FLOOD RISK MAPPING AND ASSESSMENT IN MIDDLE AWASH SUB-BASIN, ETHIOPIA

AMARE GEBRIE GETANEH

A THESIS SUBMITTED TO
THE DEPARTMENT OF WATER RESOURCES ENGINEERING
SCHOOL OF CIVIL ENGINEERING AND ARCHITECTURE

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER'S IN WATER RESOURCES ENGINEERING (SPECIALIZATION IN HYDROLOGY AND WATER RESOURCES MANAGEMENT)

OFFICE OF GRADUATE STUDIES

ADAMA SCIENCE AND TECHNOLOGY UNIVERSITY

SEPTEMBER, 2022
ADAMA, ETHIOPIA

Flood Risk Mapping and Assessment in Middle Awash Sub-basin, Ethiopia

Amare Gebrie Getaneh

Id No.: PGR/22426/13

Advisor: Mulugeta Musie Abdi (Ph.D.)

A Thesis Submitted to the Department of Water Resources Engineering

School of Civil Engineering and Architecture

Presented in Partial Fulfillment of the Requirement for the Degree of Master's in Water Resources Engineering (Specialization in Hydrology and Water Resources Management)

Office of Graduate Studies

Adama Science and Technology University

September, 2022

Adama, Ethiopia

APPROVAL PAGE

I, the advisor of the thesis entitled "Flood Risk Mapping and Assessment in Middle Awash Sub-basin, Ethiopia" and developed by Amare Gebrie, hereby certify that the recommendation and suggestions made by the board of examiners are appropriately incorporated into the final version of the thesis.

Mulugeta Musie (Ph.D.)	____________________	____________________
Major Advisor	Signature	Date

We, the undersigned, members of the Board of Examiners of the thesis by Amare Gebrie have read and evaluated the thesis entitled "Flood Risk Mapping and Assessment in Middle Awash Sub-basin, Ethiopia" and examined the candidate during open defense. This is, therefore, to certify that the thesis is accepted for partial fulfillment of the requirement of the degree of Master of Science in Hydrology and Water Resources Management.

Keneni Alemu (Ph.D.)	____________________	____________________
Chairperson	Signature	Date
Abdulkerim Bedewi (Ph.D.)	____________________	____________________
Internal Examiner	Signature	Date
Zeleke Agide (Ph.D.)	____________________	30/10/202
External Examiner	Signature	Date

Final approval and acceptance of the thesis is contingent upon submission of its final copy to the Office of Postgraduate Studies (OPGS) through the Department Graduate Council (DGC) and School Graduate Committee (SGC).

____________________	____________________	____________________
Department Head	Signature	Date
____________________	____________________	____________________
School Dean	Signature	Date
____________________	____________________	____________________
Office of Postgraduate Studies, Dean	Signature	Date

DECLARATION

I hereby declare that this Master Thesis entitled "Flood Risk Mapping and Assessment in Middle Awash Sub-basin, Ethiopia" is my original work. That is, it has not been submitted for the award of any academic degree, diploma or certificate in any other university. All sources of materials that are used for this thesis have been duly acknowledged through citation.

Amare Gebrie	________________	________________
Name of student	Signature	Date

RECOMMENDATION

I, the advisor of this thesis, hereby certify that I have read the revised version of the thesis entitled "Flood Risk Mapping and Assessment in Middle Awash Sub-basin, Ethiopia" prepared under my guidance by Amare Gebrie submitted in partial fulfillment of the requirements for the degree of Mater's of Science in Hydrology and Water Resources Management. Therefore, I recommend the submission of revised version of the thesis to the department following the applicable procedures.

Mulugeta Musie (Ph.D.)	____________________	______________________
Major Advisor	Signature	Date

ACKNOWLEDGEMENTS

I express my most sincere thanks to my advisor, Mulugeta Musie (Ph.D.), for his unlimited guidance, support, and encouragement throughout this thesis work. I am also sincerely grateful to Dereje Birhanu (Ph.D.) for his valuable advice and support throughout this study. And I thank Million Teshome (MSc.) and Mulugeta Kebede (MSc.), the two consecutive Department head of Water Resources Engineering, for their support in fulfilling facilities for a convenient study environment.

I would like to acknowledge Adama Science and Technology University (ASTU) for offering me the academic program and funds. I would like to thank the Ministry of Water and Energy (MoWE) and Ethiopian Road Authority (ERA) for their provision of the required data. I know nothing can be done without data.

I have special thanks to all my friends and families, for their ideal encouragement and financial support. Above all, my sincere thanks go to the "Almighty God" for granting me his care and blessings all along my way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

AMS	Annual Maximum Series
AOI	Area of Interest
ArcGIS	Aeronautical Reconnaissance Coverage Geographic Information System
C_V	Coefficient of Variation
DEM	Digital Elevation Model
d/s	downstream
DPPC	Disaster Prevention and Preparedness Commission
DTM	Digital Terrain Model
FEMA	Federal Emergency Management agency
FI	Flood Intensity
GEV	Generalized Extreme Value
GLO	Generalized Logistic
GNO	Generalized Normal
GoF	Goodness of Fit
GPA	Generalized Pareto
HEC-Geo RAS	Hydrologic Engineering Center's Geographic River Analysis System
HEC-RAS	Hydraulic Engineering Center for River Analysis System
H_F	Hazard Factor
IFM	Index Flood Method
L_{CK}	Linear Coefficient of Kurtosis
L_{CS}	Linear Coefficient of Skewness
L_{CV}	Linear Coefficient of Variation
LMIF	L-moment-based index-flood method
LMRD	Linear Moments Ratio Diagram
MASB	Middle Awash Sub-basin
MCDA	Multi-Criteria Decision support Approach
n	Manning's coefficient
NDRMC	National Disaster Risk Management Commission
NDWI	Normalized Difference Water Index
NMA	National Metrological Agency
NSE	Nash-Sutcliffe model efficiency
OCHA	Office for the Coordination of Humanitarian Affairs

PBIAS	percent bias
PDS	Partial Duration Series
PE3	Pearson Type III
POT	Peak Over Threshold
PWM	Probability Weighted Method
Q_P	Peak discharge
Q_T	Extreme flow under T-year return period
R^2	coefficient of determination
R_F	Risk Factor
RFFA	Regional flood frequency analysis
RGC	Regional Growth Curve
SNNPR	Southern Nations, Nationalities and Peoples Region
T	Return period
TIN	Triangular Irregular Network
UNDP	United Nations Development Programe
USACE	United States Army Corps of Engineers
u/s	upstream
VF	Vulnerability Factor
WMO	World Meteorological Organization
X_T	Estimated flow quantiles value

ABSTRACT

In Ethiopia, river flood is a major hazard causing significant damage to lives and properties. Middle Awash Sub-basin (MASB) has experienced severe flooding by the overflowing of Awash River. The objective of this study is to generate flood hazard, vulnerability and risk maps and do their assessment on MASB with the application of HEC-RAS, ArcGIS and HEC-GeoRAS. River flow data of seven gauge stations were analyzed for estimating flood quantiles on gauged and ungauged sites by applying regional flood frequency analysis (RFFA) based on L-moment-based index flood. From five candidate distributions, Pearson type III (PE3) was identified as the best fit regional distribution for MASB and the estimated regional flood quantiles are 0.956, 1.32, 1.53, 1.73, 1.96, 2.12, 2.28, 2.48 for 2-, 5-, 10-, 20-, 50-, 100-, 200-, and 500-year floods respectively. Based on the regional quantiles and the index flood, at-site quantiles are estimated and inputted to HEC-RAS as profiles for the generation of inundation extent, depth and flow velocity. HEC-RAS was calibrated using flow and water level data of Awash 7 Kilo station and it performed very good (NSE = 0.70), good (PBIAS = -13.80) and very good (R^2 = 0.99) with adjusted n value of 0.042 for the main river channel. Comparing historical flood extent from satellite image and HEC-RAS simulation is used in model validation and resulted overlapping percentage and the F statistic values of 94.74 and 86.67 respectively. Flood risk for each land unit is computed as a product of hazard and vulnerability. A considerable increase in inundated area is found with an increase in return periods. The majority floodplain area of MASB is under medium and high hazard class. Out of the 80 land units considered, 25 land units scored a vulnerability factor of less than 1 (48.98%) and the remaining 26 and 29 land units scored 1 to 3 (27.77%) and above 3 (23.26%) respectively. High and sever risks are observed on small areas of the basin covering a total percentage of 10.48 to 12.41% for 2- to 500-year flood. Towns such as Haro Adi, Melka Worer, Bole, Doni, Awash Melkasa, and Metehara have experienced high and sever flood risk for the different return periods. The study addressed that flood risk of the land units can be varied even after there is no change in flood hazard intensity. The study can provide supportive information in identifying priority areas in the flood risk zones of MASB which is particularly useful when there is a budget and time constraint. It can be also a source of baseline information for future studies.

Keywords: MASB, HEC-RAS, GIS/HEC-GeoRAS, regional flood frequency, Index flood, Hazard, Vulnerability, Risk

1. INTRODUCTION

1.1. Background

Flooding is a natural process happens when the natural or man-made drainage system cannot cope with the volume of water created by rainfall (Gezahegn & Suryabhagavan, 2018). It is a costly natural disaster causing fatalities, and damages to life, property, communications, transportation, and other infrastructures. It can happen anywhere and its effects are clearer in developing countries due to a lack of economic resources, poor housing facilities, inadequate warning systems, and preparedness (El-Naqa & Jaber, 2018). Flood accounts for 40% of all natural disasters worldwide (Abaya et al., 2009). It represented a major natural hazard in Africa, causing over 27, 000 fatalities between 1950 and 2019 (Tramblay et al., 2020). Based on the number of people affected, over the past 30 years, flood and drought are the two natural hazards that have the largest humanitarian impacts in Africa (Lumbroso, 2020). However, in the past decade, across Africa, flood has overtaken drought in terms of the number of people that they impact. It usually affects large river basins like the Congo, Niger, Nile, and Zambezi basins (Brook & Boneya, 2020).

Ethiopia had extensive experiences in both flash and riverine flood events (Mamo et al., 2019). Its frequency and magnitude had increased rapidly in the last half-century (Sinshaw et al., 2018). According to Disaster Prevention and Preparedness Commission (DPPC) and Federal Democratic Republic of Ethiopia National Disaster Risk Management Commission (NDRMC), the major river flood-prone areas in Ethiopia are parts of the Oromia and Afar regions lying along the upper, middle, and down-stream plains of the Awash River; parts of the Somali region along the Wabe Shebelle, Genale, and Dawa Rivers; low-lying areas of Gambella along the Baro, Gilo, Alwero and Akobo Rivers; down-stream areas along the Omo and Bilate Rivers in Southern nations, Nationalities and Peoples Region (SNNPR) and floodplains surrounding Lake Tana and the banks of Gumera, Rib and Megech Rivers in Amhara (Brook & Boneya, 2020; DPPC, 2006; NDRMC, 2018). The most recent figures indicate that nearly about 1,017,854 people are affected by flooding and 292,863 people are displaced across the country. As many as 447,565 people are affected in Oromia, 144,490 in Amhara, 140,892 in Somali, and 90,121 in SNNPR (NDRMC, 2020).

The most highly developed river basin with improved economic infrastructure is the Awash River Basin. Close to 70% of the country's large-scale irrigated agriculture is located along the Awash River. Therefore, high economic damage is expected to occur during flooding along this river (Achamyeleh, 2003). Lower and middle basins of the Awash River are highly susceptible to flooding and are frequently flooded by the overflowing of the river (Haile, 2009). The performance of Amibara irrigation scheme was declining from time to time and users complained about its management. Many farms and investors had left command areas due to surface and groundwater inundation in the vicinity (Wondafrash, 2021). The Wonji area of the Adama woreda is also under flood hazard. By 23 August 1996, Wonji of the Adama *woreda* not only the population but also a large state-owned sugar plantation was put at Awash River flood risk. Further downstream, critical locations include other plantations in the area of Merti and south of Metahara (Ahrens, 1996). According to United Nations Development programme (UNDP) Emergencies Unit for Ethiopia, Melka Werer Agricultural Research Centre in Melka Werer and the Melka Sedi State Farm has also been damaged by the floods (UNDP, 1999). Zone 3 of the Afar region including districts of Amibara, Dulecha and Awash Fentale of the MARB are under flooding risk (NDRMC, 2018). The United Nations Office for the Coordination of Humanitarian Affairs (OCHA) reported that about 600,000 citizens are affected by flood of the Awash River. Out of this, 25,000 people have been displaced in Fentale *woreda* and 14,000 ha of large farms in Amibara and Nahurka *woredas* inflicted huge damages (OCHA, 2020).

Unlike most types of disasters like volcanoes and earthquakes, flood is preventable and manageable through proper implementation of structural and non-structural measures. Flood risk mapping and assessment is one of the non-structural measures to combat flooding. This study conducted flood risk mapping and assessment on MASB by integrating the hydraulic model (HEC-RAS), an ArcGIS extension of HEC-GeoRAS and RS products.

1.2. Statement of the Problem

Awash River basin has been a subject of large-scale flooding for several years mainly due to heavy rains and inadequate water resource management (Tessema et al., 2015). This leads overflow of the Awash River over its banks and inundate the surrounding causing thousands of households to displace from their home, loss of thousands of livestock, destruction of infrastructures and agricultural command areas. This is aggravated by the flow coming from the tributaries of Kessem and Kebena, and high flood spills from Koka Reservoir (Achamyeleh, 2003). Therefore, the flood problem in this area is a critical issue that needs detailed assessments on the proper design of structures, flood forecasting, and flood risk assessment for the control and minimization of river flood damages.

