."

Ahmed, Ishita Afreen, Swapan Talukdar, Shahfahad, Ayesha Parvez, Mohd. Rihan, Mirza Razi Imam Baig, and Atiqur Rahman. 2022. "Flood Susceptibility Modeling in the Urban Watershed of Guwahati Using Improved Metaheuristic-Based Ensemble Machine Learning Algorithms." *Geocarto International* 37(26):12238–66. doi: 10.1080/10106049.2022.2066200.

Al-Abadi, Alaa M. 2018. "Mapping Flood Susceptibility in an Arid Region of Southern Iraq Using Ensemble Machine Learning Classifiers: A Comparative Study." *Arabian Journal of Geosciences* 11(9):218. doi: 10.1007/s12517-018-3584-5.

Al-Aizari, Ali R., Yousef A. Al-Masnay, Ali Aydda, Jiquan Zhang, Kashif Ullah, Abu Reza Md. Towfiqul Islam, Tayyiba Habib, Dawuda Usman Kaku, Jean Claude Nizeyimana, Bazel Al-Shaibah, Yasser M. Khalil, Wafaa M. M. AL-Hameedi, and Xingpeng Liu. 2022. "Assessment Analysis of Flood Susceptibility in Tropical Desert Area: A Case Study of Yemen." *Remote Sensing* 14(16):4050. doi: 10.3390/rs14164050.

AL-Areeq, Ahmed M., Hatim O. Sharif, S. I. Abba, Shakhawat Chowdhury, Mohammed Al-Suwaiyan, Mohammed Benaafi, Mohamed A. Yassin, and Isam H. Aljundi. 2023. "Digital Elevation Model for Flood Hazards Analysis in Complex Terrain: Case Study from Jeddah, Saudi Arabia." *International Journal of Applied Earth Observation and Geoinformation* 119:103330. doi: 10.1016/j.jag.2023.103330.

Albertini, Cinzia, Andrea Gioia, Vito Iacobellis, and Salvatore Manfreda. 2022. "Detection of Surface Water and Floods with Multispectral Satellites." *Remote Sensing* 14(23):6005. doi: 10.3390/rs14236005.

Aldiansyah, Septianto, and Farida Wardani. 2023. "Evaluation of Flood Susceptibility Prediction Based on a Resampling Method Using Machine Learning." *Journal of Water and Climate Change* 14(3):937–61. doi: 10.2166/wcc.2023.494.

Alin, Aylin. 2010. "Multicollinearity." *WIREs Computational Statistics* 2(3):370–74. doi: 10.1002/wics.84.

Al-Juaidi, Ahmed E. M., Ayman M. Nassar, and Omar E. M. Al-Juaidi. 2018. "Evaluation of Flood Susceptibility Mapping Using Logistic Regression and GIS Conditioning Factors." *Arabian Journal of Geosciences* 11(24):765. doi: 10.1007/s12517-018-4095-0.

Al-Kindi, Khalifa M., and Zahra Alabri. 2024. "Investigating the Role of the Key Conditioning

Factors in Flood Susceptibility Mapping Through Machine Learning Approaches." *Earth Systems and Environment*. doi: 10.1007/s41748-023-00369-7.

Amiri, Afshin, Keyvan Soltani, Isa Ebtehaj, and Hossein Bonakdari. 2024. "A Novel Machine Learning Tool for Current and Future Flood Susceptibility Mapping by Integrating Remote Sensing and Geographic Information Systems." *Journal of Hydrology* 632:130936. doi: 10.1016/j.jhydrol.2024.130936.

Andaryani, Soghra, Vahid Nourani, Ali Torabi Haghighi, and Saskia Keesstra. 2021. "Integration of Hard and Soft Supervised Machine Learning for Flood Susceptibility Mapping." *Journal of Environmental Management* 291(April):112731–112731. doi: 10.1016/j.jenvman.2021.112731.

Anon. 2022. "Lao People's Democratic Republic - Government Asian Development Bank." *Lao People's Democratic Republic - Government Asian Development Bank*. Retrieved November 17, 2023 (https://www.preventionweb.net/publication/national-strategy-disaster-risk-reduction-nsdrr-2021-2030-lao-pdr).

Anusha, N., and B. Bharathi. 2020. "Flood Detection and Flood Mapping Using Multi-Temporal Synthetic Aperture Radar and Optical Data." *The Egyptian Journal of Remote Sensing and Space Science* 23(2):207–19. doi: 10.1016/j.ejrs.2019.01.001.

Arabameri, Alireza, Khalil Rezaei, Artemi Cerdà, Christian Conoscenti, and Zahra Kalantari. 2019. "A Comparison of Statistical Methods and Multi-Criteria Decision Making to Map Flood Hazard Susceptibility in Northern Iran." *Science of the Total Environment* 660:443–58. doi: 10.1016/j.scitotenv.2019.01.021.

Arora, Aman, Alireza Arabameri, Manish Pandey, Masood A. Siddiqui, U. K. Shukla, Dieu Tien Bui, Varun Narayan Mishra, and Anshuman Bhardwaj. 2021. "Optimization of State-of-the-Art Fuzzy-Metaheuristic ANFIS-Based Machine Learning Models for Flood Susceptibility Prediction Mapping in the Middle Ganga Plain, India." *Science of the Total Environment* 750. doi: 10.1016/j.scitotenv.2020.141565.

Askar, Shavan, Sajjad Zeraat Peyma, Mohanad Mohsen Yousef, Natalia Alekseevna Prodanova, Iskandar Muda, Mohamed Elsahabi, and Javad Hatamiafkoueieh. 2022. "Flood Susceptibility Mapping Using Remote Sensing and Integration of Decision Table Classifier and Metaheuristic Algorithms." *Water (Switzerland)* 14(19). doi: 10.3390/w14193062.

Auynirundronkool, Kridsakron, Nengcheng Chen, Caihua Peng, Chao Yang, Jianya Gong, and Chaowalit Silapathong. 2012. "Flood Detection and Mapping of the Thailand Central Plain Using RADARSAT and MODIS under a Sensor Web Environment." *International Journal of Applied Earth Observation and Geoinformation* 14(1):245–55. doi: 10.1016/j.jag.2011.09.017.

Avand, Mohammadtaghi, Alban Kuriqi, Majid Khazaei, and Omid Ghorbanzadeh. 2022. "DEM Resolution Effects on Machine Learning Performance for Flood Probability Mapping." *Journal of Hydro-Environment Research* 40:1–16. doi: 10.1016/j.jher.2021.10.002.

Avand, Mohammadtaghi, Hamidreza Moradi, and Mehdi Ramazanzadeh Lasboyee. 2021. “Using Machine Learning Models, Remote Sensing, and GIS to Investigate the Effects of Changing Climates and Land Uses on Flood Probability.” *Journal of Hydrology* 595:125663. doi: 10.1016/j.jhydrol.2020.125663.

Azizi, Mojtaba, Akram Mohajerani, and Mohammadreza Akhavan. 2018. “Simulating and Prediction of Flow Using by WetSpa Model in Ziyarat River Basin, Iran.” *Open Journal of Geology* 08(03):298–312. doi: 10.4236/ojg.2018.83019.

Baghermanesh, Shadi Sadat, Shabnam Jabari, and Heather McGrath. 2022. “Urban Flood Detection Using TerraSAR-X and SAR Simulated Reflectivity Maps.” *Remote Sensing* 14(23):6154. doi: 10.3390/rs14236154.

Bahremand, A., F. De Smedt, J. Corluy, Y. B. Liu, J. Poorova, L. Velcicka, and E. Kunikova. 2007. “WetSpa Model Application for Assessing Reforestation Impacts on Floods in Margecany–Hornad Watershed, Slovakia.” *Water Resources Management* 21(8):1373–91. doi: 10.1007/s11269-006-9089-0.

Baig, Muhammad Aslam, Donghong Xiong, Mahfuzur Rahman, Md Monirul Islam, Ahmed Elbeltagi, Belayneh Yigez, Dil Kumar Rai, Muhammad Tayab, and Ashraf Dewan. 2022. “How Do Multiple Kernel Functions in Machine Learning Algorithms Improve Precision in Flood Probability Mapping?” *Natural Hazards* 113(3):1543–62. doi: 10.1007/s11069-022-05357-0.

Bakhshi Ostadkalayeh, Fatemeh, Saba Moradi, Ali Asadi, Alireza Moghaddam Nia, and Somayeh Taheri. 2023. “Performance Improvement of LSTM-Based Deep Learning Model for Streamflow Forecasting Using Kalman Filtering.” *Water Resources Management* 37(8):3111–27. doi: 10.1007/s11269-023-03492-2.

Bentivoglio, Roberto, Elvin Isufi, Sebastian Nicolaas Jonkman, and Riccardo Taormina. 2022. “Deep Learning Methods for Flood Mapping: A Review of Existing Applications and Future Research Directions.” *Hydrology and Earth System Sciences* 26(16):4345–78. doi: 10.5194/hess-26-4345-2022.

Bera, Subhas, Arup Das, and Taraknath Mazumder. 2022. “Evaluation of Machine Learning, Information Theory and Multi-Criteria Decision Analysis Methods for Flood Susceptibility Mapping under Varying Spatial Scale of Analyses.” *Remote Sensing Applications: Society and Environment* 25:100686. doi: 10.1016/j.rsase.2021.100686.