The negative impacts of floods can be mitigated through applying structural and non-structural approaches that can be applied as either proactive or reactive measures. Structural measures are through the construction of flood preventive structures such as reservoirs, embankments, flood levies, drainage channels, etc. Non-structural measures are measures not involving physical construction which use knowledge, practice or agreement to reduce risks including floodplain zoning and management, floodplain regulation, flood forecasting and warning, evacuation and shelter management, flood proofing, land use planning laws, public awareness programs, a system of flood risk assessment, flood-related databases, etc. (Kundzewicz, 2002). It is becoming critical for governments to be proactive in reducing flood risk rather than reactive by providing post-disaster response and recovery measures (Abebe, 2010). The first step towards a proactive model is to perform a flood risk assessment.

There are different studies conducted on Awash River basin, more on upper sub-basin, regarding flood hazard and risk mapping and their assessment using different approaches, as De Risi et al., (2015) based on Topographic Water Index (TWI), Wondim, (2016) using multi criteria decision analysis, Getahun & Gebre, (2015) using GIS and HEC-RAS, Namara et al., (2021) using HEC-HMS and HEC-RAS with GIS etc. But less studies are conducted on MASB. Up to this search, three MSc. thesis works are obtained namely, Gadise, (2020), Nama & Tesfaye, (2020) and Haile, (2009). The study by Gadise, (2020) has computed the inundation extent but, hazard and risk maps are not made while Nama & Tesfaye, (2020) and Haile, (2009) studied only a small portion of the study area. Thus, we can generally say that MASB is not well studied for flood hazard and risk mapping and assessment.

This study tried to generate flood hazard, vulnerability and risk maps and do an assessment of them on whole length of Awash River in MASB, considering estimated flows of the different return periods using HEC-RAS and GIS. In spite of the abundance of water resources in Ethiopia, there is a limited availability of hydrometeorological data that hinder research and development in the country (Tessema et al., 2015). There is clearly a need for implementing methods in this region that can estimate more reliable flood quantile under such limited data conditions. Regional flood frequency analysis (RFFA) is one of the frequency analysis techniques that is widely used for analyzing stream flow data and estimating flow quantiles on gauged and ungauged sites. Here, RFFA based on L-moments and index flood is used (Hosking & Wallis, 1997). This method has not been widely applied in the Middle Awash Sub-basin.

A place with high flooding hazard but not under flood risk is decided based on vulnerability of the area. This is because, risk is a function of hazard and vulnerability. Some studies consider only inundation depth as a hazard factor (Kim et al., 2020; Tingsanchali & Karim, 2010). But, factors like duration of flooding, flow velocity, timing of occurrence, rate of rise of food, etc. affects hazard intensity of a flood. This study incorporates inundation depth and flow velocity, considering the topography of the study area and available resource, as used by (Elkhrachy et al., 2021; Vojtek & Vojteková, 2016). Vulnerability can be computed by considering different assets as elements at risk, broadly classified as physical, economic, social and environmental vulnerabilities (Ali et al., 2016). Here, vulnerability of the different landuse types is considered. Land unit-based risk assessment, considering Kebele and town administrative boundaries as a land unit, is conducted as it is easy to understand and implement the results. To enhance the accuracy of our work, some structures along the river and floodplain such as blocked obstructions and bridges are incorporated during the execution of the hydraulic model. In addition, higher resolution remote sensing products such as digital elevation model (DEM) of 12.5m resolution, landuse/cover map of 10m resolution and Landsat image were used.

We should realize that no protection against floods is absolute. But an increasing role can be added by applying non-structural measures such as flood risk mapping and assessment. Flood hazard and risk maps are easy to read, rapidly accessible, which can provide significant input for planning and management of flood risk in the study area. Such a need can be aided with a proper application of hydraulic modeling integrated with GIS software and RS products.

1.3. Objective of the Study

1.3.1. General objectives

The general objective of this study is to prepare a multi-return period flood risk map and assess the flood risk on MASB along the Awash River floodplain area using Hydraulic Model (HEC-RAS), HEC-GeoRAS, GIS and RS products.

1.3.2. Specific objectives

In connection with the above general objective, the following specific objectives are outlined:

- To perform hydrological data quality check;
- To estimate multi-return period flood quantiles;
- To build a hydrodynamic model for MASB and simulate flow of estimated quantiles;
- To generate flood hazard, vulnerability and risk maps and do assessment.

1.4. Research Questions

- How good are the quality of collected hydrological data?
- What are the magnitudes of flood quantiles for different return periods?
- How satisfactory is the HEC-RAS model for simulating floods on MASB?
- Which parts of the MASB suffer from more flood hazards and which area is highly vulnerable to flood risk?

1.5. Significance of the Study

Flood mitigation measures, which include structural as well as non-structural measures, can be selected effectively once the risk assessment for the area has been done accurately (Malakeel et al., 2021). Flood risk mapping facilitates the administrators and planners to identify areas vulnerable to food hazards and determine infrastructure at risk and the degree they might be affected and map their capacity to respond and recover. Most importantly, it helps in identifying and prioritizing the mitigation and response efforts and helps to inform emergency responses (Baky et al., 2019). It also makes land use planning decisions in floodplains be informed for choices to incorporate scientifically derived information in their decision-making process.

This study has performed flood hazard and risk mapping and assessment for MASB which can be applied for the benefit of the local community, government agencies and researchers. It can help in the planning of more effective emergency response (early warning application) for flood hazard areas such as Amibara, Adama, Boset, Fentale, Awash and Dulecha *woredas* during high flood events on upstream. It enables people to protect some property and infrastructure before flood hazard. As a developing country with budget constraints, it is useful to identify priority areas in the risk zone for mitigation action. The land unit-based flood risk assessment enables to identify the level of risk on each *kebeles*. In addition, the study can be used as a point of reference for policymakers and a source of baseline information for those researchers who intend to investigate further study on MASB.

1.6. Scope of the Study

The scope of the study is delimited to both geographical and thematic scopes. Although seasonal flooding has been affecting the upper, middle and lower sub-basins of Awash Sub-basin, the scope of this study is delimited to the geographical coverage of MASB due to time and resource limitations. Regarding the thematic scope, the study focused on regional flood frequency analysis (RFFA) based on the L-moment-based index flood method and mapping and assessment of the flood hazard and risk for the different return periods. The flood damage assessment is not conducted because of the requirement of intensive data and budget allocation for on-site data survey. For the same reason, a field survey for the input geometric data, manning's roughness determination and the dimension of the structures on the floodplain were not conducted, instead, these data were extracted from available remote sensing (RS) products and secondary data sources. The risk mapping and assessment were conducted through the integration of HEC-RAS, HEC-GeoRAS and ArcGIS software with RS products.

1.7. Thesis Organization

This thesis study is organized into five chapters. The first chapter is the introduction which consist the background, problem of the statement, objective of the study, research questions, significance of the study and scope of the study. The second chapter is the literature review which reviews articles and other necessary documents conducted on various thematic areas to fully address the topic of the study and easily understand it. The review includes a description on different tittles and terminologies of floods, flood hazard, vulnerability and risk mapping and their assessment, flood magnitude estimation methods, flood frequency analysis, probability distributions, regionalization and the index flood method, hydraulic models and flood modelling, and finally reviews studies on flood risk mapping and assessment. Materials and methods in the third chapter discussed the methodology from the data collection phase to the flood risk mapping and assessment phase. In middle of the methodology, details of preparation phase and execution phase have also been presented including description of the study area, statistical tests, regional flood frequency analysis, HEC-RAS/HEC-GeoRAS execution, and flood hazard, vulnerability and risk mapping and assessment. The fourth chapter describes the results of the study and a discussion made on them. Finally, the fiveth chapter draws conclusions from analysis of results and has made recommendations based on the analysis from this study.

2. LITERATURE REVIEW

2.1. Definition, Type and Causes of Flood

2.1.1. Floods and floodplain

Flood is an area of study in the discipline of hydrology. It can be defined as a water flow with high discharge that rises and drowns out the surrounding area that is normally dry (El-Naqa & Jaber, 2018). Floodplains are relatively flat alluvial geographic areas located adjacent to rivers and streams that carry large volumes of water during high flow events. When water enters a floodplain, it can pose a safety and personal property risk to individuals who reside in and structures that are located in the floodplain (Alzahrani, 2017; Benito & Hudson, 2010). Floodplains are often preferred place for human use and settling, and can provide amenities, risks and rewards. During flooding, there is a diminution of natural resources and their functions. In the long term, the residents on the floodplain will be the loser, meaning that the risk will exceed all the advantages obtained from settling on the floodplain (Pojani & Keçi, 2016).

2.1.2. Types of floods

Before a hazard and risk assessment is carried out, it is necessary to determine which types of floods are dominant and catastrophic in the catchment area, because in practice the selection of hazard and risk assessment methods applied varies depending on the type of flood (Wright, 2015). Several different types of floods are categorized based on their origin of the water (source), the geography of the receiving area, cause and speed of onset (Pojani & Keçi, 2016). The types of floods include river (fluvial) flood, pluvial flood (surface/ overland) flood, coastal (surge) flood, flash flood and ice jam flooding.

River flood is one of the most common types of flooding, it usually occurs when excessive rainfall over an extended period of time causes a river to exceed its capacity (Tazin, 2018). Owing to the long-time scale of the rising waters, river floods pose a lower risk of fatalities; people have more time to take proper actions (Brook & Boneya, 2020). Pluvial floods occur when surface runoff generation exceeds infiltration rates and drainage capacity, often during high-intensity short-duration storm rainfall events (Miller & Hutchins, 2017). Coastal flooding occurs in areas that lie on the coast of a sea, ocean, or other large body of open water. It is caused by higher-than-average high tide and worsened by heavy rainfall and onshore winds (i.e., wind blowing landward from the ocean) (Tazin, 2018). Flash flooding is characterized by an intense, high-velocity torrent of water that occurs in an existing river channel with little to no notice caused by very heavy rain when water can't be absorbed quickly enough (Davis, 2001; Tazin, 2018). Ice jam flooding is similar to flash flooding in that the formation of a jam results in a rapid rise of water both at the point of the jam and upstream. Failure of the jam results in sudden flooding downstream.

2.1.3. Causes of flooding and their impacts

Flooding is experienced all over the globe and for a variety of reasons - but why exactly does flooding occur? Floods are caused by extreme hydrometeorological actions while their evolution depends on geomorphologic agents, such as permeability and soil stability, vegetation cover, and the geometric characteristics of the river basins (Khattak et al., 2016). Factors that cause floods include changes in land use, waste management, erosion and sedimentation, slums along rivers, improper flood control systems, high rainfall, river physiography, inadequate river capacity, effects of high tides, land subsidence, water structures, and damage to flood control structures.

Water overflow is usually not a problem if it does not cause losses, deaths, or injuries, and does not remain in settlements for a long time or cause other problems in daily life. But on the contrary, if water is pooled with a high enough, and occurs in a long time of course this can complicate human activity (Sholihah et al., 2020). Floods can have significant consequences including human and animal life losses, agricultural crops destruction and soil loss, damages of infrastructures, communication networks, and transport of sediment loads and pollutants (Zoka et al., 2018).

2.2. Flood Hazard, Vulnerability and Risk Assessments

2.2.1. Flood hazard and assessment

Flood hazard is defined as the probability of potentially damaging flood situations in a given area and within a specified time horizon (Prinos, 2008). Flood hazard assessment aims to determine the probability of flood exceeding a particular intensity over an extended period of time. It requires determination of flood characteristics and the probability of occurrence of each flood. A flood hazard map shows flood characteristics with a given return period. The flood characteristics include the flood frequency (the probability of an occurrence in a specific time period), extent (the geographical area of impact), severity (intensity, the water level or water depths), flow velocity and/or direction of flow and duration of inundation (indicating the period that an area was or may be under water) (Benito & Hudson, 2010; Maskong et al., 2019; Prinos, 2008). The map delineates the flood hazard areas and creates easily-read, rapidly accessible charts and maps, which facilitates the administers and planners to identify areas of risk and prioritize their mitigation or response efforts(Chakraborty & Biswas, 2020). It can be used to ensure preparedness, response and recovery efficiently undertaken to mitigate the impact of flooding (Ekeu-wei & Blackburn, 2018). A flood hazard map can be generated from a variety of tools such as vertical aerial photographs (Elkhrachy et al., 2021), a remote sensing and GIS based flood index (Gan et al., 2012; Rahman & Thakur, 2018), field surveying, analytic hierarchy process (Ogato et al., 2020; Rahmati et al., 2016) and GIS with hydraulic simulation tools (Erena et al., 2018; Kim et al., 2020).

2.2.2. Flood vulnerability and assessment

Vulnerability is the degree to which system is susceptible to the adverse effects (Antony et al., 2021). Flood hazard is the physical and statistical aspects of the actual flooding (e.g. return period of the flood, extent and depth of inundation and flow velocity); whereas, flood vulnerability is the exposure of people and assets to floods and the susceptibility of the elements at risk to suffer from flood damage (Vojtek & Vojteková, 2016). For large floodplains, it may be preferable to model vulnerability at the scale of larger administrative units (Liu et al., 2021; Tingsanchali & Karim, 2005), census tracts (Armenakis et al., 2017), or using existing land-use maps (Banstola & Sapkota, 2019; Wright, 2015).

For assessing vulnerability, at first the approach should address the elements at risk which can be classified under different vulnerability classes such as physical vulnerability, economic vulnerability, social vulnerability and environmental vulnerability. According to Ali et al., (2016), physical vulnerability relates to buildings, infrastructure and agriculture. Economic vulnerability assesses the risk of hazard-causing losses to economic assets and processes. Socially, floods can cause death directly via drowning, physical trauma, or secondary effects such as the failure of water and sanitation services, the spread of waterborne diseases and decreased nutrition (Wright, 2015). Social vulnerable groups include women, mentally and physically handicapped persons, children, and elderly persons, the poor people, refugees, and livestock (Ali et al., 2016).

2.2.3. Flood risk and assessment

Ologunorisa & Abawua, (2005) defined risk as future potential condition, a function of the magnitude of the natural hazard and of the vulnerability of all the exposed elements in a determined moment. Risk assessment is a methodology to determine the nature and extent of risk by analyzing potential hazards and evaluating existing conditions of vulnerability that could pose a potential threat or harm to people, property, livelihoods and the environment on which they depend (Vojtek & Vojteková, 2016). Since risk cannot be completely eliminated, the only option is to manage it. Risk assessment is the first step in risk management (Ologunorisa & Abawua, 2005). It comprises of three distinct steps: hazard assessment, vulnerability assessment and finally risk assessment as risk is a function of hazard and vulnerability. Flood risk of a particular land unit is the expected level of damage to properties and loss of human lives of that area (Tingsanchali & Karim, 2005).

Flood risk analyses enable agents to evaluate the cost-effectiveness of the flood control projects, and provide risk information to residents, insurance companies and municipalities to prepare for disasters (Shah et al., 2018).

2.3. Flood Magnitude Estimation Methods

A common task in hydrology is the estimation of the design flood, i.e., a value of river discharge corresponding to a given exceedance probability that is often expressed as a return period in years (Okoli, 2019).This is crucial for the design and operation of flood control structures (dams, retaining basins) infrastructure objects (flood defenses, bridges, roads and dams), as well for flood risk management, planning, flood risk mapping and improvement of alert methods within a region (Leščešen & Dolinaj, 2019). For example, according to Hailegeorgis & Alfredsen, (2017) review, in Norway, a 200-year flood is used for flood hazard mapping for roads and railroads, and a 500-year, a 1000-year and probable maximum floods are used for dam safety analysis and in France, hazard mapping typically uses a 100-year return period, while some civil engineering structures (large dams and nuclear power plants) may require 10^3–10^4 target return periods. According to Tingsanchali & Karim, (2010), extreme events ranging from 20- to 400-year return periods are considered for planning and designing of water resource systems based on purpose and type of structure.

There are three basic approaches to the estimation of design flood: hydro-meteorological (rainfall-runoff) approach, statistical (probabilistic) approach and use of empirical formulae (Husain, 2017). Statistical methods, generally referred to as “flood frequency analysis”, requires the fitting of a probability distribution function to a record of annual maximum flows (Okoli, 2019). To use the statistical approach, a long records of discharge measurements must be available. The statistical calculations are based on observed historical data. The analyst fits several statistical distributions and selects a distribution that better fits the data (Pojani & Keçi, 2016). The studies by Khaing et al., (2021) and Baky et al., (2019) used statistical methods for design flood prediction.

In many cases there are insufficient data for discharge frequency analysis and sometimes there are missing at all. In these cases, an indirect hydro-meteorological approach known as rainfall-runoff modeling is used (Pojani & Keçi, 2016). Design flood estimation by rainfall-runoff approach mainly involves the estimation of probable maximum flood by applying unit hydrograph theory. While using statistical frequency analysis, the peak discharge that passes a certain location during a flood is assessed. But, using rainfall-runoff models, not only the peak discharge but also a design hydrograph over a period of time is obtained as well (Pojani & Keçi, 2016). One of the most widely used methods for determining the amount of runoff is the Natural Resource Conservation Service

(NRCS) curve number (CN) method. The studies by Abdulrazzak et al., (2019) and Liu et al., (2021) used rainfall-runoff modeling achieved using the HEC-HMS software program to get design hydrographs for different return periods.

In absence of the real field discharge data, empirical methods can be used to estimate flood discharge. The calculation uses various empirical formulae involving area of the catchment and some coefficients depending upon the location of the catchment. For example, Tamang & Tamrakar, (2017) used DHM/WECS (Nepal) formula and Modified Dickens method. The available formulae such as Dicken's formula, Ryves formula and Inglis formula are applicable, respectively, to Northern India, Western India and Southern India (Jena & Nath, 2020). Jena & Nath, (2020) developed an empirical formula for design flood estimation of ungauged catchments in Brahmani Basin, Odisha. Banstola & Sapkota, (2019) also used empirical methods to calculate high flow discharge for return periods of 2-, 5-, 10-, 50-, 100-, and 200-years in Gorkha, Nepal. The major limitation of applying the empirical formula is that they cannot be applied universally (Malakeel et al., 2021).

2.4. Statistical tests

When conducting a flood frequency analysis, an initial step is to undertake basic analysis of the peak flow time series to check for obvious errors and to check that the data conform to the assumptions used in the frequency analysis (England et al., 2019). A stationary stochastic process is based on two assumptions, namely the independence (without serial correlation and trends) and identical distribution (being from the same population). i.e., iid of the time series (Hesarkazzazi et al., 2021). Independence shows absence of self-correlation and identically distributed implies the homogeneity, stationarity and absence of outsiders of the time series (Boucefiane & Meddi, 2019).

A statistical test is a way to evaluate the evidence that the data provides against a hypothesis. This hypothesis is called the null hypothesis and is often referred to as H_0. H_0 is usually opposed to a hypothesis called the alternative hypothesis, referred to as H1 or Ha. If the data does not provide enough evidence against H_0, H_0 is not rejected. A threshold value that specifies H_0 should not be rejected is known as significance level (alpha, α) that lies between 0 and 1. Level of significance is probability of rejecting a hypothesis when it is in fact true. At a "10-percent" level of significance, the probability is 1/10 (England Jr et al., 2019). Very often, alpha is set to 1%, 5% and 10% of significance level or 99%, 95% and 90% confidence level respectively, in statistical tests (Hailegeorgis & Alfredsen, 2017; Rezende de Souzaa et al., 2021).

The statistical test produces a number called p-value (that is also bounded between 0 and 1). The p-value measures the believability of the null hypothesis; the smaller the p-value, the less likely the observed test statistic is when H_0 is true, and, therefore, the stronger the evidence for rejection of the null hypothesis. The p-value is also referred to as the "attained significance level" (Helsel et al., 2020). More practically, the p-value should be compared to alpha:

- If p-value < α, we reject H_0 and accept Ha with a risk proportional to p-value of being wrong.
- If p-value > α, we do not reject H_0, but this does not necessarily imply that we should accept it. It either means that H_0 is true, or that H_0 is false but our experiment and statistical test were not "strong" enough to lead to a p-value lower than alpha.

2.5. Flood Frequency Analysis

Flood frequency analysis is a statistical method used by hydrologists to estimate flow values corresponding to specific return periods or probabilities which requires the fitting of a probability distribution function to a record of annual maximum flows (Okoli, 2019). It is essential to provide hazard probability of flood risk assessment as well as to determine design flood for sites along the river (Thoummalangsy et al., 2019). Furthermore, the catchment characteristics, water availability and possible extreme hydrological conditions like floods and droughts at a river system may be illustrated through frequency analysis (Romali & Yusop, 2017).

2.5.1. Data series models

The study of extreme hydrologic events involves the selection of a sequence of the largest or smallest observations from sets of data. These data models can be extracted from either from mean daily discharge records (Leščešen & Dolinaj, 2019) or instantaneous discharge records (Jingyi & Hall, 2004). There are three types of flow data series models i.e., annual maximum series (AMS) model, partial duration series (PDS), also called as peaks over the threshold (POT) model, and complete duration series (TS) model. Of these, flood frequency analysis uses either AMS or PDS (Leščešen & Dolinaj, 2019). The AMS is based on the instantaneous maximum flood peak for each year while PDS is obtained by taking all flood peaks equal to or greater than a predefined base flood. Thus, an n-year record can produce m peaks with m > n (England et al., 2019).

If more than one flood per year must be considered, a partial duration series may be appropriate. The PDS models, however, are limited because observations may not be independent, which colors the assumption of independence of flood peaks for statistical analysis (Gebregiorgis et al., 2013). No specific guidelines are recommended for defining flood events to be included in a partial series.

For example, Desalegn & Mulu, (2021) and Getahun & Gebre, (2015) used 5% highest data from the gauging stations. Flood frequency analysis by using AMS of streamflow records are required to reduce the uncertainty due to time resolution of the sampling (Hailegeorgis & Alfredsen, 2017). Because, fluctuations of streamflow with in the day due to diurnal variations and effects of runoff delay would result in the instantaneous flood peaks, which are higher than their daily average values. However, the use of an AM flow series may involve some loss of information. For example, the second or third peak within a year may be greater than the maximum flow in other years, and yet they are ignored (Rao & Hamed, 2000).

2.5.2. Regionalization and the Index Flood Method (IFM)

In the context of flood frequency analysis, regionalization refers to grouping of basins into homogeneous regions and selection of appropriate frequency distributions for the identified regions (Kachroo et al., 2000). A homogeneous region is an area consisting sites with the same standardized frequency distributional form and parameters (Sine & Ayalew, 2004). Since observed flood data are rarely available at the specific location, flood estimation in ungauged or poorly gauged basins is a common problem in engineering hydrology (Lee & Kim, 2019). The advantages of working with a homogeneous region is that the historical data available within the region can be pooled to get an efficient estimate of parameters of a chosen distribution and hence a more robust quantile estimate. If a homogeneous region is geographically continuous then it has an added advantage. The distributional form and parameters of the region, chosen on the analysis of the gauged data, can be used for flood frequency estimation at ungauged sites within the region (Kachroo et al., 2000).

RFFA is based on the idea that if the frequencies of events of a certain size in different river basins are similar, then a more accurate conclusion on their distribution can be achieved by the common analysis of data from basins with similar hydrological reaction than data from one river basin (Garmdareh et al., 2018; Vojtek & Vojteková, 2016). Hosking & Wallis (1997) demonstrated that even with the presence of some degree of heterogeneity, misspecification of distribution and inter site dependence; the regional estimation yields estimated quantiles that are considerably more accurate than the at-site quantile estimates especially for higher return periods and it can give more reliable estimation of distributions than would be possible using only a single site data. Because extrapolation of the at-site short records for estimation of quantiles for longer return periods may provide unreliable results (Hailegeorgis & Alfredsen, 2017).

RFFA needs development of regional flood prediction equations using commonly used techniques such as the rational method, index flood method (IFM) and quantile regression technique

(Garmdareh et al., 2018). The statistical prediction of design floods by regionalization was pioneered by Dalrymple (1960) who established the conceptual basis of the IFM (Bocchiola et al., 2003). It is based on the main hypothesis that the frequency distributions of sites within a homogeneous region are identical apart from a site-specific scaling factor, the index flood (Lee & Kim, 2019). The IFM is commonly used to develop regional flood frequency models for ungauged sites or gauged sites where hydrologic information is not sufficient (Saf, 2009). The IFM based on L-moments proposed by Hosking & Wallis, (1997) has been increasingly used in many countries (Eregno, 2014; Gebregiorgis et al., 2013; Lee & Kim, 2019; Rezende de Souzaa et al., 2021). L-moment based IFM allows for a diversity of statistical distributions being fitted to the observed data unlike Bulletin 17B (Interagency Advisory Committee on Water Data, 1982) procedure, recommended procedure by U.S. Water Resources Council, which derive quantiles from a single LP3 distribution (Lim & Voeller, 2009).

Bocchiola et al. (2003) has reviewed different methodologies used for the estimation of index flood that broadly classified into two as direct and indirect methods. The direct method involves estimating the index flood from observed (recorded) flood peaks at the station of interest such as mean of AMS, mean of the PDS of flood peaks and median of AMS, while the indirect methods involve estimation of the index flood from other hydrological, meteorological, and geomorphological features based on the different approaches.

2.6. Statistical Probability Distributions and Parameter Estimation

2.6.1. Probability distributions

Flood records describe a succession of natural events that do not fit any one specific known statistical distribution (England et al., 2019). Probability distributions are mathematical functions that describes the probability of different possible values of a variable. For flood frequency analysis, different candidate distribution can be used. According to Cunnane (1989), Normal, lognormal, three-parameter lognormal, Extreme Value Type I and II, Log-Gumbel, Pearson Type III (PE3), Gamma, Exponential, Weibull, 2 Parameter Log-normal, 3 Parameter Log-normal, 2 Component Extreme Value, Wakeby, Log Pearson Type III (LP3), Log-logistic and Generalized logistic can be used for flood frequency analysis. The most frequently used or recommended frequency distribution types for extremes of both precipitation and floods are the Extreme Value Type II and the log-normal distributions, and PE3 and the LP3 distributions are mainly for floods (Cunnane, 1989). The selection of appropriate probability distribution and associated parameter procedure is important in flood frequency analysis to avoid under- or over-estimation of design floods (Romali & Yusop, 2017).

There is no common agreement on a preferred technique of model choice and no single distribution is universally accepted (Gebregiorgis et al., 2013). Which distribution will be selected and used for FFA in some countries is defined by official guidelines. For example, in U.S., the LP3 distribution was recommended by US Water Resources Council (England et al., 2019), also known as Bulletin 17C. Hosking and Wallis, (1997), proposed a five-parameter Wakeby distribution as a default regional distribution if none of the considered candidate distributions fulfills the requirements of goodness of-fit statistics. Five distributions with three parameters, namely Generalized Logistic (GLO), Generalized extreme-value (GEV), Generalized normal (GNO), Pearson type III (PE3) and Generalized pareto (GPA) are usually used as a candidate distribution in Hosking & Wallis (1997). Distributions with their cumulative distribution functions is attached in Appendix A.

2.6.2. Parameter estimation methods

Fitting distributions can be accomplished by estimating the distribution's parameters. The parameters include shape, location, and scale etc., A number of methods can be used for parameter estimation. Commonly used methods include method of moments (MoM), maximum likelihood (ML), probability weighted moments (PWM) and methods of L-Moments (LM) (Cunnane, 1989; Leščešen & Dolinaj, 2019).