BEVEN, K. J., and M. J. KIRKBY. 1979. “A Physically Based, Variable Contributing Area Model of Basin Hydrology / Un Modèle à Base Physique de Zone d’appel Variable de l’hydrologie Du Bassin Versant.” *Hydrological Sciences Bulletin* 24(1):43–69. doi: 10.1080/02626667909491834.

Beven, K. J., M. J. Kirkby, N. Schofield, and A. F. Tagg. 1984. “Testing a Physically-Based Flood Forecasting Model (TOPMODEL) for Three U.K. Catchments.” *Journal of Hydrology* 69(1):119–43. doi: 10.1016/0022-1694(84)90159-8.

Bhattarai, Yogesh, Sunil Duwal, Sanjib Sharma, and Rocky Talchabhadel. 2024. "Leveraging Machine Learning and Open-Source Spatial Datasets to Enhance Flood Susceptibility Mapping in Transboundary River Basin." *International Journal of Digital Earth* 17(1):2313857. doi: 10.1080/17538947.2024.2313857.

Bovenga, Fabio, Antonella Belmonte, Alberto Refice, Guido Pasquariello, Raffaele Nutricato, Davide Nitti, and Maria Chiaradia. 2018. "Performance Analysis of Satellite Missions for Multi-Temporal SAR Interferometry." *Sensors* 18(5):1359. doi: 10.3390/s18051359.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45(1):5–32. doi: 10.1023/A:1010933404324.

Bruno, Leonardo Souza, Tiago Souza Mattos, Paulo Tarso Sanches Oliveira, André Almagro, and Dulce Buchala Bicca Rodrigues. 2022. "Hydrological and Hydraulic Modeling Applied to Flash Flood Events in a Small Urban Stream." *Hydrology* 9(12):223. doi: 10.3390/hydrology9120223.

Bui, Quang-Thanh, Quoc-Huy Nguyen, Xuan Linh Nguyen, Vu Dong Pham, Huu Duy Nguyen, and Van-Manh Pham. 2020. "Verification of Novel Integrations of Swarm Intelligence Algorithms into Deep Learning Neural Network for Flood Susceptibility Mapping." *Journal of Hydrology* 581:124379. doi: 10.1016/j.jhydrol.2019.124379.

Buitinck, Lars, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake Vanderplas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. 2013. "API Design for Machine Learning Software: Experiences from the Scikit-Learn Project." 1–15.

Chai, Soo See, Wei Keat Wong, and Kok Luong Goh. 2016. "Backpropagation Vs. Radial Basis Function Neural Model: Rainfall Intensity Classification For Flood Prediction Using Meteorology Data." *Journal of Computer Science* 12(4):191–200. doi: 10.3844/jcssp.2016.191.200.

Chang, Li-Chiu, Mohd Amin, Shun-Nien Yang, and Fi-John Chang. 2018. "Building ANN-Based Regional Multi-Step-Ahead Flood Inundation Forecast Models." *Water* 10(9):1283. doi: 10.3390/w10091283.

Chapi, Kamran, Vijay P. Singh, Ataollah Shirzadi, Himan Shahabi, Dieu Tien Bui, Binh Thai Pham, and Khabat Khosravi. 2017. "A Novel Hybrid Artificial Intelligence Approach for Flood Susceptibility Assessment." *Environmental Modelling and Software* 95:229–45. doi: 10.1016/j.envsoft.2017.06.012.

Chau, K. W., and J. H. W. Lee. 1991. "A Microcomputer Model for Flood Prediction with Applications." *Computer-Aided Civil and Infrastructure Engineering* 6(2):109–21. doi: 10.1111/j.1467-8667.1991.tb00181.x.

Chau, K. W., C. L. Wu, and Y. S. Li. 2005. "Comparison of Several Flood Forecasting Models in Yangtze River." *Journal of Hydrologic Engineering* 10(6):485–91. doi: 10.1061/(ASCE)1084-0699(2005)10:6(485).

Chen, Chien-Yuan, and Fan-Chieh Yu. 2011. "Morphometric Analysis of Debris Flows and Their Source Areas Using GIS." *Geomorphology* 129(3–4):387–97. doi: 10.1016/j.geomorph.2011.03.002.

Chen, Wei, Yang Li, Weifeng Xue, Himan Shahabi, Shaojun Li, Haoyuan Hong, Xiaojing Wang, Huiyuan Bian, Shuai Zhang, Biswajeet Pradhan, and Baharin Bin Ahmad. 2020. "Modeling Flood Susceptibility Using Data-Driven Approaches of Naïve Bayes Tree, Alternating Decision Tree, and Random Forest Methods." *Science of The Total Environment* 701:134979. doi: 10.1016/j.scitotenv.2019.134979.

Chormanski, Jaroslaw, Tim Van De Voorde, Tim De Roeck, Okke Batelaan, and Frank Canters. 2008. "Improving Distributed Runoff Prediction in Urbanized Catchments with Remote Sensing Based Estimates of Impervious Surface Cover." *Sensors* 8(2):910–32. doi: 10.3390/s8020910.

Choubin, Bahram, Ehsan Moradi, Mohammad Golshan, Jan Adamowski, Farzaneh Sajedi-Hosseini, and Amir Mosavi. 2019. "An Ensemble Prediction of Flood Susceptibility Using Multivariate Discriminant Analysis, Classification and Regression Trees, and Support Vector Machines." *Science of the Total Environment* 651:2087–96. doi: 10.1016/j.scitotenv.2018.10.064.

Cirella, Giuseppe T., and Felix O. Iyalomhe. 2018. "Flooding Conceptual Review: Sustainability-Focalized Best Practices in Nigeria." *Applied Sciences* 8(9):1558. doi: 10.3390/app8091558.

Cortes, Corinna, and Vladimir Vapnik. 1995. "Support-Vector Networks." *Machine Learning* 20(3):273–97. doi: 10.1007/BF00994018.

Costache, Romulus, Quoc Bao Pham, Ehsan Sharifi, Nguyen Thi Thuy Linh, S. I. Abba, Matej Vojtek, Jana Vojteková, Pham Thi Thao Nhi, and Dao Nguyen Khoi. 2020. "Flash-Flood Susceptibility Assessment Using Multi-Criteria Decision Making and Machine Learning Supported by Remote Sensing and GIS Techniques." *Remote Sensing* 12(1). doi: 10.3390/RS12010106.

Costache, Romulus, and Dieu Tien Bui. 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." *Science of The Total Environment* 691:1098–1118. doi: 10.1016/j.scitotenv.2019.07.197.

Dantzig, George B., and Mukund N. Thapa. 2013. *Linear Programming 1: Introduction*. Springer Publishing Company, Incorporated.

Das, Sumit. 2020. "Flood Susceptibility Mapping of the Western Ghat Coastal Belt Using Multi-Source Geospatial Data and Analytical Hierarchy Process (AHP)." *Remote Sensing Applications: Society and Environment* 20:100379. doi: 10.1016/j.rsase.2020.100379.

Davis, Jesse, and Mark Goadrich. 2006. "The Relationship between Precision-Recall and ROC Curves." Pp. 233–40 in *Proceedings of the 23rd international conference on Machine* *learning - ICML '06*. Pittsburgh, Pennsylvania: ACM Press.

De León Pérez, David, Rick Acosta Vega, Sergio Salazar Galán, José Ángel Aranda, and Félix Francés García. 2024. "Toward Systematic Literature Reviews in Hydrological Sciences." *Water* 16(3):436. doi: 10.3390/w16030436.

De Roo, A. P. J., C. G. Wesseling, and W. P. A. Van Deursen. 2000. "Physically Based River Basin Modelling within a GIS: The LISFLOOD Model." *Hydrological Processes* 14(11–12):1981–92. doi: 10.1002/1099-1085(20000815/30)14:11/12<1981::AID-HYP49>3.0.CO;2-F.

Dhungana, Santosh, Sangam Shrestha, Tuan Pham Van, Saurav Kc, Ashim Das Gupta, and Thi Phuoc Lai Nguyen. 2023. "Evaluation of Gridded Precipitation Products in the Selected Sub-Basins of Lower Mekong River Basin." *Theoretical and Applied Climatology* 151(1–2):293–310. doi: 10.1007/s00704-022-04268-1.

Di Baldassarre, Giuliano, Alberto Montanari, Harry Lins, Demetris Koutsoyiannis, Luigia Brandimarte, and Günter Blöschl. 2010. "Flood Fatalities in Africa: From Diagnosis to Mitigation." *Geophysical Research Letters* 37(22). doi: 10.1029/2010GL045467.

Dodangeh, Esmaeel, Bahram Choubin, Ahmad Najafi Eigdir, Narjes Nabipour, Mehdi Panahi, Shahaboddin Shamshirband, and Amir Mosavi. 2020. "Integrated Machine Learning Methods with Resampling Algorithms for Flood Susceptibility Prediction." *Science of The Total Environment* 705:135983. doi: 10.1016/j.scitotenv.2019.135983.

Durand, M., Gleason, C. J., Pavelsky, T. M., Prata de Moraes Frasson, R., Turmon, M., David, C. H., ... & Wang, J. (2023). A framework for estimating global river discharge from the Surface Water and Ocean Topography satellite mission. Water Resources Research, 59(4), e2021WR031614. doi: 10.1029/2021WR031614.

Duro, Dennis C., Steven E. Franklin, and Monique G. Dubé. 2012. "A Comparison of Pixel-Based and Object-Based Image Analysis with Selected Machine Learning Algorithms for the Classification of Agricultural Landscapes Using SPOT-5 HRG Imagery." *Remote Sensing of Environment* 118:259–72. doi: 10.1016/j.rse.2011.11.020.

Duwal, Sunil, Dedi Liu, and Prachand Man Pradhan. 2023. "Flood Susceptibility Modeling of the Karnali River Basin of Nepal Using Different Machine Learning Approaches." *Geomatics, Natural Hazards and Risk* 14(1). doi: 10.1080/19475705.2023.2217321.

El-Haddad, Bosy A., Ahmed M. Youssef, Hamid R. Pourghasemi, Biswajeet Pradhan, Abdel-Hamid El-Shater, and Mohamed H. El-Khashab. 2021. "Flood Susceptibility Prediction Using Four Machine Learning Techniques and Comparison of Their Performance at Wadi Qena Basin, Egypt." *Natural Hazards* 105(1):83–114. doi: 10.1007/s11069-020-04296-y.

Elmoulat, Meryem, and Lahcen Ait Brahim. 2018. "Landslides Susceptibility Mapping Using GIS and Weights of Evidence Model in Tetouan-Ras-Mazari Area (Northern Morocco)." *Geomatics, Natural Hazards and Risk* 9(1):1306–25. doi: 10.1080/19475705.2018.1505666.

El-Telbany, Mohammed E. 2017. "Improving the Predictability of GRNN Using Fruit Fly Optimization and PCA: The Nile Flood Forecasting." Pp. 310–19 in *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016*, edited by A. E. Hassanien, K. Shaalan, T. Gaber, A. T. Azar, and M. F. Tolba. Cham: Springer International Publishing.

Fang, Sidun, and Hongdong Wang. 2021. *Optimization-Based Energy Management for Multi-Energy Maritime Grids*. Vol. 11. Singapore: Springer Singapore.

Fang, Zhice, Yi Wang, Ling Peng, and Haoyuan Hong. 2021. "Predicting Flood Susceptibility Using LSTM Neural Networks." *Journal of Hydrology* 594:125734–125734. doi: 10.1016/j.jhydrol.2020.125734.

Farhadi, Hadi, and Mohammad Najafzadeh. 2021. "Flood Risk Mapping by Remote Sensing Data and Random Forest Technique." *Water* 13(21):3115. doi: 10.3390/w13213115.

Felix, A. Yovan, and T. Sasipraba. 2019. "Flood Detection Using Gradient Boost Machine Learning Approach." Pp. 779–83 in *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*. Dubai, United Arab Emirates: IEEE.

Fenicia, Fabrizio, Dmitri Kavetski, Hubert H. G. Savenije, Martyn P. Clark, Gerrit Schoups, Laurent Pfister, and Jim Freer. 2014. "Catchment Properties, Function, and Conceptual Model Representation: Is There a Correspondence?" *Hydrological Processes* 28(4):2451–67. doi: 10.1002/hyp.9726.

FitzGerald, Gerry, Weiwei Du, Aziz Jamal, Michele Clark, and Xiang-Yu Hou. 2010. "Flood Fatalities in Contemporary Australia (1997–2008)." *Emergency Medicine Australasia* 22(2):180–86. doi: 10.1111/j.1742-6723.2010.01284.x.

Florinsky, Igor V. 2017. "An Illustrated Introduction to General Geomorphometry." *Progress in Physical Geography: Earth and Environment* 41(6):723–52. doi: 10.1177/0309133317733667.

Fobert, Mary-Anne, Vern Singhroy, and John G. Spray. 2021. "InSAR Monitoring of Landslide Activity in Dominica." *Remote Sensing* 13(4):815. doi: 10.3390/rs13040815.

Fortin, Jean-Pierre, Richard Turcotte, Serge Massicotte, Roger Moussa, Josée Fitzback, and Jean-Pierre Villeneuve. 2001. "Distributed Watershed Model Compatible with Remote Sensing and GIS Data. I: Description of Model." *Journal of Hydrologic Engineering* 6(2):91–99. doi: 10.1061/(ASCE)1084-0699(2001)6:2(91).

French, Mark. 2018. *Fundamentals of Optimization*. Cham: Springer International Publishing.

García-Pintado, Javier, Jeff C. Neal, David C. Mason, Sarah L. Dance, and Paul D. Bates. 2013. "Scheduling Satellite-Based SAR Acquisition for Sequential Assimilation of Water Level Observations into Flood Modelling." *Journal of Hydrology* 495:252–66. doi: 10.1016/j.jhydrol.2013.03.050.

Ghorbanzadeh, Omid, Thomas Blaschke, Khalil Gholamnia, Sansar Meena, Dirk Tiede, and Jagannath Aryal. 2019. “Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection.” *Remote Sensing* 11(2):196. doi: 10.3390/rs11020196.

Gude, Vinayaka, Steven Corns, and Suzanna Long. 2020. “Flood Prediction and Uncertainty Estimation Using Deep Learning.” *Water* 12(3):884. doi: 10.3390/w12030884.

Gudiyangada Nachappa, Thimmaiah, Sepideh Tavakkoli Piralilou, Khalil Gholamnia, Omid Ghorbanzadeh, Omid Rahmati, and Thomas Blaschke. 2020. “Flood Susceptibility Mapping with Machine Learning, Multi-Criteria Decision Analysis and Ensemble Using Dempster Shafer Theory.” *Journal of Hydrology* 590:125275. doi: 10.1016/j.jhydrol.2020.125275.

Gupta, Vijay K., Ricardo Mantilla, Brent M. Troutman, David Dawdy, and Witold F. Krajewski. 2010. “Generalizing a Nonlinear Geophysical Flood Theory to Medium-Sized River Networks.” *Geophysical Research Letters* 37(11). doi: 10.1029/2009GL041540.

Hamidi, Ebrahim, Brad G. Peter, David F. Munoz, Hamed Moftakhari, and Hamid Moradkhani. 2023. “Fast Flood Extent Monitoring With SAR Change Detection Using Google Earth Engine.” *IEEE Transactions on Geoscience and Remote Sensing* 61:1–19. doi: 10.1109/TGRS.2023.3240097.

Hammami, Salma, Lahcen Zouhri, Dhekra Souissi, Ali Souei, Adel Zghibi, Amira Marzougui, and Mahmoud Dlala. 2019. “Application of the GIS Based Multi-Criteria Decision Analysis and Analytical Hierarchy Process (AHP) in the Flood Susceptibility Mapping (Tunisia).” *Arabian Journal of Geosciences* 12(21):653. doi: 10.1007/s12517-019-4754-9.

Hanley, J. A. 1989. “Receiver Operating Characteristic (ROC) Methodology: The State of the Art.” *Critical Reviews in Diagnostic Imaging* 29(3):307–35.

Hansana, Phonekham, Xin Guo, Shuo Zhang, Xudong Kang, and Shutao Li. 2023. “Flood Analysis Using Multi-Scale Remote Sensing Observations in Laos.” *Remote Sensing* 15(12):3166. doi: 10.3390/rs15123166.

Hasanuzzaman, Md, Aznarul Islam, Biswajit Bera, and Pravat Kumar Shit. 2022. “A Comparison of Performance Measures of Three Machine Learning Algorithms for Flood Susceptibility Mapping of River Silabati (Tropical River, India).” *Physics and Chemistry of the Earth, Parts A/B/C* 127:103198. doi: 10.1016/j.pce.2022.103198.

Hillier, Frederick S., and Gerald J. Lieberman. 2015. *Introduction to Operations Research*. McGraw-Hill.

Hitouri, Sliman, Meriame Mohajane, Meriam Lahsaini, Sk Ajim Ali, Tadesual Asamin Setargie, Gaurav Tripathi, Paola D’Antonio, Suraj Kumar Singh, and Antonietta Varasano. 2024. “Flood Susceptibility Mapping Using SAR Data and Machine Learning Algorithms in a Small Watershed in Northwestern Morocco.” *Remote Sensing* 16(5):858. doi: 10.3390/rs16050858.

Hochreiter Sepp and Schmidhuber Jürgen. n.d. "Long Short-Term Memory." *Neural Computation* 9(8):1735–80. doi: https://doi.org/10.1162/neco.1997.9.8.1735.

Horritt, M. 2003. "Waterline Mapping in Flooded Vegetation from Airborne SAR Imagery." *Remote Sensing of Environment* 85(3):271–81. doi: 10.1016/S0034-4257(03)00006-3.

Horritt, M. S., D. C. Mason, and A. J. Luckman. 2001. "Flood Boundary Delineation from Synthetic Aperture Radar Imagery Using a Statistical Active Contour Model." *International Journal of Remote Sensing* 22(13):2489–2507. doi: 10.1080/01431160116902.

Horritt, MatthewS. 1999. "A Statistical Active Contour Model for SAR Image Segmentation." *Image and Vision Computing* 17(3–4):213–24. doi: 10.1016/S0262-8856(98)00101-2.

Huang, Jun, Pute Wu, and Xining Zhao. 2013. "Effects of Rainfall Intensity, Underlying Surface and Slope Gradient on Soil Infiltration under Simulated Rainfall Experiments." *CATENA* 104:93–102. doi: 10.1016/j.catena.2012.10.013.