The MoM selects values for the parameters of the probability density function so that its moments are equal to those of the sample data. The first three sample moments are used to estimate the parameters (Stedinger, 2000). These include the mean (μ), standard deviation (σ), and skewness coefficient (γ). Although MoM is easy to apply, it does not use all the sample information in an exhaustive manner and is not as efficient as ML estimation, especially in three parameter distributions (Cunnane, 1989). ML method is based on the reason that the best value of a parameter of a probability distribution should be that value which maximizes the likelihood or joint probability of occurrence of the observed sample (Chow et al., 1988). It is most efficient that the sampling variance of the estimated parameters and hence of Q_T is asymptotically smaller than by any other estimation method (Cunnane, 1989). In general, ML estimators are best when fitting a distribution with large samples. However, with flood records of 50 years or less, and when fitting distributions that do not exactly match the data, the choice is generally less clear (Stedinger, 2000). Parameter estimation based on PWMs is of potential interest for distributions which may be expressed in inverse form, i.e. if X is a random variable with cumulative distribution function F(x) then, X may be written as a function of F as X = X(F) (Hooda et al., 2018). PWMs are easy to apply, almost as efficient as ML and can be easily used in regional estimation schemes (Cunnane, 1989). L-moments are dimensionless PWM ratios or linear combinations of PWMs and can be directly interpreted as measures of scale and shape of probability distributions. The location, scale, and shape estimators of the L-moment are unbiased irrespective of the probability distribution (Lee & Kim, 2019). According to (Cunnane, 1989; Noto & La Loggia, 2009):

- Compared to conventional moments, L-moments can characterize a wide range of distributions;
- Sample estimates of L-moments are so robust that they are not affected by the presence of outlier in the data set;
- L-moments yield more accurate estimates of the parameters of a fitted distribution, even sometime parameter estimated from samples are more accurate than maximum likelihood;
- L-moment ratio estimators such as L-CV, L-skewness, and L-kurtosis can exhibit lower bias than conventional product moment ratios, especially for highly skewed samples;
- with small and moderate samples, method of L-moments is often more efficient than ML.

2.7. GIS and RS in Flood Studies

Remote sensing (RS), GIS and hydraulic modelling techniques have emerged as a powerful tool to deal with various aspects of flood management in prevention, preparedness and relief management of flood disaster. GIS consists a computer environment that joins graphical elements (points, lines, polygons) with associated tabular attribute descriptions. It provide various functions that motivated the integration of GIS and hydraulic models that included pre-processing of spatial depth, creation of digital terrain models, calculation of watershed characteristics, image overlaying, network analysis, database management and visualization of results (Selvanathan & Dymond, 2010). GIS is ideally suited for various floodplain management activities such as base mapping, topographic mapping, and post-disaster verification of mapped floodplain extents and depths (Khanna et al., 2005). GIS also enables extracting topographically correct cross-section data from a DTM (Anderson & Maidment, 2000).

Nowadays, RS has been introduced as a key solution in which the data collection process is performed faster and cost-effectively without human presence in the area. Satellite-based RS images have been used to map the extent of flood inundation since the early 1970 (Bhatt et al., 2014). The possibility of providing multi-source, charge-free data and generating various attributes/parameters for efficient representation of flood hazard zones, makes remote sensing an attractive approach for low-cost flood hazard mapping (Parsian et al., 2021). Satellite images can also be used for hydraulic model calibration and validation and accuracy assessment of generated flood hazard maps (Khattak et al., 2016; Parsian et al., 2021). RS satellite imagery support the validation of flood risk maps due to their high spatial coverage and can thus complement the field surveys that are undertaken following a flood (Quirós & Gagnon, 2020). Water bodies can be delineated from Landsat images using different indexes. The water body extent can be detected using Water Ratio Index (WRI), Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), supervised classification and wetness component of K-T transformation can be used. NDWI proved to work well in separating water body and vegetation (Gautam et al., 2015).

Digital terrain models (DTMs) and Land use land cover (LULC) maps can be sourced from RS applications. DTM is a vector dataset that is the representation of bare earth ground surface (Rangari et al., 2019). A DTM allows water surface profiles calculated by hydraulic models to be combined with terrain models to generate 3-D topographic floodplain representations (Jie Yang et al., 2006). The Shuttle Radar Topography Mission (SRTM) data, ASTER Global Digital Elevation Model, JAXA's Global ALOS 3D World and Light Detection and Ranging (LiDAR) are freely available global DEM data sources. In addition to the above RS sources, DEM can be sourced from ground survey techniques and existing topographic maps. Remote sensing has a long history of being used for deriving land-cover maps. Aerial photography and satellite imagery (such as Landsat images) are important source of land-cover information (Sohl & Sleeter, 2012). As DEMs do not vary in spatial resolution, they are not generally suitable for the large-scale terrain representation required for floodplain delineation activities (Anderson & Maidment, 2000). TIN is well suited to represent linear features, such as channel banks, roads and levees (Cameron & Ackerman, 2011). The amount of data is also a problem. Huge amount of data is not easy to handle. It slows down the computer and can be the cause for storage problems (Prinos, 2008). Therefore, it is preferred to convert DEM to a TIN format.

2.8. Hydraulic Modelling

A peak discharge value or hydrograph is transformed into an estimate of flood water elevation, known as flood stage, and an estimate of velocity using a hydraulic model (also called a hydrodynamic model). Hydrodynamic modelling is the study of floods in motion. Hydraulic models are developed for all kinds of purposes such as (Brunner et al., 2020):

- Developing a model to produce rough answers quickly
- A detailed planning study used to evaluate study alternatives
- A design study in which the model will be directly used to design a structure
- Real-time modeling and mapping, consequence mapping for dam or levee failure

The application of hydraulic model in flood hazard assessment is becoming popular (Baky et al., 2019). There are numerous tools that offer a river analysis component to simulate floods. Example, HEC-RAS, LISFLOOD-FP (Shustikova et al., 2019), TUFLOW (Huxley, 2004), Storm Water Management Model (SWMM), SOBEK (San et al., 2020), MIKE-11 etc., just to mention a few.

2.8.1. Dimensions of hydraulic models

The modelling of fluids is complex due to the three dimensions involved and the accompanying time-dependency. Hydrodynamic modeling of water bodies can be performed using one-dimensional (1D), two-dimensional (2D) or in three-dimensional approaches.

i. One-Dimensional models

One-dimensional flood modelling considers the flow in only one of the three coordinate dimensions namely along the central streamlines (along longitudinal downstream direction) in the channel means, no lateral flows are considered. The 1D model just computes how deep that water is going to get, it does not determine the direction (Alzahrani, 2017). In 1D modeling, simplified equations of continuity and momentum are solved in one direction. These models characterize the terrain using a set of cross sections and for each section perpendicular to it, they compute the flow depth and velocity (Pojani & Keçi, 2016). The flow underlying 1D can be steady and unsteady state condition (Alzahrani, 2017). Regarding to the flow velocity, 1D models only produce horizontally and vertically averaged velocities (Brunner et al., 2020).

They are preferred way for the steady flow analysis (Vojtek & Vojteková, 2016). 1D hydraulic model principles are simple to apply and there is a wide selection of software packages available. Also, they reduce computational time and do not need utilization of super computers. 1D models are sufficient to identify areas that are located within flood prone areas, but not for the development of disaster managements plans where various stakeholders would be involved. They can be a good choice in mountain and low mountain range areas when the stream flow takes mainly one direction (Prinos, 2008).

ii. Two-Dimensional (2D) models

In 2D models, the velocity of the flow changes along the longitudinal and lateral directions in the channel with the assumption that the third dimension , water depth, is shallow in comparison to the other two dimension (Teng et al., 2017). Thus, gradients of the velocity exist in two dimensions of the horizontal plane. They are used for the modelling of more complex flood parameters (e.g., flow velocity, flood wave propagation, inundation duration, flow direction and water rise rate). In

2D modeling, equations of continuity and momentum are solved in two dimensions and results are calculated at each grid point in the solution domain.

2D hydraulic models are recommended for complex topography such as urban settings, wider floodplains estuary areas or areas that are affected by tide. They also produces an accurate description of complex conditions such as multiple channels, tidal rivers, and skewed roadway or bridges (Elkhrachy et al., 2021; Prinos, 2008). But they are usually highly demanding in terms of computational time and required data input (Teng et al., 2017). It was reported that they are not viable for areas larger than 1000 km^2, if a resolution of less than 10m and/or multiple model runs are required (Ongdas et al., 2020). Because of their greater sophistication, most 2D models are not freely available (Pojani & Keçi, 2016).

iii. Three-Dimensional (3D) models

Accurate representation of complex dynamics of a river may only be fully captured by the three-dimensional (3D) models that use Reynolds-averaged Navier stokes equations in finite differences, finite elements or finite volume techniques. In 3D hydrodynamic modeling, the target water body is divided into computational cells both in horizontal and vertical dimensions, which enables an accurate description of complex surface water bodies and thus solves the three-dimensional governing equations (X, Y, Z momentum and continuity) in all the three directions. This model is normally applied to complicated reaches as it requires detailed data input, significant computer power and engineering expertise (Dyhouse et al., 2003).

2.8.2. HEC-RAS

The Hydrologic Engineering Center's River Analysis System (HEC-RAS) is one of the Federal Emergency Management Agency (FEMA)-approved software models under the National Flood Insurance Program (NFIP) (Selvanathan & Dymond, 2010), developed by U.S. Army Corps of Engineers. The hydraulic analysis function of the HEC-RAS system contains several river analysis components for steady flow water surface profile computations, one- and two-dimensional unsteady flow simulation, movable boundary sediment transport computations and water quality analysis. It can perform 1D steady and unsteady flow calculation, as well as 2D unsteady flow calculation. HEC-RAS can handle a full network of constructed channels or a natural network such as a dendritic system, single river reaches or looping river network system.

2.8.3. HEC-GeoRAS

Hydrological Engineering Center – Geospatial River Analysis System (HEC-GeoRAS) was developed in a joint effort between the HEC and the Environmental Systems Research Institute

(ESRI) as an extension to ArcGIS (Tazin, 2018). HEC-GeoRAS provides the extra capability to capture the geometric data from existing DTM according to the HEC-RAS format required for the hydraulic modelling. It exports and imports the spatial data to different formats between ArcMap and HEC-RAS by using a data exchange format called a RAS GIS File (Hajibayov et al., 2017; Pathan & Agnihotri, 2019).

HEC-GeoRAS provides pre-processing and post-processing functionality before and after the hydraulic analysis respectively. In pre-processing, it enables the user to create a series of line themes pertinent to developing geometric data for HEC-RAS. This HEC-GeoRAS file gathers geometric data of the study area including the river bed, cross-sections, water flow lines, etc. When this file is imported to HEC-RAS, we can obtain velocity and depth results through hydraulic calculations. Finally, pos-processing using HEC-GeoRAS done to process water surface profile data and velocity data for floodplain mapping, flood damage computations, ecosystem restoration, and flood warning response and preparedness.

2.8.4. RAS layers and their digitization

Creating RAS layers is digitizing RAS GIS Import Files such as Stream Centerline, Bank Lines, Flow Paths, Cross-sectional cut lines, Manning's n value, Ineffective flow areas and defining bridges, culverts and other structures along the river.

i. Stream Centerline, Bank lines and Flow Paths

Stream centerline is the first layer to be digitized that are drawn through the river center from upstream towards downstream flow direction. It is used to establish the river reach network for HEC-RAS. Left and right banks of the river are defined by lines parallel to the Stream Centerlines. They are used to distinguish the main channel from the overbank floodplain areas. The Flow Path Lines layer is used to determine the d/s reach lengths between cross-section in the channel and overbank areas. It contains three lines: centerline, left overbank and right overbank.

ii. Cross sections

Cross sections are the lines drawn to mark the transversal sections of the river, which have the purpose of extracting altitude values from the terrain model, and create a terrain profile, along the river's drainage channel. They are needed to be spatially located to accurately capture the terrain, friction losses, and contraction and expansion coefficients. Too low or too high spacing leads to inaccuracies in the result. The number of cross-sections needed for 1D modelling is highly dependent on the slope of the stream that steeper streams require more cross-sections at shorter distance apart. Additionally, they must be placed at slope changes, around hydraulic structures, at significant changes in roughness, junctions and flow splits, and locations where flow will change. As a rule of thumb, Samuels's equation as used by Vojtek et al., (2019):

$$Dx = \frac{0.15D}{S_O} \quad \text{………………..……..……….. (2.1)}$$

where, D_X is the maximum cross section spacing, D is the average bank full depth and S_O is the average bed slope calculated as the slope of the river centerline as suggested by Vojtek et al., (2019). Sharkey, (2014) suggested the bank full depth as the slimmest bank-to-bank distance in the reach due to the fact that it yielded the most conservative overall spacing.

iii. Bridges/Culverts

Bridges and culverts are treated much like cross-sections. The creation of the bridges layer is based on the cross sections. Four cross-sections are typically needed in the vicinity of a bridge to adequately represent the bridge in HEC-RAS. One cross-section downstream of the expansion effect of flow coming out of the bridge, one cross-section upstream of the contraction effect of the bridge and two cross-sections at the toe of the bridge fill on the downstream and upstream side of the bridge face (Siavash et al., 2016).

iv. Blocked obstructions and Ineffective flow areas

Blocked obstructions layer identifies portions of the flood plain that permanently remove flow area from the cross section. The position of blocked obstructions will be extracted based on the intersection of blocked obstructions layer and the cross-sectional cut lines layer (Cameron & Ackerman, 2011). Buildings in the floodplain and levees can be considered as blocked obstructions. If elevation height data of the blocked obstruction is not available, the default elevation will be defined by the highest elevation on the terrain model at intersection of the blocked polygon with cross sections.