Ibarra-Zavaleta, Sara, Rosario Landgrave, Rabindranarth Romero-López, Annie Poulin, and Raúl Arango-Miranda. 2017. "Distributed Hydrological Modeling: Determination of Theoretical Hydraulic Potential & Streamflow Simulation of Extreme Hydrometeorological Events." *Water* 9(8):602. doi: 10.3390/w9080602.

Ighile, Eseosa Halima, Hiroaki Shirakawa, and Hiroki Tanikawa. 2022. "Application of GIS and Machine Learning to Predict Flood Areas in Nigeria." *Sustainability* 14(9):5039. doi: 10.3390/su14095039.

Imamura, Yoshiyuki. 2022. "Development of a Method for Assessing Country-Based Flood Risk at the Global Scale." *International Journal of Disaster Risk Science* 13(1):87–99. doi: 10.1007/s13753-021-00388-w.

Ingle Smith, D. 1999. "Floods: Physical Processes and Human Impacts by K. Smith and R. Ward, John Wiley, Chichester 1998. No. of Pages: 382." *Earth Surface Processes and Landforms* 24(13):1261–1261. doi: 10.1002/(SICI)1096-9837(199912)24:13<1261::AID-ESP22>3.0.CO;2-#.

J Janizadeh, S., Avand, M., Jaafari, A., Phong, T. V., Bayat, M., Ahmadisharaf, E., ... & Lee, S. (2019). Prediction success of machine learning methods for flash flood susceptibility mapping in the Tafresh watershed, Iran. Sustainability, 11(19), 5426. doi: 10.3390/su11195426.

Januário, Tomásio Eduardo, Augusto José Pereira Filho, and Marcos Figueiredo Salviano. 2022. "Hydrometeorological Modeling of Limpopo River Basin in Mozambique with TOPMODEL and Remote Sensing." *Open Journal of Modern Hydrology* 12(2):55–73. doi: 10.4236/ojmh.2022.122004.

Jayakrishnan, R., R. Srinivasan, C. Santhi, and J. G. Arnold. 2005. "Advances in the Application of the SWAT Model for Water Resources Management." *Hydrological Processes*

19(3):749–62. doi: 10.1002/hyp.5624.

Jiang, X., Liang, S., He, X., Ziegler, A. D., Lin, P., Pan, M., ... & Zeng, Z. (2021). Rapid and large-scale mapping of flood inundation via integrating spaceborne synthetic aperture radar imagery with unsupervised deep learning. ISPRS journal of photogrammetry and remote sensing, 178, 36-50. doi: 10.1016/j.isprsjprs.2021.05.019.

Jimeno-Sáez, Patricia, Raquel Martínez-España, Javier Casalí, Julio Pérez-Sánchez, and Javier Senent-Aparicio. 2022. "A Comparison of Performance of SWAT and Machine Learning Models for Predicting Sediment Load in a Forested Basin, Northern Spain." *CATENA* 212:105953. doi: 10.1016/j.catena.2021.105953.

Kaelbling, L. P., M. L. Littman, and A. W. Moore. 1996. "Reinforcement Learning: A Survey." *Journal of Artificial Intelligence Research* 4:237–85. doi: 10.1613/jair.301.

Kaiser, Maria, Stephan Günnemann, and Markus Disse. 2022. "Regional-Scale Prediction of Pluvial and Flash Flood Susceptible Areas Using Tree-Based Classifiers." *Journal of Hydrology* 612:128088. doi: 10.1016/j.jhydrol.2022.128088.

Karjalainen, Mika, Ville Kankare, Mikko Vastaranta, Markus Holopainen, and Juha Hyyppä. 2012. "Prediction of Plot-Level Forest Variables Using TerraSAR-X Stereo SAR Data." *Remote Sensing of Environment* 117:338–47. doi: 10.1016/j.rse.2011.10.008.

Kasiviswanathan, K. S., Jianxun He, K. P. Sudheer, and Joo-Hwa Tay. 2016. "Potential Application of Wavelet Neural Network Ensemble to Forecast Streamflow for Flood Management." *Journal of Hydrology* 536:161–73. doi: 10.1016/j.jhydrol.2016.02.044.

Kavzoglu, Taskin, Emrehan Kutlug Sahin, and Ismail Colkesen. 2015. "Selecting Optimal Conditioning Factors in Shallow Translational Landslide Susceptibility Mapping Using Genetic Algorithm." *Engineering Geology* 192:101–12. doi: 10.1016/j.enggeo.2015.04.004.

Kaya, Cagla Melisa, and Leyla Derin. 2023. "Parameters and Methods Used in Flood Susceptibility Mapping: A Review." *Journal of Water and Climate Change* 14(6):1935–60. doi: 10.2166/wcc.2023.035.

Keophila, V., A. Promwungkwa, and K. Ngamsanroaj. 2019. "Effectiveness of Cascades Reservoir for Flood Control Operation and Electricity Production in Nam Ngum River." *Journal of Physics: Conference Series* 1175:012276. doi: 10.1088/1742-6596/1175/1/012276.

Khalil, Usman, and Noor Muhammad Khan. 2017. "Floodplain Mapping for Indus River: Chashma –Taunsa Reach."

Khosravi, Khabat, Ebrahim Nohani, Edris Maroufinia, and Hamid Reza Pourghasemi. 2016. "A GIS-Based Flood Susceptibility Assessment and Its Mapping in Iran: A Comparison between Frequency Ratio and Weights-of-Evidence Bivariate Statistical Models with Multi-Criteria Decision-Making Technique." *Natural Hazards* 83(2):947–87. doi: 10.1007/s11069-016-2357-2.

Khosravi, Khabat, Binh Thai Pham, Kamran Chapi, Ataollah Shirzadi, Himan Shahabi, Inge Revhaug, Indra Prakash, and Dieu Tien Bui. 2018. "A Comparative Assessment of Decision Trees Algorithms for Flash Flood Susceptibility Modeling at Haraz Watershed, Northern Iran." *Science of The Total Environment* 627:744–55. doi: 10.1016/j.scitotenv.2018.01.266.

Khosravi, Khabat, Himan Shahabi, Binh Thai Pham, Jan Adamowski, Ataollah Shirzadi, Biswajeet Pradhan, Jie Dou, Hai Bang Ly, Gyula Gróf, Huu Loc Ho, Haoyuan Hong, Kamran Chapi, and Indra Prakash. 2019. "A Comparative Assessment of Flood Susceptibility Modeling Using Multi-Criteria Decision-Making Analysis and Machine Learning Methods." *Journal of Hydrology* 573(March):311–23. doi: 10.1016/j.jhydrol.2019.03.073.

Kim, Sooyoul, Yoshiharu Matsumi, Shunqi Pan, and Hajime Mase. 2016. "A Real-Time Forecast Model Using Artificial Neural Network for after-Runner Storm Surges on the Tottori Coast, Japan." *Ocean Engineering* 122:44–53. doi: 10.1016/j.oceaneng.2016.06.017.

Kim, Yunjee, and Moung-Jin Lee. 2020. "Rapid Change Detection of Flood Affected Area after Collapse of the Laos Xe-Pian Xe-Namnoy Dam Using Sentinel-1 GRD Data." *Remote Sensing* 12(12):1978. doi: 10.3390/rs12121978.

Kimmany, Bounhome, Piyatida Ruangrassamee, and Supattra Visessri. 2020. "Optimal Multi-Reservoir Operation for Hydropower Production in the Nam Ngum River Basin." *Engineering Journal* 24(5):1–13. doi: 10.4186/ej.2020.24.5.1.

Kron, Wolfgang. 2005. "Flood Risk = Hazard • Values • Vulnerability." *Water International* 30(1):58–68. doi: 10.1080/02508060508691837.

Kulithalai Shiyam Sundar, Parthasarathy, and Subrahmanya Kundapura. 2023. "Spatial Mapping of Flood Susceptibility Using Decision Tree–Based Machine Learning Models for the Vembanad Lake System in Kerala, India." *Journal of Water Resources Planning and Management* 149(10):04023052. doi: 10.1061/JWRMD5.WRENG-5858.

Kumar, Rajesh, and Prasenjit Acharya. 2016. "Flood Hazard and Risk Assessment of 2014 Floods in Kashmir Valley: A Space-Based Multisensor Approach." *Natural Hazards* 84(1):437–64. doi: 10.1007/s11069-016-2428-4.

Kundzewicz, Zbigniew W., Shinjiro Kanae, Sonia I. Seneviratne, John Handmer, Neville Nicholls, Pascal Peduzzi, Reinhard Mechler, Laurens M. Bouwer, Nigel Arnell, Katharine Mach, Robert Muir-Wood, G. Robert Brakenridge, Wolfgang Kron, Gerardo Benito, Yasushi Honda, Kiyoshi Takahashi, and Boris Sherstyukov. 2014. "Le Risque d'inondation et Les Perspectives de Changement Climatique Mondial et Régional." *Hydrological Sciences Journal* 59(1):1–28. doi: 10.1080/02626667.2013.857411.

Lavers, David A., Adrian Simmons, Freja Vamborg, and Mark J. Rodwell. 2022. "An Evaluation of ERA5 Precipitation for Climate Monitoring." *Quarterly Journal of the Royal Meteorological Society* 148(748):3152–65. doi: 10.1002/qj.4351.

Lei, Tianjie, Jiabao Wang, Xiangyu Li, Weiwei Wang, Changliang Shao, and Baoyin Liu. 2022.

"Flood Disaster Monitoring and Emergency Assessment Based on Multi-Source Remote Sensing Observations." *Water* 14(14):2207. doi: 10.3390/w14142207.

Li, Kaizhong, Shaohong Wu, Erfu Dai, and Zhongchun Xu. 2012. "Flood Loss Analysis and Quantitative Risk Assessment in China." *Natural Hazards* 63(2):737–60. doi: 10.1007/s11069-012-0180-y.

Li, Yuting, and Haoyuan Hong. 2023. "Modelling Flood Susceptibility Based on Deep Learning Coupling with Ensemble Learning Models." *Journal of Environmental Management* 325:116450. doi: 10.1016/j.jenvman.2022.116450.

Löwe, Roland, Julian Böhm, David Getreuer Jensen, Jorge Leandro, and Søren Højmark Rasmussen. 2021. "U-FLOOD – Topographic Deep Learning for Predicting Urban Pluvial Flood Water Depth." *Journal of Hydrology* 603:126898. doi: 10.1016/j.jhydrol.2021.126898.

Ma, Meihong, Gang Zhao, Bingshun He, Qing Li, Haoyue Dong, Shenggang Wang, and Zhongliang Wang. 2021. "XGBoost-Based Method for Flash Flood Risk Assessment." *Journal of Hydrology* 598:126382. doi: 10.1016/j.jhydrol.2021.126382.

Mangkhaseum, Sackdavong, Yogesh Bhattarai, Sunil Duwal, and Akitoshi Hanazawa. 2024. "Flood Susceptibility Mapping Leveraging Open-Source Remote-Sensing Data and Machine Learning Approaches in Nam Ngum River Basin (NNRB), Lao PDR." *Geomatics, Natural Hazards and Risk* 15(1):2357650. doi: 10.1080/19475705.2024.2357650.

Mangkhaseum, Sackdavong, and Akitoshi Hanazawa. 2021. "Comparison of Machine Learning Classifiers for Land Cover Changes Using Google Earth Engine." Pp. 1–7 in *2021 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES)*. Bali, Indonesia: IEEE.

Marjanović, Miloš, Miloš Kovačević, Branislav Bajat, and Vít Voženílek. 2011. "Landslide Susceptibility Assessment Using SVM Machine Learning Algorithm." *Engineering Geology* 123(3):225–34. doi: 10.1016/j.enggeo.2011.09.006.

Martinis, Sandro, Simon Plank, and Kamila Ćwik. 2018. "The Use of Sentinel-1 Time-Series Data to Improve Flood Monitoring in Arid Areas." *Remote Sensing* 10(4):583. doi: 10.3390/rs10040583.

Masafu, Christopher, and Richard Williams. 2024. "Satellite Video Remote Sensing for Flood Model Validation." *Water Resources Research* 60(1):e2023WR034545. doi: 10.1029/2023WR034545.

Mastro, Pietro, Guido Masiello, Carmine Serio, and Antonio Pepe. 2022. "Change Detection Techniques with Synthetic Aperture Radar Images: Experiments with Random Forests and Sentinel-1 Observations." *Remote Sensing* 14(14):3323. doi: 10.3390/rs14143323.

Matgen, P., R. Hostache, G. Schumann, L. Pfister, L. Hoffmann, and H. H. G. Savenije. 2011. "Towards an Automated SAR-Based Flood Monitoring System: Lessons Learned from

Two Case Studies." *Physics and Chemistry of the Earth, Parts A/B/C* 36(7–8):241–52. doi: 10.1016/j.pce.2010.12.009.

Matgen, P., G. Schumann, J. B. Henry, L. Hoffmann, and L. Pfister. 2007. "Integration of SAR-Derived River Inundation Areas, High-Precision Topographic Data and a River Flow Model toward near Real-Time Flood Management." *International Journal of Applied Earth Observation and Geoinformation* 9(3):247–63. doi: 10.1016/j.jag.2006.03.003.

Meema, Thatkiat, Yasuto Tachikawa, Yutaka Ichikawa, and Kazuaki Yorozu. 2021. "Uncertainty Assessment of Water Resources and Long-Term Hydropower Generation Using a Large Ensemble of Future Climate Projections for the Nam Ngum River in the Mekong Basin." *Journal of Hydrology: Regional Studies* 36:100856. doi: 10.1016/j.ejrh.2021.100856.

Megahed, Hanaa A., Amira M. Abdo, Mohamed A. E. AbdelRahman, Antonio Scopa, and Mohammed N. Hegazy. 2023. "Frequency Ratio Model as Tools for Flood Susceptibility Mapping in Urbanized Areas: A Case Study from Egypt." *Applied Sciences* 13(16):9445. doi: 10.3390/app13169445.

Mehravar, Soroosh, Seyed Vahid Razavi-Termeh, Armin Moghimi, Babak Ranjgar, Fatemeh Foroughnia, and Meisam Amani. 2023. "Flood Susceptibility Mapping Using Multi-Temporal SAR Imagery and Novel Integration of Nature-Inspired Algorithms into Support Vector Regression." *Journal of Hydrology* 617(PC):129100–129100. doi: 10.1016/j.jhydrol.2023.129100.

Meyer, Volker, Sebastian Scheuer, and Dagmar Haase. 2009. "A Multicriteria Approach for Flood Risk Mapping Exemplified at the Mulde River, Germany." *Natural Hazards* 48(1):17–39. doi: 10.1007/s11069-008-9244-4.

Miles, Jeremy. 2014. "Tolerance and Variance Inflation Factor." in *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd.

Minar, Matiur Rahman, and Jibon Naher. 2018. "Recent Advances in Deep Learning: An Overview."

Mirzaei, Sajjad, Mehdi Vafakhah, Biswajeet Pradhan, and Seyed Jalil Alavi. 2021. "Flood Susceptibility Assessment Using Extreme Gradient Boosting (EGB), Iran." *Earth Science Informatics* 14(1):51–67. doi: 10.1007/s12145-020-00530-0.

Mojaddadi, Hossein, Biswajeet Pradhan, Haleh Nampak, Noordin Ahmad, and Abdul Halim bin Ghazali. 2017. "Ensemble Machine-Learning-Based Geospatial Approach for Flood Risk Assessment Using Multi-Sensor Remote-Sensing Data and GIS." *Geomatics, Natural Hazards and Risk* 8(2):1080–1102. doi: 10.1080/19475705.2017.1294113.

Mollaei, Zeinab, Kamran Davary, Seyed Majid Hasheminia, Alireza Faridhosseini, and Yavar Pourmohamad. 2018. "Enhancing Flood Hazard Estimation Methods on Alluvial Fans Using an Integrated Hydraulic, Geological and Geomorphological Approach." *Natural Hazards and Earth System Sciences* 18(4):1159–71. doi: 10.5194/nhess-18-1159-2018.

Moore, I. D., R. B. Grayson, and A. R. Ladson. 1991. “Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications.” *Hydrological Processes* 5(1):3–30. doi: 10.1002/hyp.3360050103.

Morrison, A., C. j. Westbrook, and B. f. Noble. 2018. “A Review of the Flood Risk Management Governance and Resilience Literature.” *Journal of Flood Risk Management* 11(3):291–304. doi: 10.1111/jfr3.12315.

Mucsi, László, and Dang Hung Bui. 2023. “Evaluating the Performance of Multi-Temporal Synthetic-Aperture Radar Imagery in Land-Cover Mapping Using a Forward Stepwise Selection Approach.” *Remote Sensing Applications: Society and Environment* 30:100975. doi: 10.1016/j.rsase.2023.100975.

Muthu, Kalidhas, and Sivakumar Ramamoorthy. 2024. “Evaluation of Urban Flood Susceptibility through Integrated Bivariate Statistics and Geospatial Technology.” *Environmental Monitoring and Assessment* 196(6):526. doi: 10.1007/s10661-024-12676-1.

Na, Xiaodong, and Wenliang Li. 2022. “Modeling Hydrological Regimes of Floodplain Wetlands Using Remote Sensing and Field Survey Data.” *Water* 14(24):4126. doi: 10.3390/w14244126.

Nagler, Thomas, Helmut Rott, Markus Hetzenecker, Jan Wuite, and Pierre Potin. 2015. “The Sentinel-1 Mission: New Opportunities for Ice Sheet Observations.” *Remote Sensing* 7(7):9371–89. doi: 10.3390/rs70709371.

Nguyen, Hong Ngoc, Hiroatsu Fukuda, and Minh Nguyet Nguyen. 2024. “Assessment of the Susceptibility of Urban Flooding Using GIS with an Analytical Hierarchy Process in Hanoi, Vietnam.” *Sustainability* 16(10):3934. doi: 10.3390/su16103934.

Nguyen, Huu Duy. 2022. “Flood Susceptibility Assessment Using Hybrid Machine Learning and Remote Sensing in Quang Tri Province, Vietnam.” *Transactions in GIS* 26(7):2776–2801. doi: 10.1111/tgis.12980.

Notti, Davide, Daniele Giordan, Fabiana Caló, Antonio Pepe, Francesco Zucca, and Jorge Pedro Galve. 2018. “Potential and Limitations of Open Satellite Data for Flood Mapping.” *Remote Sensing* 10(11):1673. doi: 10.3390/rs10111673.