Ineffective flow represents areas where flow velocities are very low (i.e., areas having a combination of flow velocities less than 0.5 feet per second and depths less than three feet) such as the dead-water zones upstream and downstream of bridges and culverts, backwater eddies, overbank areas shadowed by obstructions etc. (Cameron & Ackerman, 2011; Goodell & Warren, 2006; Siavash et al., 2016). Areas behind bridge abutments representing contraction and expansion zones can be considered as ineffective flow areas.

v. Land Use layer

The Land Use layer is a polygon dataset used to establish roughness coefficients for each cross-section. Assigning Manning's n values to individual cross-sections is accomplished by using a land use feature class with Manning's n stored for different land use types. The attribute table of this feature class must have a descriptive field identifying land use type and a field for corresponding Manning's n values.

2.8.5. Steady and unsteady flow conditions

Steady flow assume flow depth, velocity and discharge stays constant at a specific point with respect to time throughout the simulation period (Hajibayov et al., 2017). In general, the assumption is true for shorter reaches of a river system and for events in which the water surface rises and falls slowly. However, as the modeling domain gets larger/or flow rate rises and falls quickly, peak flow rates are not occurring at the same time spatially. Steady flow is recommended for medium to steep sloping stream, but not advisable on extremely flat river systems (Brunner et al., 2020). During unsteady flow, the water discharges change throughout the simulation time (Hajibayov et al., 2017). Unsteady flow routing is the process of determining depths and flows at various locations along within a channel at various times. This means that velocity, discharge and depth are functions of location (distance along the stream channel) and time (Alzahrani, 2017).

If the design discharge is a peak discharge estimate, then the models must be run in a steady flow mode. This is the least intensive mode from a computation standpoint. While steady flow mode is frequently used in flood hazard assessment, it may not capture the complicated flow dynamics in complex terrain such as urban floodplains. If the design discharge is a complete hydrograph, then the models may be run in unsteady mode, in which discharge rates and water levels across the model area can vary over time. Unsteady simulations can require significant computational resources in some cases, particularly when 2D or 3D models are used (Huţanu et al., 2020; Wright, 2015). When the change over time is important for the distribution of the water volume in the area, unsteady calculating model should be used (Prinos, 2008). According to Cheung (2014) review, although real-world situations are better represented by unsteady flow, changes in depth and velocity at a given point normally occur very slowly, even during a flood event, and therefore satisfactory results can be obtained by using a steady flow model.

2.8.6. Flow regimes

HEC-RAS has the capability to model subcritical, critical, and supercritical flow as defined by the Froude number (Fr). Subcritical flow occurs when $Fr < 1$; supercritical flow occurs when $Fr > 1$. Critical flow is defined at the point where the total energy head is a minimum and the Fr equals 1. Both subcritical and supercritical flow can be experienced during floods and therefore, a mixed flow regime can be selected for HEC-RAS simulation (Haestad, 2003). Froude number is computed as:

$$Fr = V/\sqrt{gy} \text{ ……………………............................... (2.2)}$$

where, V = mean fluid velocity (m/s), g = gravitational acceleration (m/s^2), y = water depth (m).

Supercritical flow is characterized by relatively shallow depths and high velocities while subcritical flow is characterized by relatively deep depths and slower velocities. Natural streams rarely have significant lengths of supercritical flow. Steep mountain streams that are alluvial in nature will generally form a step-pool morphology, where the very short "step" goes supercritical, but the reach overall is subcritical (Goodell, 2014).

2.8.7. Boundary conditions

Boundary condition is the conditions or phenomenon occurring at the boundaries of the model. A downstream boundary condition is needed if the flow considered is subcritical regime, upstream boundary if the flow is supercritical and both at upstream and downstream if the flow considered is mixed type of flow. For steady flow analysis, known water surface elevations, critical depth, normal depth and a rating curve can be used as boundary condition. For unsteady flow boundary conditions must be established at all of the open end of the river system being modeled, such as flow hydrograph, stage hydrograph, flow and stage hydrograph for u/s end and rating curve, normal depth, stage hydrograph, and stage and flow hydrograph for d/s end (Brunner, 2020).

At the d/s boundary, the normal depth option is commonly be used. It requires the input of the energy slope that will be used in calculating normal depth (Manning’s equation) at that location (the slope of the channel is often used where the energy slope is unknown). The critical depth requires no input from the user as HEC-RAS automatically calculates it for each. Critical depth is only accurate when used at a significant grade break, a drop structure, a waterfall, etc., where flow passes through critical depth. A rating curve boundary condition requires the user to add elevations against a flow rating curve (Getahun & Gebre, 2015). Using Known water surface elevation as a boundary condition need entering of water surface elevation for the entered flow profile values. A stage hydrograph of water surface elevation versus time may be used as the d/s boundary condition if the stream flows into a backwater environment, such as an estuary or bay, or where it flows into a lake or reservoir of known stage(s) (Brunner, 2020).

2.8.8. Fundamental hydraulic equations

Hydraulic modelling software is based on three basic equations from the physical laws regarding the conservation of mass, energy and momentum in fluid dynamics**.**

2.8.8.1. Manning's equation

In 1D HEC-RAS, the determination of total conveyance is calculated through subdividing flow in the overbank areas using input cross-section n-value break points (location where n-values change) as the basis for subdivision. The main channel conveyance is normally computed as a single conveyance element. The total conveyance is obtained by summing the left, channel and right conveyances (Brunner, 2020). Then, Manning's equation is used to calculate the conveyance as:

$$Q = \frac{1.486}{n} AR^{2/3}S^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.3)$$

where, n = Manning's roughness coefficient for subdivision, A = flow area for subdivision, S = slope of energy grade line and R = hydraulic radius for subdivision (area/wetted perimeter) and n = coefficient of roughness, specifically known as Manning's n.

2.8.8.2. Energy equation

The energy equation is, also known as the law of conservation of energy, is in general used by most of the 1D steady flow programs to solve the one dimensional energy equation. Water surface profiles are computed from one cross-section to the next by solving the energy equation with an iterative procedure called standard step method. The steady flow 1D energy equation (often called Bernoulli's equation) is written as (Brunner, 2020):

$$\mathrm{Z}_2 + Y_2 + \frac{\alpha_2 \mathrm{V}_2^2}{2g} = Z_1 + Y_1 + \frac{\alpha_1 \mathrm{V}_1^2}{2g} + he \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.4)$$

where: Z_1, Z_2 = elevation of the main channel inverts, V_1, V_2 = average velocities (total discharge/ total flow area), Y_1, Y_2 = depth of water at cross-sections, g = gravitational acceleration, α_1, α_2 = velocity weighting coefficients and he = energy head loss.

$$h_e = L\bar{S}_f + C\left|\frac{a_2 V_2^2}{2g} - \frac{a_1 V_1^2}{2g}\right| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.5)$$

Where: L = distance weighted reach length, S_f = representative friction slope between two sections and C = expansion or contraction loss coefficient.

$$L = \frac{L_{lob}\bar{Q}_{lob} + L_{ch}\bar{Q}_{ch} + L_{rob}\bar{Q}_{rob}}{\bar{Q}_{lob} + \bar{Q}_{ch} + \bar{Q}_{rob}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.6)$$

where, $\bar{Q}_{lob}$, $\bar{Q}_{ch}$, $\bar{Q}_{rob}$ are arthimetic average of the laws between sections for the left overbank, main channel, and right overbank respectively, L_{lob}, L_{ch}, L_{rob}, are cress section reach lengths specified for flow in the left overbank, main channel, and right overbank respectively

2.8.8.3. Momentum equation

The momentum equation is based on Newton's second law of motion which states that linear momentum is conserved. The steady flow form of the one-dimensional momentum equation can be written as (Brunner et al., 2020):

$$\frac{\partial Q}{\partial t} + \frac{\partial QV}{\partial x} + g\,A\left(\frac{\partial z}{\partial x} - s_f\right) = 0 \quad \text{……….……………………..…… (2.7)}$$

Where, Q = discharge, A =cross-sectional area, x =distance in the direction of flow, $\frac{\partial z}{\partial x}$ is the water surface slope, z is elevation of the water surface, S_f is the friction slope (from Maning's equation)

2.9. Previous Studies on Flood Risk Mapping and Assessment

The concepts of hazard, vulnerability and risk have been extensively used in flood studies to express facing hazardous flood events. But no unique definitions and assessment procedures have been widely accepted. Thus, different studies used various procedures and techniques. To mention a few, multicriteria analysis (e.g., Analytical Hierarchy Process, AHP) based on GIS and RS, a hybrid approach of integrating hydraulic model, GIS and RS, Topographic Wetness Index (TWI) and using Machine Learning (e.g., Artificial Neural Network, ANN). Studies also use different methodologies for the estimation of input flows and processing software for flood hazard and vulnerability computations by considering various hazard and vulnerability factors. Table 2.1 summarizes some literatures on flood risk mapping and assessment.

Bulti et al. (2017) used multicriteria approach for the assessment of Adama city flood risk map. The results revealed that 10.4% of total area of the city is within high hazard zone and shown that southern part of the city highlights ecological flood risk. This is primarily attributed to low drainage density, low elevation and imperviousness of the land surface. Getahun & Gebre (2015) used HEC-RAS model and a multi criteria evaluation technique for flood hazard assessment and mapping of flood inundation on Awash River Basin. The major findings in the study revealed that inundated areas in the upper and middle part of Awash River Basin are low as compared to downstream part. The other study using multi-criteria analysis is the study by Mahmoud & Gan, (2018). This study developed flood susceptibility maps in Riyadh province, the central region of Saudi Arabia by incorporating 10 susceptibility factors: flow accumulation, annual rainfall, slope, runoff, LULC, elevation, geology, soil type distance from the drainage network and drainage density. Based on the sensitivity analysis, it recommended to consider six or more susceptibility factors in developing susceptibility maps especially factors related to surface runoff and flow accumulation. Khaing et al., (2021) made a case study on Nyaung-U Township in Myanmar to generate flood risk map and made risk assessment using HEC-RAS, for a 10- and 100-year return periods which estimated by flood frequency analysis of gage flow data. The results showed that high risk regions were located near and along the river in the western part of the study area.

Table 2.1. Some literatures on flood risk mapping and assessment

Author	Country	Quantile estimation	Software used	Vulnerability factors	Hazard factors
Quesada-Román et al., (2022)	Costa Rica	RFFA	2D IBER model	infrastructure, population density, social development index	Topographic wetness index (TWI)
Baky et al., (2019)	Bangladesh	Statistical FFA	2D Delft3D, GIS	Cropland and settlement	Inundation depth
Ogato et al., (2020)	Ethiopia	-	GIS	Land use and human population density	LULC, elevation, slope, drainage density, soil, rainfall
Abdulrazzak et al., (2019).	Saudi Arabia	Rainfall-runoff/ HEC-HMS	HEC-HMS, 2D HEC-RAS, GIS	-	Inundation depth
Vojtek & Vojteková, (2016)	Slovakia	Empirical formula	1D HEC-RAS	Land use type	Inundation depth and velocity
Armenakis et al., (2017)	Canada	-	GIS	Large building structures, population demography	-
Liu et al., (2021)	Taiwan	Rainfall-runoff/ HEC-HMS +HEC-GeoHMS	HEC-HMS, FLO-2D, GIS Model	Demographic characteristic, social economy, Rescue equipment, Special agency	Inundation depth
Khaing et al., (2021)	Myanmar	Statistical FFA	HEC-RAS plus GIS	Population Density, Land use, Gender ratio and Age ratio, Urban and rural map and road density	land use, elevation, slope, rainfall, soil, drainage density, inundation duration and depth,
Banstola & Sapkota,(2019)	Nepal	Empirical WECS/DHM	HEC-RAS plus GIS	Landuse type	Inundation depth
Jalayer et al., (2014)	Ethiopia	Rainfall-runoff model	FLO-2D, GIS	Residential, major road corridors, pop. density	TWI
Li et al., (2021)	England	Not mentioned	QGIS	land uses, road networks, buildings	Inundation depth, velocity

Kourgialas & Karatzas (2017) introduced a national scale flood hazard assessment methodology using multi-criteria analysis and artificial neural networks (ANNs) techniques in a GIS environment for Greece. Flood generation factors such as flow accumulation, land use, flow altitude, slope, soil erodibility, rainfall intensity and available water capacity. Those factors are combined and processed using the appropriate ANN algorithm and finally gave a four-category flood hazard map. Quesada-Román et al., (2022) conducted regional flood risk assessment for

Térraba catchment, in Costa Rica. The study used 2D hydrodynamic model (IBER) to compute flood depth for quantiles estimated using RFFA using Bayesian Markov Monte Carlo Chain (MCMC) algorithm. Flood risk is computed as a function of hazard, exposure, and vulnerability. The exposure to floods was defined with the population density and infrastructure, social development index as vulnerability factor and Topographic Wetness Index (TWI) as a hazard factor. It showed that regional flood risk assessments can be performed in large-scale catchments if both coarse and detailed inputs are used.

Regarding the studies on flood hazard and risk assessment on Middle Awash Sub-basin, less studies are conducted than on Upper River Basin, till this review is conducted. Gadise, (2020) conducted a thesis study on MASB to map flood inundation using HEC-RAS by applying frequency analysis based on L-moments for the estimation of flood quantiles. The study revealed that all sites on the MASB form a homogeneous region and indicated that all parts of Awash below Koka Dam up to Wonji, some parts of Awash Nura Hera up to Metehara and some parts of Awash at Metehara up to Melka Werer River reaches were under the impact of 25, 50, 100-, 200-, 500- and 1000-year flood problem. Although this study showed that the region is homogeneous, it has estimated flood quantiles based on selected distribution for each stations instead of regional data. Also, the hazard, vulnerability and risk condition of the basin is not assessed. Nama & Tesfaye, (2020) studied on analysis of flood hazard on Dulecha *Woreda* which is found on MASB using HEC-RAS for inundation computation and HEC-HMS for runoff computation of the 25-Aug-08 flood event. Finally, it found that the flood event gave a maximum flood depth of 3.51m. Another study conducted on MASB is by Haile, (2009) using a 1D- 2D Sobek – Rural hydrodynamic model to carryout flood risk assessment of the area. It estimated flood damage using the Hydro-Economic Expected Annual Damage (EAD) method. The study showed that, for Metehara area, the most effective strategy is regulating a 10-year return period flow and water level since the relative increment of expected annual damage cost is the highest at 10-year return period flow.

3. MATERIALS AND METHODS

3.1. Study Area Description

3.1.1. Location

The Awash River Basin is divided into three distinct zones based on its climatological, physical, socio-economic, agricultural, and water resource characteristics as the Upper Awash, the Middle, and the Lower Basins (Duguma et al., 2021; Mitiku et al., 2020). The Middle Awash River Basin (MARB) is defined by the Ministry of Water, Irrigation and Energy (MoWIE) as the watershed area downstream of Koka and upstream of Melka Werer town in Afar Region (MoWIE, 2015).

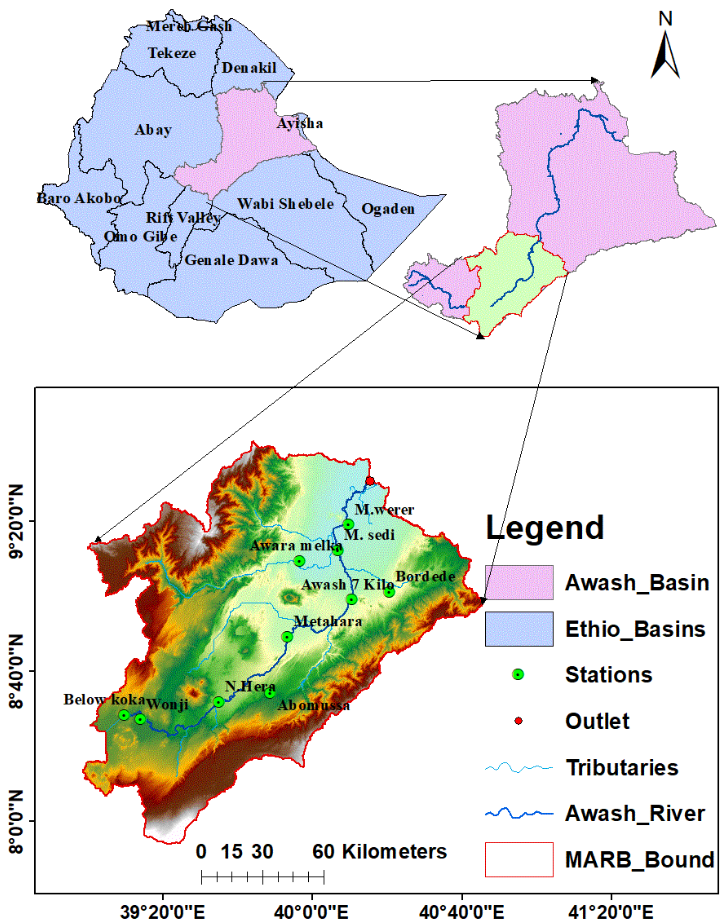


Figure 3.1. Location Map of the Study Area

For this study, the MARB considered lies between latitude of 7° 53′ 49″ N - 9° 41′ 1″ N and longitude of 38° 58′ 36″ E - 40° 46′ 1″ E with a total drainage area of about 32270 km^2 and delineated area of about 20,552 km^2 at the outlet on Dulecha *Woreda* at 9° 30′ 52″ latitude and 40° 16′ 10″. The river reach considered starts from downstream of Koka reservoir to the outlet having a reach length of 297kms. The elevation of the MARB considered ranges from 696 to 4190 meters, which shows a higher altitudinal difference within the basin.

3.1.2. Climate

The rainfall distribution is generally bimodal in the MASB. Based on the available rainfall and temperature data of seven gauge stations in the sub-basin (shown in Figure 3.2 and 3.3), the area receives low rainfall on the months of May-June and November-January. The main rainy season from July- September accounts for about 53.28% of the annual average rainfall. Another minor rainy season is March-April accounting 17.75% of the total average annual rainfall. November-February is less sever dry season with relatively cool temperatures. August is the wettest month with monthly average maximum rainfall of 98 to 244mm respectively on Awara Melka and Nazeret stations. The monthly average maximum temperature ranges from 25.975˚C to 39.25ºC while the monthly average minimum temperature ranges from 12.23˚C to 24.06˚C. The monthly average maximum temperature records are observed on May and June while the minimum is on December for all stations.

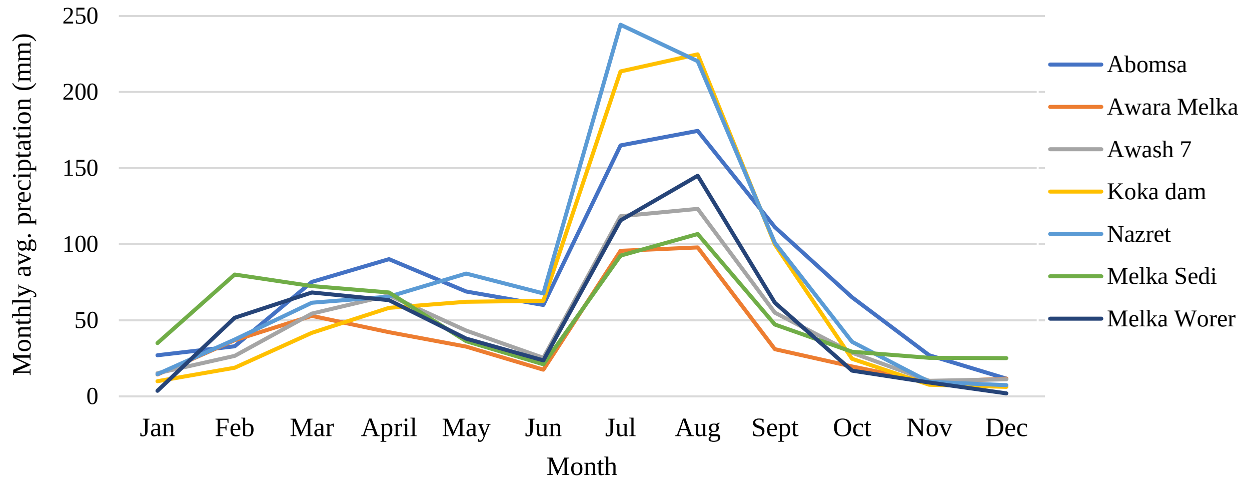


Figure 3.2. Monthly average precipitation distribution of the area

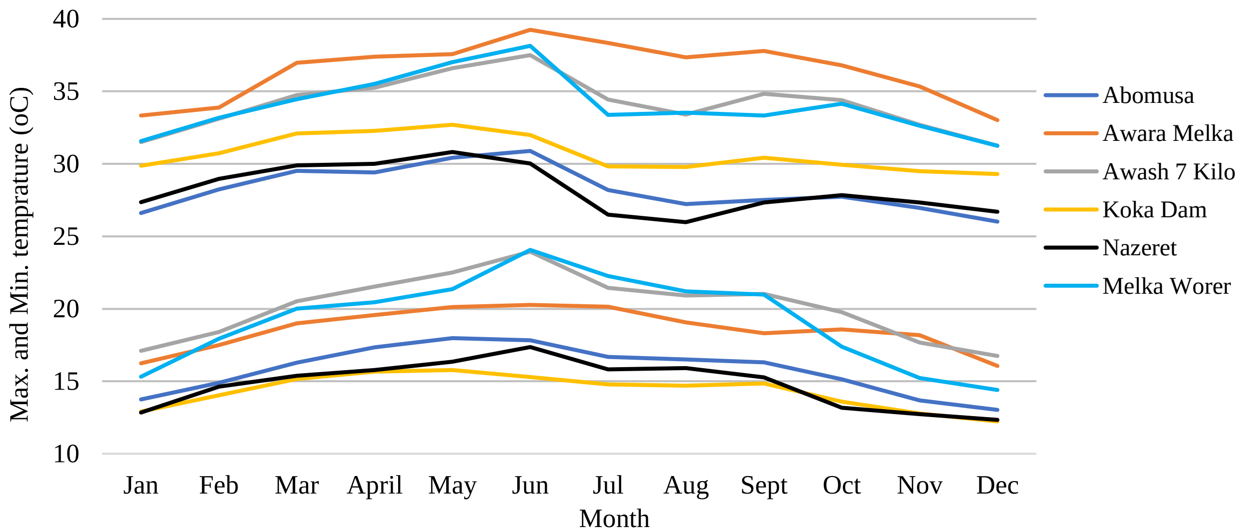


Figure 3.3. Mean monthly maximum and minimum temprature distribution of the area

3.1.3. Soil and geology

According to FAO soil classification, there are 11 soil types in the study area. About 50% of the area is covered with Eutric cambisol and Lithosols. The remaining part is covered with Calcic xerosols (11.57%), chromic vertisols (5.74%), Pellic vertisols (5.78), Haplic yermosols (11.86%), Calcaric fluvisols (1.09%), Haplic xerosols (5.09%), Eutric nitosols (5.67%), Water (0.35%) and Dystric regosols (0.42%). The soil texture comprises clay (50.47%), clay-loam (11.86%), loam (13.08%), sandy-clay-loam (19.14%), silt-loam (5.09%) and water (0.35%).

World Geologic Maps from USGS (*https://certmapper.cr.usgs.gov/data/apps/world-maps/*) shows that about 70.59%, of the MARB is covered with Tertiary extrusive and intrusive rock. The remaining part of the geology is covered with Quaternary extrusive and intrusive rock, Quaternary, Holocene, Water and Pleistocene with an area coverage of 21.97%, 2.43%, 4.27%, 0.44% and 0.30% respectively. Soil map and geological map of the area is shown in Figure 3.4.

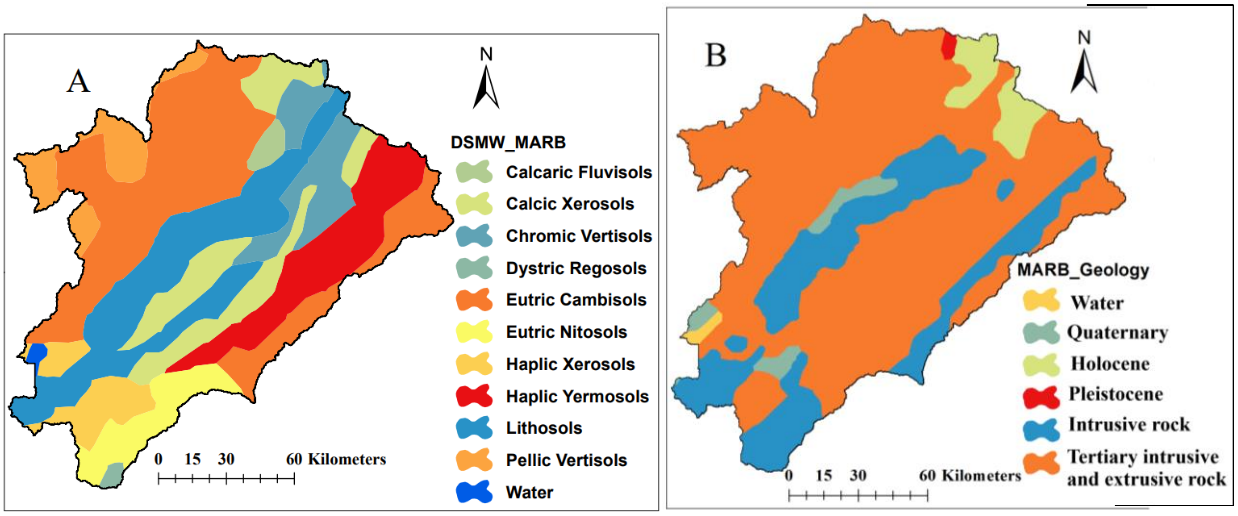


Figure 3.4. Soil map (A) and geological map (B) of the study area

3.1.4. LULC type

According to the LULC map of Sentinel-2 satellite image (10m resolution), there are seven types: dense vegetation, flooded vegetation, water bodies, crops, built area, bare ground and rangeland as shown in Figure 3.3. About 50% of the study area is occupied by rangeland followed by crop land, which cover about 41% of the total area. Land use types of dense vegetation and built-up area covers 5.31% and 2.44% respectively of the total study area coverage. The remaining portion is covered with flooded vegetation (0.06%), bare ground (0.17%), and water bodies (0.94%). Description of each landcover type of Sentinel-2 data is found in Appendix E.

The common types of vegetation's of the study area such as *P. julifora*, Acacia, shrubs, thorny bushes, grassland and riverine woodland trees are found along the Middle Awash River floodplains (Lammesa, 2017). Historically sugar and cotton have been the major crops grown in Middle and Lower Awash Valley and they are major cotton producing areas of Ethiopia. In addition, other crops such as banana, onion, and tomato are found with small area coverage. Currently, *P. juliflora* poses a threat to indigenous biodiversity because of its weedy and invasive nature (Nanesa, 2021).

3.2. Data Collection and Materials Used

Before any hydraulic modelling commences, data needed for the study should be collected and analyzed for data quality. HEC-RAS needs the three specific input parameters to generate the water surface level along the river: flow data, stream geometry, and the model plan. Therefore, secondary data from different organizations and online sources are collected including stream flow, topographic data (DEM) and land use (land cover) data. In addition, historical flood satellite image, administrative shape files, bridge data and recorded water level data were collected.