Oberstadler, R., H. Hönsch, and D. Huth. 1997. “Assessment of the Mapping Capabilities of ERS-1 SAR Data for Flood Mapping: A Case Study in Germany.” *Hydrological Processes* 11(10):1415–25. doi: 10.1002/(SICI)1099-1085(199708)11:10<1415::AID-HYP532>3.0.CO;2-2.

Opperman, Jeffrey J., Gerald E. Galloway, Joseph Fargione, Jeffrey F. Mount, Brian D. Richter, and Silvia Secchi. 2009. “Sustainable Floodplains Through Large-Scale Reconnection to Rivers.” *Science* 326(5959):1487–88. doi: 10.1126/science.1178256.

Palanisamy, Prathiba A., Kamal Jain, and Stefania Bonafoni. 2023. “Machine Learning Classifier Evaluation for Different Input Combinations: A Case Study with Landsat 9 and Sentinel-2

Data." *Remote Sensing* 15(13):3241. doi: 10.3390/rs15133241.

Pande, Chaitanya B. 2022. "Land Use/Land Cover and Change Detection Mapping in Rahuri Watershed Area (MS), India Using the Google Earth Engine and Machine Learning Approach." *Geocarto International* 37(26):13860–80. doi: 10.1080/10106049.2022.2086622.

Panigrahi, Bharati, Krishna Chaitanya Rao Kathala, and M. Sujatha. 2023. "A Machine Learning-Based Comparative Approach to Predict the Crop Yield Using Supervised Learning With Regression Models." *Procedia Computer Science* 218:2684–93. doi: 10.1016/j.procs.2023.01.241.

Pang, Q., and Dimailig, L. A. (2019). Distribution of Disasters in ASEAN, in the AHA Centre, 2019. ASEAN Risk Monitor and Disaster Management Review (ARMOR),.

Poudyal, Chandra Prakash, Chandong Chang, Hyun-Joo Oh, and Saro Lee. 2010. "Landslide Susceptibility Maps Comparing Frequency Ratio and Artificial Neural Networks: A Case Study from the Nepal Himalaya." *Environmental Earth Sciences* 61(5):1049–64. doi: 10.1007/s12665-009-0426-5.

Pourghasemi, Hamid Reza, Narges Kariminejad, Mahdis Amiri, Mohsen Edalat, Mehrdad Zarafshar, Thomas Blaschke, and Artemio Cerda. 2020. "Assessing and Mapping Multi-Hazard Risk Susceptibility Using a Machine Learning Technique." *Scientific Reports* 10(1):1–11. doi: 10.1038/s41598-020-60191-3.

Pradhan, Biswajeet. 2013. "A Comparative Study on the Predictive Ability of the Decision Tree, Support Vector Machine and Neuro-Fuzzy Models in Landslide Susceptibility Mapping Using GIS." *Computers & Geosciences* 51:350–65. doi: 10.1016/j.cageo.2012.08.023.

Pradhan, Biswajeet, Mahyat Shafapour Tehrany, and Mustafa Neamah Jebur. 2016. "A New Semiautomated Detection Mapping of Flood Extent From TerraSAR-X Satellite Image Using Rule-Based Classification and Taguchi Optimization Techniques." *IEEE Transactions on Geoscience and Remote Sensing* 54(7):4331–42. doi: 10.1109/TGRS.2016.2539957.

Priscillia, Stela, Calogero Schillaci, and Aldo Lipani. 2021. "Flood Susceptibility Assessment Using Artificial Neural Networks in Indonesia." *Artificial Intelligence in Geosciences* 2:215–22. doi: 10.1016/j.aiig.2022.03.002.

Pulvirenti, Luca, Giuseppe Squicciarino, Elisabetta Fiori, Laura Candela, and Silvia Puca. 2023. "Analysis and Processing of the COSMO-SkyMed Second Generation Images of the 2022 Marche (Central Italy) Flood." *Water* 15(7):1353. doi: 10.3390/w15071353.

Rahmati, Omid, Hamid Reza Pourghasemi, and Hossein Zeinivand. 2016. "Flood Susceptibility Mapping Using Frequency Ratio and Weights-of-Evidence Models in the Golastan Province, Iran." *Geocarto International* 31(1):42–70. doi: 10.1080/10106049.2015.1041559.

Rappaport, Edward N. 2014. "Fatalities in the United States from Atlantic Tropical Cyclones: New Data and Interpretation." *Bulletin of the American Meteorological Society* 95(3):341–46. doi: 10.1175/BAMS-D-12-00074.1.

Razavi Termeh, Seyed Vahid, Aiding Kornejady, Hamid Reza Pourghasemi, and Saskia Keesstra. 2018. "Flood Susceptibility Mapping Using Novel Ensembles of Adaptive Neuro Fuzzy Inference System and Metaheuristic Algorithms." *Science of The Total Environment* 615:438–51. doi: 10.1016/j.scitotenv.2017.09.262.

Razavi-Termeh, Seyed Vahid, MyoungBae Seo, Abolghasem Sadeghi-Niaraki, and Soo-Mi Choi. 2023. "Flash Flood Detection and Susceptibility Mapping in the Monsoon Period by Integration of Optical and Radar Satellite Imagery Using an Improvement of a Sequential Ensemble Algorithm." *Weather and Climate Extremes* 41:100595. doi: 10.1016/j.wace.2023.100595.

Refsgaard, Jens Christian. 1997. "Parameterisation, Calibration and Validation of Distributed Hydrological Models." *Journal of Hydrology* 198(1–4):69–97. doi: 10.1016/S0022-1694(96)03329-X.

Reichenbach, Paola, Mauro Rossi, Bruce D. Malamud, Monika Mihir, and Fausto Guzzetti. 2018. "A Review of Statistically-Based Landslide Susceptibility Models." *Earth-Science Reviews* 180:60–91. doi: 10.1016/j.earscirev.2018.03.001.

Rezaie, Fatemeh, Mahdi Panahi, Sayed M. Bateni, Changhyun Jun, Christopher M. U. Neale, and Saro Lee. 2022. "Novel Hybrid Models by Coupling Support Vector Regression (SVR) with Meta-Heuristic Algorithms (WOA and GWO) for Flood Susceptibility Mapping." *Natural Hazards* 114(2):1247–83. doi: 10.1007/s11069-022-05424-6.

Saber, Mohamed, Tayeb Boulmaiz, Mawloud Guermoui, Karim I. Abdrabo, Sameh A. Kantoush, Tetsuya Sumi, Hamouda Boutaghane, Tomoharu Hori, Doan Van Binh, Binh Quang Nguyen, Thao T. P. Bui, Ngoc Duong Vo, Emad Habib, and Emad Mabrouk. 2023. "Enhancing Flood Risk Assessment through Integration of Ensemble Learning Approaches and Physical-Based Hydrological Modeling." *Geomatics, Natural Hazards and Risk* 14(1). doi: 10.1080/19475705.2023.2203798.

Saber, Mohamed, Tayeb Boulmaiz, Mawloud Guermoui, Karim I. Abdrabo, Sameh A. Kantoush, Tetsuya Sumi, Hamouda Boutaghane, Daisuke Nohara, and Emad Mabrouk. 2022. "Examining LightGBM and CatBoost Models for Wadi Flash Flood Susceptibility Prediction." *Geocarto International* 37(25):7462–87. doi: 10.1080/10106049.2021.1974959.

Saha, Tamal Kanti, Swades Pal, Swapan Talukdar, Sandipta Debanshi, Rumki Khatun, Pankaj Singha, and Indrajit Mandal. 2021. "How Far Spatial Resolution Affects the Ensemble Machine Learning Based Flood Susceptibility Prediction in Data Sparse Region." *Journal of Environmental Management* 297:113344. doi: 10.1016/j.jenvman.2021.113344.

Salem, Abdella, and Leila Hashemi-Beni. 2022. "Inundated Vegetation Mapping Using SAR Data: A Comparison of Polarization Configurations of UAVSAR L-Band and Sentinel C-Band."

Remote Sensing 14(24):6374. doi: 10.3390/rs14246374.

Samanta, Sailesh, Dilip Kumar Pal, and Babita Palsamanta. 2018. “Flood Susceptibility Analysis through Remote Sensing, GIS and Frequency Ratio Model.” *Applied Water Science* 8(2):66. doi: 10.1007/s13201-018-0710-1.

Samui, Pijush. 2008. “Slope Stability Analysis: A Support Vector Machine Approach.” *Environmental Geology* 56(2):255–67. doi: 10.1007/s00254-007-1161-4.

Sanyal, Joy, and X. X. Lu. 2004. “Application of Remote Sensing in Flood Management with Special Reference to Monsoon Asia: A Review.” *Natural Hazards* 33(2):283–301. doi: 10.1023/B:NHAZ.0000037035.65105.95.

Saravanan, Subbarayan, Devanantham Abijith, Nagireddy Masthan Reddy, Parthasarathy Kss, Niraimathi Janardhanam, Subbarayan Sathiyamurthi, and Vivek Sivakumar. 2023. “Flood Susceptibility Mapping Using Machine Learning Boosting Algorithms Techniques in Idukki District of Kerala India.” *Urban Climate* 49:101503. doi: 10.1016/j.uclim.2023.101503.