3.2.1. Raw data collected

i. Flow and water level data

Table 3.1. Flow gauge stations

River @ station	Lat. (N)	Long. (E)	Record years	No. of years (n)	Catchment Area (km^2)	Missed data (%)[1]
Awash @ Melka sedi	9º12’03.0’’	40º07’13.2’’	1990-2014	25	21520	0.00
Awash @ Melka werer	9º18’56.7’’	40º10’00.4’’	1990-2010	21	31183	1.05
Awash @ Wonji	8º27’05.2’’	39º14’10.1’’	1990-2010	25	11690	1.61
Awash @ Metahara	8º49’00.2’’	39º53’35.1’’	1990-2014	21	16416.8	13.19
Awash @ Nura Hera	8º31’38.3’’	39º35’07.6’’	1980-1995	16	14173	0.87
Arba @ Abomussa	8º34’00.0’’	39º49’00.0’’	1995-2010	16	140	1.46
Kessem @ Awara melka	9º09’14.6’’	39º57’00.3’’	1990-2009	20	3113	6.68
Arba @ Bordede	9º00’57.1’’	40º21’00.1’’	1993-2014	22	72	26.94
Burka @ Asebe teferi	-	-	1980-1989	10	75	17.54
Awash below koka dam	8º28’15.1’’	39º09’47.8’’	1990-2013	24	11219	2.46
Awash @ 7 kilo	8º58’57.1’’	40º11’00.1’’	1990-2009	20	19110.8	7.09

A daily instantaneous flow data of 11 stream flow gauging stations lying in the MARB of the Awash River and its tributaries varying over 10 – 25 years in record length were collected from

[1] The rainy season is considered (June-September)

Ethiopian Ministry of Water and Energy (MoWE). The annual peak floods of these sites vary from 7.51m^3/s (on Arba @[2] Abomusa) to 721.5 m^3/s (on Kessem @ Awara Melka) and their catchments areas range from 72 km^2 (Arba @ Bordede) to 31,183 km^2 (Awash @ melka werer). Water level data of the station Awash @ 7 kilo for the year 1968 - 2000 was collected from MoWE and combined with river flow data to generate a rating curve.

ii. Remote Sensing (RS) products

The downloaded RS products for this study are topographic data (DEM)[3], historical Landsat image[4], aerial photograph and land use land cover (LULC) map[5]. On-site collection or extraction from topographic data of cross-sectional data needs high budget allocation and is time consuming. ALOS-PALSAR (Advanced Land Observing Satellite-Phased Array-Type L-Band Synthetic Aperture Radar) Radiometrically Terrain Corrected (RTC) DEM, 12.5m resolution, is downloaded from Alaska Satellite Facility (ASF) and used to create TIN which in turn used for RAS-layers building and final inundation mapping. Due to its higher resolution, it is proven to be very effective in extracting stream network and are suggested for morphometric studies over other DEM products. Also, it can identify a much shallower and narrower drainage networks which are unidentified in other coarse resolution products (Nitheshnirmal et al., 2019). The downloaded raster files of the DEM are downloaded, merged and then clipped to the area of interest (AOI).

For calibration and validation of hydraulic models historical maps of past flood events from RS sources for a flood event can be used (Nogherotto et al., 2022). Historical Landsat image (Landsat 4-5 TM) of the Aug 18, 1996 flood event was downloaded from USGS Earth Explorer for validation of HEC-RAS inundation extent map output based on Normalized Difference Water Index (NDWI). High resolution aerial photograph of Google Earth image is downloaded and georeferenced in GIS to aid in identifying the location of buildings, bridges and the river centerline and its right and left banks.

LULC data of Sentinel-2 on a sensing date of 2022-02-29 from Copernicus Access hub. Sentinel-2 is a wide-swath, high resolution, multi-spectral imaging mission, supporting Copernicus Land Monitoring studies, including the monitoring of vegetation, soil and water cover, as well as observation of inland waterways and coastal areas. They are excellent candidates for LULC mapping due to the high spatial, spectral, and temporal resolution (Karra et al., 2021). After merged and clipped to the AOI, i.e., MARB, it shows a cloud cover of almost 0% coverage.

[2] The location of the gauging stations
[3] *https://search.asf.alaska.edu/*
[4] *https://earthexplorer.usgs.gov/*
[5] *https://scihub.copernicus.eu/dhus/#/home*

Sentinel-2 comprises 10 landuse/ cover types and only seven types are found on the study area as shown in Figure 3.3. LULC map for this study was used in the extraction of the Manning roughness coefficient (n) in HEC-GeoRAS software.

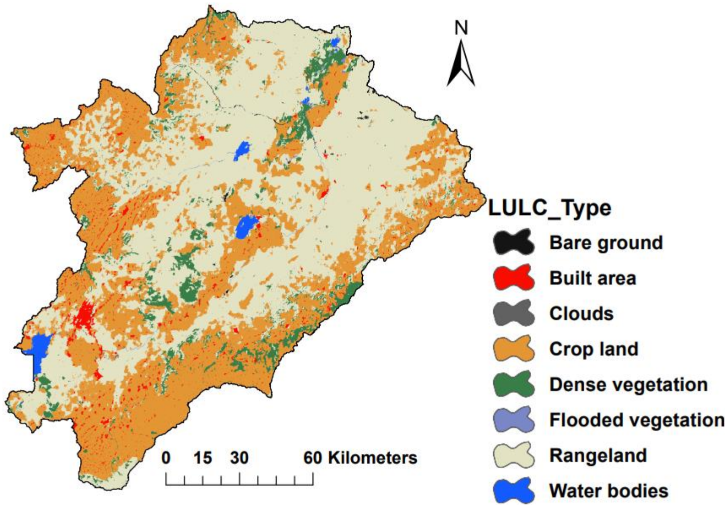


Figure 3.5. Sentinel-2 image for LULC classification

iii. Bridge data and administrative shapefiles

Freely available administrative shapefiles from "The Ethiopia GeoPortal"[6] is used to download *woreda*, Town and *kebele* shapefiles of the study area. This enables the administrative level hazard and risk assessment for easy understanding and implementation of the results. Bridge data for three bridges was collected from Ethiopian Roads Authority - Bridge Management System (ERA-BMS)[7] which is in house developed web-based software having information and database of more than 3,600 bridges and 30,000 Culverts. The bridge data is used to enter dimensions of the bridge in HEC-RAS after HEC-GeoRAS pre-processing.

Table 3.2. Collected bridge data

Bridge name	x-cord. (m)	y-cord. (m)	Altitude (m)	Bridge Length (m)	Bridge Width (m)	No of pier	Pier width (m)	Deck depth (m)
Awash – Ras Mekonen WM	630089	997955	823	157.54	8.90	2	10.00	1.36
ELO/AWASH	582931	956075	1155	45.10	4.40	2	9.20	0.70
Awash	562252	940726	1242	41.00	6.00	2	8	1.40

[6] *https://ethiopia.africageoportal.com/*
[7] *http://213.55.9.172:8087/*

3.2.2. Data analysis tools

To aid the computation of proposed methods of the study, the software ArcGIS 10.4, HEC-RAS 5.0.7, HEC-GeoRAS 10.1, RStudio, Rainbow and Microsoft Excel were used. R (R Core Team, 2021) is a well-known language for statistical computing and graphics. R provides a suite of software facilities for reading and manipulating data, computation, conducting statistical analyses and displaying the results (Kuhnert & Venables, 2005). RStudio is an Integrated Development Environment (IDE) which makes using R easier.

Table 3.3. Data analysis tools

Software used	Use
ArcGIS 10.4	Interpretation and visualization of HEC-RAS results, support HEC-GeoRAS
HEC-RAS 5.0.7	Inundation map generation
HEC-GeoRAS 10.1	Geometric layers extraction and data exchange between HEC-RAS and GIS
RStudio	Trend, independence and stationarity test and RFFA
Rainbow	Data homogeneity test
MS Excel	Data consistency check

RAINBOW[8] is designed to study meteorological or hydrologic records by means of a frequency analysis and to test the homogeneity of the record. The homogeneity test is based on Buishand, (1982) homogeneity test. RAINBOW offers statistical tests for investigating whether data follow a certain distribution (Chi-square and the Kolmogorov-Smirnov test). Also, RAINBOW allows to analyze time-series with zero or near zero events (the so-called nil values) by separating temporarily the nil values from the non-nil values (Raes et al., 2006).

3.3. Exploratory Flow Data Analysis

Exploratory data analysis is used to perform initial investigations on the collected flow data to discover patterns, to spot animalities, to test hypothesis and to check the data whether it fulfills the assumptions of flood frequency analysis. This included data screening based on visual inspection and criteria settled, data model selection, consistency check, outlier test and statistical tests.

3.3.1. Data screening

The critical issue during data collection is to find sufficient data of good quality. The short data size can affect the choice of distributions, the quantile estimations, particularly those

[8] Freely available from: *http://www.iupware.be*

corresponding to large return periods and the extent of confidence intervals (Benameur et al., 2017). Setting some criteria and visual inspection of the daily time series can be used to screen out data that are not suitable for the frequency analysis.

A daily flow data with a minimum record length of 10 years is taken as a sufficiency condition according to the recommendation of Mfwango et al., (2018) and Bulletin 17C (England et al., 2019). Only sites having observation records longer than 10 years were used to avoid errors due to short record length (Saf, 2009). Each station is initially screened using the following 4 criteria:

1. Stations which have a short record length (<10 years),
2. Stations with missed value of >15% during rainy season,
3. Stations which consist a greater number of "no data" in the series (contain more consecutive gaps for months or years),
4. Rivers which reflect not natural phenomenon.

As shown in Table 3.1, the stations have a record length 10 (Asebe Teferi) to 25 years (Melka Sedi) which fulfills the minimum record length requirement. Stream flow stations were accepted with missing values of less than 15% in order to have better representative station distribution within the basin (Tadese et al., 2019). Due to a missing value greater than 15%, two stations (Bordede and Asebe Teferi) were screened out from subsequent step of data analysis. But no consecutive gaps for months and more have been seen on the data series of all stations.

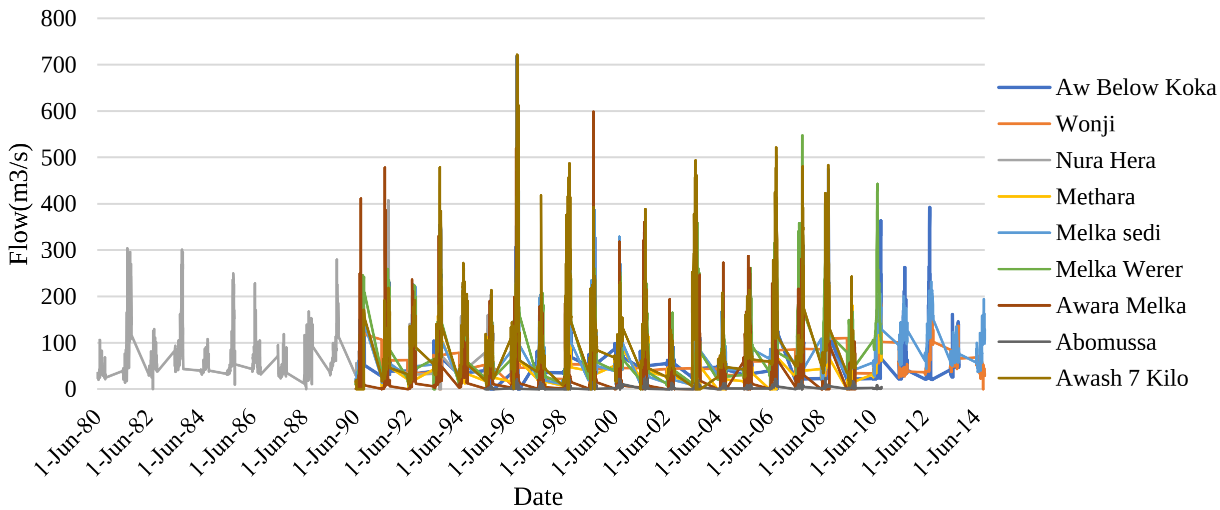


Figure 3.6. Daily time series plot of flow stations

A visual inspection of a plot of the time series should be the first step in assessing the time series (England et al., 2019). The plot of a series of a hydrologic variable versus time over the total length of the observation period is useful for visual investigation that usually provides a qualitative evaluation. A possible trend, a sudden change, and even a periodicity may reveal themselves by a careful visual observation of the plotted values (Haktanir et al., 2013).

Conventional methods of flood frequency analysis are not suitable for obtaining flood estimates from regulated flow data series (Ishak et al., 2013). If stations have repeated values for long period of time and/ or some constant fluctuations, this could be in catchments which are under control at somewhere upstream or in stations that the gauging instruments are not able to measure high magnitude floods. The instantaneous daily water flow of Kessem @ Awera Melka shows a sharp increase and decrease with a great variation in flow for few days showing unnatural flow during the rainy season. Also, repeating values are visualized in the daily data time series. A similar behavior of the series is visualized on Awash below Koka station and thus excluded from the subsequent analysis. This regulation of the stations is also mentioned by Tadese et al., (2019) coated as “Koka and Kessem streamflow stations are regulated at Koka and Kessem reservoirs, respectively, while Tendaho and Mille stations are monitored at Tendaho reservoir”.

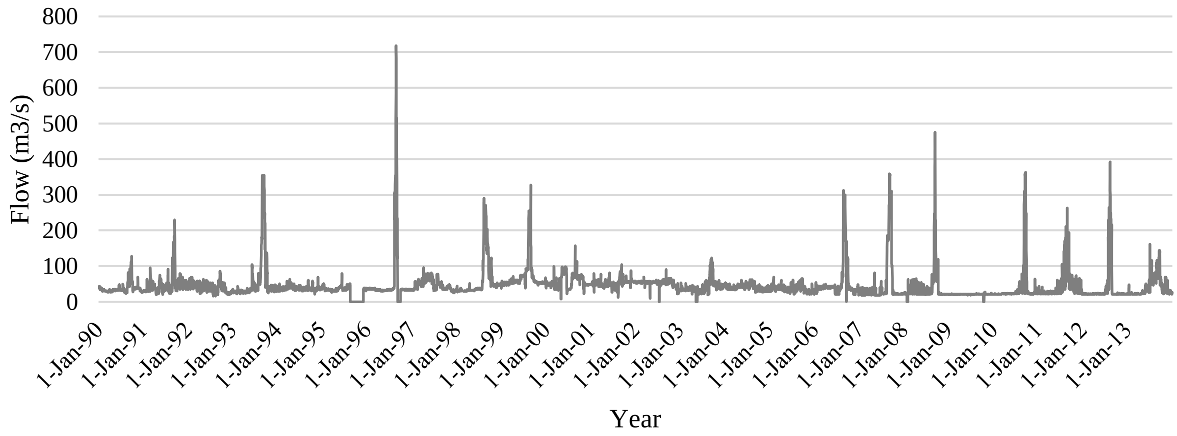


Figure 3.7. Daily time series plot for Awash below Koka station

After this initial screening, the remaining seven stations (Awash @ Melka sedi, Awash @ Melka Werer, Awash @ Wonji, Awash @ Metahara, Awash @ Nura Hera, Arba @ Abomussa and Awash @ station 7 kilo) have passed to the subsequent data quality tests.