Šaur, David. 2017. “Forecasting of Convective Precipitation Through NWP Models and Algorithm of Storms Prediction.” Pp. 125–36 in *Artificial Intelligence Trends in Intelligent Systems*, edited by R. Silhavy, R. Senkerik, Z. Kominkova Oplatkova, Z. Prokopova, and P. Silhavy. Cham: Springer International Publishing.

Schlaffer, Stefan, Patrick Matgen, Markus Hollaus, and Wolfgang Wagner. 2015. “Flood Detection from Multi-Temporal SAR Data Using Harmonic Analysis and Change Detection.” *International Journal of Applied Earth Observation and Geoinformation* 38:15–24. doi: 10.1016/j.jag.2014.12.001.

Selvam, Richard Abishek, and Antony Ravindran Antony Jebamalai. 2023. “Application of the Analytical Hierarchy Process (AHP) for Flood Susceptibility Mapping Using GIS Techniques in Thamirabarani River Basin, Srivaikundam Region, Southern India.” *Natural Hazards* 118(2):1065–83. doi: 10.1007/s11069-023-06037-3.

Shafapourtehrany, Mahyat, Fatemeh Rezaie, Changhyun Jun, Essam Heggy, Sayed M. Bateni, Mahdi Panahi, Haluk Özener, Farzin Shabani, and Hamidreza Moeini. 2023. “Mapping Post-Earthquake Landslide Susceptibility Using U-Net, VGG-16, VGG-19, and Metaheuristic Algorithms.” *Remote Sensing* 15(18):4501. doi: 10.3390/rs15184501.

Shafizadeh-Moghadam, Hossein, Roozbeh Valavi, Himan Shahabi, Kamran Chapi, and Ataollah Shirzadi. 2018. “Novel Forecasting Approaches Using Combination of Machine Learning and Statistical Models for Flood Susceptibility Mapping.” *Journal of Environmental Management* 217:1–11. doi: 10.1016/j.jenvman.2018.03.089.

Shahabi, Himan., Ataollah. Shirzadi, Kayvan. Ghaderi, Ebrahim. Omidvar, Nadhir. Al-Ansari, John J. Clague, Marten. Geertsema, Khabat. Khosravi, Ata. Amini, Sepideh. Bahrami, Omid. Rahmati, Kyoumars. Habibi, Ayub. Mohammadi, Hoang. Nguyen, Assefa M. Melesse, Baharin Bin. Ahmad, and Anuar. Ahmad. 2020. “Flood Detection and

Susceptibility Mapping Using Sentinel-1 Remote Sensing Data and a Machine Learning Approach: Hybrid Intelligence of Bagging Ensemble Based on K-Nearest Neighbor Classifier." *Remote Sensing* 12(2). doi: 10.3390/rs12020266.

Sharma, Manju, Vipin Kumar, and Sandeep Kumar. 2024. "A Systematic Review of Urban Sprawl and Land Use/Land Cover Change Studies in India." *Sustainable Environment* 10(1):2331269. doi: 10.1080/27658511.2024.2331269.

Shastry, Apoorva, Elizabeth Carter, Brian Coltin, Rachel Sleeter, Scott McMichael, and Jack Eggleston. 2023. "Mapping Floods from Remote Sensing Data and Quantifying the Effects of Surface Obstruction by Clouds and Vegetation." *Remote Sensing of Environment* 291:113556. doi: 10.1016/j.rse.2023.113556.

Shrestha, Ajay, and Ausif Mahmood. 2019. "Review of Deep Learning Algorithms and Architectures." *IEEE Access* 7:53040–65. doi: 10.1109/ACCESS.2019.2912200.

Singh, Ankita, and Vipin Vyas. 2022. "A Review on Remote Sensing Application in River Ecosystem Evaluation." *Spatial Information Research* 30(6):759–72. doi: 10.1007/s41324-022-00470-5.

Sørensen, R., U. Zinko, and J. Seibert. 2006. "On the Calculation of the Topographic Wetness Index: Evaluation of Different Methods Based on FIeld Observations." *Hydrology and Earth System Sciences*.

Swain, Kishore Chandra, Chiranjit Singha, and Laxmikanta Nayak. 2020. "Flood Susceptibility Mapping through the GIS-AHP Technique Using the Cloud." *ISPRS International Journal of Geo-Information* 9(12):720. doi: 10.3390/ijgi9120720.

Talukdar, Swapan, Bonosri Ghose, Shahfahad, Roquia Salam, Susanta Mahato, Quoc Bao Pham, Nguyen Thi Thuy Linh, Romulus Costache, and Mohammadtaghi Avand. 2020. "Flood Susceptibility Modeling in Teesta River Basin, Bangladesh Using Novel Ensembles of Bagging Algorithms." *Stochastic Environmental Research and Risk Assessment* 34(12):2277–2300. doi: 10.1007/s00477-020-01862-5.

Tang, Xianzhe, Jiufeng Li, Minnan Liu, Wei Liu, and Haoyuan Hong. 2020. "Flood Susceptibility Assessment Based on a Novel Random Naïve Bayes Method: A Comparison between Different Factor Discretization Methods." *Catena* 190(September):104536–104536. doi: 10.1016/j.catena.2020.104536.

Tariq, Muhammad Atiq Ur Rehman, Rashid Farooq, and Nick Van De Giesen. 2020. "A Critical Review of Flood Risk Management and the Selection of Suitable Measures." *Applied Sciences* 10(23):8752. doi: 10.3390/app10238752.

Tehrany, Mahyat Shafapour, Simon Jones, and Farzin Shabani. 2019. "Identifying the Essential Flood Conditioning Factors for Flood Prone Area Mapping Using Machine Learning Techniques." *Catena* 175(April 2018):174–92. doi: 10.1016/j.catena.2018.12.011.

Tehrany, Mahyat Shafapour, and Lalit Kumar. 2018. "The Application of a Dempster–Shafer-

Based Evidential Belief Function in Flood Susceptibility Mapping and Comparison with Frequency Ratio and Logistic Regression Methods." *Environmental Earth Sciences* 77(13):490. doi: 10.1007/s12665-018-7667-0.

Tehrany, Mahyat Shafapour, Lalit Kumar, and Farzin Shabani. 2019. "A Novel GIS-Based Ensemble Technique for Flood Susceptibility Mapping Using Evidential Belief Function and Support Vector Machine: Brisbane, Australia." *PeerJ* 2019(10). doi: 10.7717/peerj.7653.

Tehrany, Mahyat Shafapour, Moung-Jin Lee, Biswajeet Pradhan, Mustafa Neamah Jebur, and Saro Lee. 2014. "Flood Susceptibility Mapping Using Integrated Bivariate and Multivariate Statistical Models." *Environmental Earth Sciences* 72(10):4001–15. doi: 10.1007/s12665-014-3289-3.

Tehrany, Mahyat Shafapour, Biswajeet Pradhan, and Mustafa Neamah Jebur. 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. doi: 10.1016/j.jhydrol.2013.09.034.

Tehrany, Mahyat Shafapour, Biswajeet Pradhan, and Mustafa Neamah Jebur. 2015. "Flood Susceptibility Analysis and Its Verification Using a Novel Ensemble Support Vector Machine and Frequency Ratio Method." *Stochastic Environmental Research and Risk Assessment* 29(4):1149–65. doi: 10.1007/s00477-015-1021-9.

Tien Bui, Dieu, Nhat-Duc Hoang, Francisco Martínez-Álvarez, Phuong-Thao Thi Ngo, Pham Viet Hoa, Tien Dat Pham, Pijush Samui, and Romulus Costache. 2020. "A Novel Deep Learning Neural Network Approach for Predicting Flash Flood Susceptibility: A Case Study at a High Frequency Tropical Storm Area." *Science of The Total Environment* 701:134413. doi: 10.1016/j.scitotenv.2019.134413.

Tien Bui, Dieu, Khabat Khosravi, Shaojun Li, Himan Shahabi, Mahdi Panahi, Vijay Singh, Kamran Chapi, Ataollah Shirzadi, Somayeh Panahi, Wei Chen, and Baharin Bin Ahmad. 2018. "New Hybrids of ANFIS with Several Optimization Algorithms for Flood Susceptibility Modeling." *Water* 10(9):1210. doi: 10.3390/w10091210.

Towfiqul Islam, Abu Reza Md, Swapan Talukdar, Susanta Mahato, Sonali Kundu, Kutub Uddin Eibek, Quoc Bao Pham, Alban Kuriqi, and Nguyen Thi Thuy Linh. 2021. "Flood Susceptibility Modelling Using Advanced Ensemble Machine Learning Models." *Geoscience Frontiers* 12(3). doi: 10.1016/j.gsf.2020.09.006.

Townsend, Philip A., and Stephen J. Walsh. 1998. "Modeling Floodplain Inundation Using an Integrated GIS with Radar and Optical Remote Sensing." *Geomorphology* 21(3–4):295–312. doi: 10.1016/S0169-555X(97)00069-X.

Tsyganskaya, Viktoriya, Sandro Martinis, and Philip Marzahn. 2019. "Flood Monitoring in Vegetated Areas Using Multitemporal Sentinel-1 Data: Impact of Time Series Features." *Water* 11(9):1938. doi: 10.3390/w11091938.

TUCKER, C. J., and P. J. SELLERS. 1986. “Satellite Remote Sensing of Primary Production.” *International Journal of Remote Sensing* 7(11):1395–1416. doi: 10.1080/01431168608948944.

Twele, André, Wenxi Cao, Simon Plank, and Sandro Martinis. 2016. “Sentinel-1-Based Flood Mapping: A Fully Automated Processing Chain.” *International Journal of Remote Sensing* 37(13):2990–3004. doi: 10.1080/01431161.2016.1192304.

Uddin, Shahadat, Arif Khan, Md Ekramul Hossain, and Mohammad Ali Moni. 2019. “Comparing Different Supervised Machine Learning Algorithms for Disease Prediction.” *BMC Medical Informatics and Decision Making* 19(1):281. doi: 10.1186/s12911-019-1004-8.

UN, World Bank, GFDRR & EU. 2018a. “UN, World Bank, GFDRR & EU.” *UNDP*. Retrieved November 29, 2023 (https://www.undp.org/laopdr/publications/post-disaster-needs-assessment-2018-floods-lao-pdr).

UN, World Bank, GFDRR & EU. 2018b. “UN, World Bank, GFDRR & EU.” *UNDP*. Retrieved November 29, 2023 (https://www.undp.org/laopdr/publications/post-disaster-needs-assessment-2018-floods-lao-pdr).

Usama, Muhammad, Junaid Qadir, Aunn Raza, Hunain Arif, Kok-lim Alvin Yau, Yehia Elkhatib, Amir Hussain, and Ala Al-Fuqaha. 2019. “Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges.” *IEEE Access* 7:65579–615. doi: 10.1109/ACCESS.2019.2916648.

Van Der Knijff, J. M., J. Younis, and A. P. J. De Roo. 2010. “LISFLOOD: A GIS-based Distributed Model for River Basin Scale Water Balance and Flood Simulation.” *International Journal of Geographical Information Science* 24(2):189–212. doi: 10.1080/13658810802549154.

Van Houdt, Greg, Carlos Mosquera, and Gonzalo Nápoles. 2020. “A Review on the Long Short-Term Memory Model.” *Artificial Intelligence Review* 53(8):5929–55. doi: 10.1007/s10462-020-09838-1.

Vidyarthi, Vikas Kumar, Ashu Jain, and Shikha Chourasiya. 2020. “Modeling Rainfall-Runoff Process Using Artificial Neural Network with Emphasis on Parameter Sensitivity.” *Modeling Earth Systems and Environment* 6(4):2177–88. doi: 10.1007/s40808-020-00833-7.

Vilandone Keophila. 2018. “Multi-Objective Optimization for Flood Control Operation and Electricity Production of Nam Ngum 1 and 2 Hydropower Plants.” *Journal of Thai Interdisciplinary Research* 13:58. doi: 10.14456/JTIR.2018.52.

Wagenaar, Dennis, Alex Curran, Mariano Balbi, Alok Bhardwaj, Robert Soden, Emir Hartato, Gizem Mestav Sarica, Laddaporn Ruangpan, Giuseppe Molinario, and David Lallemant. 2020. “Invited Perspectives: How Machine Learning Will Change Flood Risk and Impact Assessment.” *Natural Hazards and Earth System Sciences* 20(4):1149–61. doi: 10.5194/nhess-20-1149-2020.

Wang, Xizhao, Yanxia Zhao, and Farhad Pourpanah. 2020. “Recent Advances in Deep Learning.” *International Journal of Machine Learning and Cybernetics* 11(4):747–50. doi: 10.1007/s13042-020-01096-5.

Wang, Y., Hong, H., Chen, W., Li, S., Panahi, M., Khosravi, K., ... & 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, 712-729. doi: 10.1016/j.jenvman.2019.06.102.

Wang, Yue, Deliang Sun, Haijia Wen, Hong Zhang, and Fengtai Zhang. 2020. “Comparison of Random Forest Model and Frequency Ratio Model for Landslide Susceptibility Mapping (LSM) in Yunyang County (Chongqing, China).” *International Journal of Environmental Research and Public Health* 17(12):4206. doi: 10.3390/ijerph17124206.

Wu, Nitu, Luís Guilherme Teixeira Crusiol, Guixiang Liu, Deji Wuyun, and Guodong Han. 2023. “Comparing Machine Learning Algorithms for Pixel/Object-Based Classifications of Semi-Arid Grassland in Northern China Using Multisource Medium Resolution Imageries.” *Remote Sensing* 15(3):750. doi: 10.3390/rs15030750.

Wubalem, Azemeraw, Gashaw Tesfaw, Zerihun Dawit, Belete Getahun, Tamrat Mekuria, and Muralitharan Jothimani. 2021. “Comparison of Statistical and Analytical Hierarchy Process Methods on Flood Susceptibility Mapping: In a Case Study of the Lake Tana Sub-Basin in Northwestern Ethiopia.” *Open Geosciences* 13(1):1668–88. doi: 10.1515/geo-2020-0329.

Xie, Haowen, Mark Randall, and Kwok-wing Chau. 2022. “Green Roof Hydrological Modelling With GRU and LSTM Networks.” *Water Resources Management* 36(3):1107–22. doi: 10.1007/s11269-022-03076-6.

Yao, X., L. G. Tham, and F. C. Dai. 2008. “Landslide Susceptibility Mapping Based on Support Vector Machine: A Case Study on Natural Slopes of Hong Kong, China.” *Geomorphology* 101(4):572–82. doi: 10.1016/j.geomorph.2008.02.011.

Yariyan, Peyman, Mohammadtaghi Avand, Rahim Ali Abbaspour, Ali Torabi Haghighi, Romulus Costache, Omid Ghorbanzadeh, Saeid Janizadeh, and Thomas Blaschke. 2020. “Flood Susceptibility Mapping Using an Improved Analytic Network Process with Statistical Models.” *Geomatics, Natural Hazards and Risk* 11(1):2282–2314. doi: 10.1080/19475705.2020.1836036.

Youssef, Ahmed M., Biswajeet Pradhan, and Abdallah Mohamed Hassan. 2011. “Flash Flood Risk Estimation along the St. Katherine Road, Southern Sinai, Egypt Using GIS Based Morphometry and Satellite Imagery.” *Environmental Earth Sciences* 62(3):611–23. doi: 10.1007/s12665-010-0551-1.

Youssef, Ahmed M., Biswajeet Pradhan, and Saleh A. Sefry. 2016. “Flash Flood Susceptibility Assessment in Jeddah City (Kingdom of Saudi Arabia) Using Bivariate and Multivariate Statistical Models.” *Environmental Earth Sciences* 75(1):12. doi: 10.1007/s12665-015-4830-8.

Yu, Han, Zengliang Luo, Lunche Wang, Xiangyi Ding, and Shaoqiang Wang. 2023. "Improving the Accuracy of Flood Susceptibility Prediction by Combining Machine Learning Models and the Expanded Flood Inventory Data." *Remote Sensing* 15(14):3601. doi: 10.3390/rs15143601.

Zeleňáková, Martina, Rastislav Fijko, Slavomír Labant, Erik Weiss, Gabriel Markovič, and Roland Weiss. 2019. "Flood Risk Modelling of the Slatvinec Stream in Kružlov Village, Slovakia." *Journal of Cleaner Production* 212:109–18. doi: 10.1016/j.jclepro.2018.12.008.

Zhang, Bo, Guoxiang Liu, Xiaowen Wang, Yin Fu, Qiao Liu, Bing Yu, Rui Zhang, and Zhilin Li. 2022. "Semi-Automated Mapping of Complex-Terrain Mountain Glaciers by Integrating L-Band SAR Amplitude and Interferometric Coherence." *Remote Sensing* 14(9):1993. doi: 10.3390/rs14091993.

Zhao, Gang, Bo Pang, Zongxue Xu, Dingzhi Peng, and Depeng Zuo. 2020. "Urban Flood Susceptibility Assessment Based on Convolutional Neural Networks." *Journal of Hydrology* 590:125235–125235. doi: 10.1016/j.jhydrol.2020.125235.

Zhou, Linghao, Hok Fok, Zhongtian Ma, and Qiang Chen. 2019. "Upstream Remotely-Sensed Hydrological Variables and Their Standardization for Surface Runoff Reconstruction and Estimation of the Entire Mekong River Basin." *Remote Sensing* 11(9):1064. doi: 10.3390/rs11091064.

Zhu, Kaili. 2024. "A Novel Framework for Feature Simplification and Selection in Flood Susceptibility Assessment Based on Machine Learning." *Journal of Hydrology*.

Zou, Yongsong, Jin Wang, Peng Lei, and Yi Li. 2023. "A Novel Multi-Step Ahead Forecasting Model for Flood Based on Time Residual LSTM." *Journal of Hydrology* 620:129521. doi: 10.1016/j.jhydrol.2023.129521.

Appendix

Data for: Flood susceptibility mapping leveraging open-source remote-sensing data and machine learning approaches in Nam Ngum River Basin (NNRB), Lao PDR

https://data.mendeley.com/datasets/vhyykrs5pr/1

Journal Publication:

Mangkhaseum, S., Bhattarai, Y., Duwal, S., & Hanazawa, A. (2024). "Flood susceptibility mapping leveraging open-source remote-sensing data and machine learning approaches in Nam Ngum River Basin (NNRB), Lao PDR". Geomatics, Natural Hazards and Risk, 15(1). https://doi.org/10.1080/19475705.2024.2357650