3.3.2. Data model selection

After the collection of hydrologic data, it is necessary to select the data series model type. From the two data models, AMS and PDS, AMS is widely used and is selected for the computation flow quantiles in the RFFA. It enables easy selection of the values and assures the values are independent. The use of PDS needs subjective judgment of the threshold value, as there is no standard guideline for selection and the analysis complex. The AMS allows independence of values and is a suggested model for flood risk evaluation (Leščešen & Dolinaj, 2019).

Table 3.4. AMS of selected stations for frequency analysis

Year	Wonji	Nura Hera	Methara	Melka Sedi	Melka Werer	Abomussa	Awash 7 Kilo
1980	-	106.20	-	-	-	-	-
1981	-	303.22	-	-	-	-	-
1982	-	129.30	-	-	-	-	-
1983	-	300.80	-	-	-	-	-
1984	-	107.70	-	-	-	-	-
1985	-	249.90	-	-	-	-	-
1986	-	228.30	-	-	-	-	-
1987	-	118.49	-	-	-	-	-
1988	-	167.20	-	-	-	-	-
1989	-	279.40	-	-	-	-	-
1990	136.4	222.00	167.518	175.559	245.33	-	170.938
1991	199.1	407.50	231.028	201.962	259.174	-	218.455
1992	126.1	197.60	208.087	225.127	224.06	-	191.108
1993	209.9	356.10	248.078	170.931	263.54	-	478.365
1994	96.88	226.20	166.943	177.434	261.351	-	272.638
1995	89.27	209.60	112.115	141.854	199.537	7.51	213.788
1996	262.8	-	214.114	426.628	345.778	21.284	721.5
1997	65.61	-	132.871	203.563	206.874	14.572	418.683
1998	228.106	-	193.326	348.733	359.202	20.914	487.17
1999	166.185	-	168.363	398.56	261.351	14.572	392.147
2000	56.359	-	125.694	328.823	261.351	21.658	231.114
2001	81.693	-	138.495	142.262	261.351	22.801	388.132
2002	87.215	-	91.221	104.589	164.739	13.979	127.297
2003	103.37	-	97.904	178.851	359.202	19.471	493.82
2004	79.357	-	32.736	168.375	272.425	12.284	208.452
2005	106.783	-	114.054	192.986	261.351	9.261	213.79
2006	225.608	-	199.824	225.384	387.126	22.035	521.393
2007	262.993	-	228.151	210.306	547.763	18.426	480.456
2008	268.66	-	225.375	193.378	397.243	23.582	483.626
2009	110.053	-	179.929	121.74	167.131	18.085	242.494
2010	194.393	-	182.855	205.564	442.683	8.359	-
2011	100.492	-	-	199.628	-	-	-
2012	228.398	-	-	231.962	-	-	-
2013	136.388	-	-	135.209	-	-	-

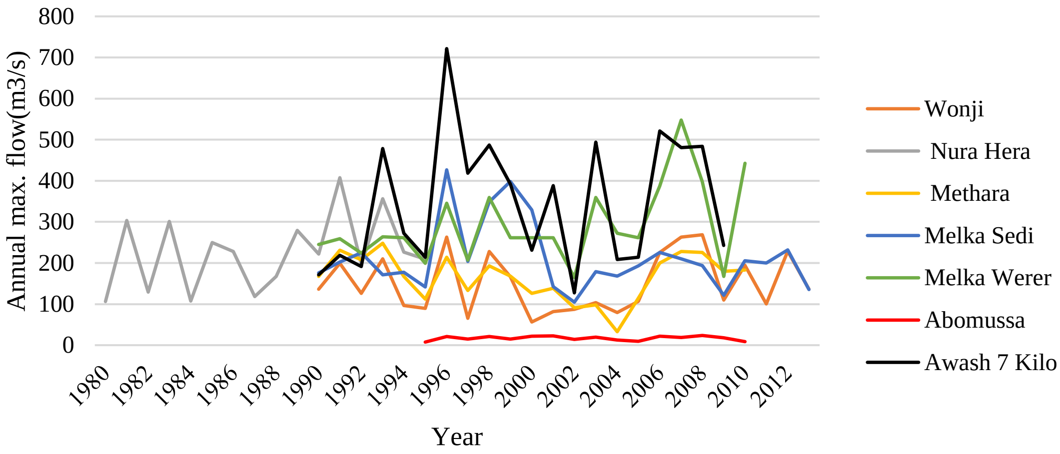


Figure 3.8. Annual maximum series plot

3.3.3. Consistency of data

Inconsistency of streamflow records can be due to a change in the method of collecting streamflow data or major changes in water use, storage, or evapotranspiration in the basin (Searcy & Hardison, 1960). Double mass curve (DMC) analysis can be used to check the consistency of the available data. A DMC is a graph of the cumulative catch at the discharge gage of interest versus the cumulative catch of one or more gages in the region that has been subjected to similar hydrometeorological occurrences and are known to be consistent. The selection of the stations depends on the geographical location of the stations. i.e., its neighboring stations can be used to check the reliability of that station and correction is then made accordingly (Sine & Ayalew, 2004).

When plotted, the double mass diagram should show a linear relationship when the suspect data set is consistent (Adoleye & Montaseri, 2002). It should not be assumed that all inconsistencies shown by a double mass curve represent inconsistencies due to change in methods of collecting the data. Unless the inconsistency shown by the DMC starts at the same time as a change in method of data collection, and unless the direction of the change shown by the curve could reasonably result from the change in method, no adjustment should be applied to the observed data, because the inconsistency could be due to other causes, such as works of man or vagaries of nature (Searcy & Hardison, 1960). The inconsistent data can be adjusted by applying the following formula (Searcy & Hardison, 1960):

$$P_a = \frac{b_a}{b_o} P_0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

where, P_a = adjusted flow, P_0 = observed flow, b_a = slope of graph to which records are adjusted, b_o = slope of graph at time P_0 was observed

3.3.4. Statistical tests on the annual maximum data series

Statistical tests are conducted on the AMS to test if the data conform to the assumptions used in the frequency analysis. i.e., independence and identical distribution (iid) of the time series. The null hypothesis (H_0) is tested at alpha value of 5% and 10% compared with p-values. In this study, idd of the time series tested using R Packages for Mann-Kendall trend test (with “Kendall” package[9]) and W-W test of independence and stationarity (*ww.test*[10] in R with the “trend” package), and RAINBOW software is used for Buishand’s homogeneity test.

3.3.4.1. Test for trend

A ‘trend’ is defined as a unidirectional and gradual change (falling or rising) in the mean value of a variable (Machiwal & Jha, 2012). Trend exists in a data set if there is a significant correlation (positive or negative) between the observations and time (Adoleye & Montaseri, 2002). The non-parametric Mann-Kendall trend test (MK test) is widely in use due to its independence in the data distribution and not being affected by missing values and outliers (Tadese et al., 2019). Also, the test has low sensitivity to abrupt breaks due to inhomogeneous time series (Likoko et al., 2019). It identifies the significant monotonic upward or downward trends in hydro-meteorological data series without specifying the type of trend (i.e., linear or nonlinear) (Hesarkazzazi et al., 2021; Machiwal & Jha, 2012). The data are ranked according to time and then each data point is successively treated as a reference data point and is compared to all data points that follow in time (Douglas et al., 2000).

The test statistics Z, S and tau are measures of significance used for Kendall test. The test statistic, Kendall's S, is calculated as (Douglas et al., 2000):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign(x_i - x_j) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.2)}$$

where x is the data point (in this case, a flood) at times i and j and

$$sgn(x_i - x_j) = \begin{cases} +1, \; if + (x_j - x_i) > 0 \\ \;\; 0, \; if (x_j - x_i) = 0 \\ -1, \; if (x_j - x_i) < 0 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \text{(3.3)}$$

While the amount of data series is larger than or equivalent to ten ($n \geq 10$), since $n \geq 10$, the MK test is then categorized by a standard distribution with the mean $E(S) = 0$ and variance Var(S) is given as (Daba et al., 2020):

---

[9] *https://www.rdocumentation.org/packages/kendall/version 2.2*

[10] *https:/rdrr.io/cran/trend/man/ww.test.html*

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{k=1}^{m} t_k(t_k-1)(2t_k+5)}{18} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.4)$$

where m is the number of the tied groups in the time series and t_k is the number of ties in the k^{th} tied group. From this, the test Z statistics is obtained using an approximation as follows:

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(S)}}, \; if \; S > 0 \\ 0, \; if \; S = 0 \\ \frac{S+1}{\sqrt{Var(S)}}, \; if \; S < 0 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.5)$$

tau is calculated as S/D, and D is the “denominator” which is the maximum possible value of S.

$$S = \frac{n(n-1)}{2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.6)$$

Local (at-site) p-values for each trend test can be obtained as (Douglas et al., 2000):

$$p = 2[1 - \Phi(|z_s|)] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.7)$$

where Φ ($|z_s|$) denotes the cumulative distribution function (cdf) of a standard normal variate.

The hypotheses for the test are as follows:

- H_0: There is no trend present in the data.
- H_a: A trend is present in the data (a positive or negative trend)

3.3.4.2. Tests for independence and stationarity

The hydrologic interpretation of the mathematical stationarity is that the series does not exhibit any significant changes over time (Chebana et al., 2010). A formal definition of stationary time series is one in which its mean, Variance and covariance do not vary with time. For example, a series of annual river flow would be considered stationary if there is no significant increase or decrease of mean and/or scale throughout the observation record, except for the usual stochastic variation. A hydrological time series is stationary if is free of trends, shifts (jumps), or periodicity (cyclicity) (Chebana et al., 2010; Machiwal & Jha, 2012; Stone, 2001). Trends in a time series can result from gradual natural and human-induced disruptive and evolutionary changes in the environment, whereas a jump may result from sudden catastrophic natural events (Haan, 2002). In nature, strictly stationary time series does not exist, and weakly stationary time series is practically considered as stationary time series (Machiwal & Jha, 2012).

In order to test if the elements of a recorded series are realizations of independent and stationary occurrences, as originally proposed by Wald and Wolfowitz in 1943, the non-parametric Wald–Wolfowitz’s (W-W) test based on the Z test-statistic is used as suggested by Rao & Hamed, (2000).

Let x_1, x_2, ..., x_n denote the sampled data, then the test statistic of the W-W test is calculated as:

$$R = \sum_{i=1}^{n-1} x_i x_{i+1} + x_1 x_n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.8)}$$

The expected value of R is:

$$E(R) = \frac{s_1^2 - s_2}{n-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.9)}$$

The expected variance is:

$$V(R) = \frac{s_2^2 - s_4}{n-1} - E(R)^2 + \frac{s_1^4 - 4s_1^2 s_2 + 4s_1 s_3 + s_2^2 - 2s_4}{(n-1)(n-2)} \quad \ldots\ldots\ldots \text{(3.10)}$$

with: $$S_t = \sum_{i=1}^{n} x_i^t \,, \ t = 1,2,3,4 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.11)}$$

For n > 10 the test statistic is normally distributed, with:

$$z = \frac{R - E(R)}{\sqrt{V(R)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.12)}$$

ww.test calculates p-values from the standard normal distribution for the two-sided case. The hypotheses for the test are as follows:

- H_0 = The series is independence and stationarity
- H_a = The series is significantly different from independence and stationarity

3.3.4.2. Test for homogeneity

The term ‘homogeneity’ implies that the data in the series belong to one population, and therefore have a time invariant mean (Machiwal & Jha, 2012). Non-homogeneity is a change in the statistical properties of the time series (Dahmen & Hall, 1990). Floods may be from different populations because they occurred before the building of a dam and after the building of a dam, or before the watershed was urbanized and after it became urbanized, or because some are generated by summer storms and others by snowmelt, or because some were generated in El Nino˜ years and some were in other years. It may be difficult in some cases to definitively say if the flood record is homogeneous (England et al., 2019).

There are different methods to perform homogeneity test such as standard normal homogeneity test (SNHT), Pettitt’s test, Buishand range test (BRT), and Von Neumann ratio (VNR) as used by Daba et al., (2020). For this study a statistical software package called RAINBOW was employed to generate the residual mass curve and compute the homogeneity statistics which uses Buishand’s homogeneity test. RAINBOW offers a test of homogeneity which is based on the cumulative deviations from the mean. By evaluating the maximum (Q) and the range (R) of the cumulative deviations from the mean, the homogeneity of the data of a time series is tested. High values of Q or R are an indication that the data of the time series is not from the same population and that the fluctuations are not purely random (Kisaka et al., 2015; Raes et al., 2006). Mathematically,

$$S_k = \sum_{i=1}^{k}(X_i - \bar{X}) \text{ k = 1, ---------- n, ……………... (3.14)}$$

where; X_i = the runoff record for the series X_1 X_2 ---- X_n, $\bar{X}$= the mean and S_k = the residual mass curve and

$$Q = \max\left[\frac{Sk}{S}\right] \text{ ……………………………..…... (3.15)}$$

$$R = \max\left[\frac{Sk}{S}\right] - \min\left[\frac{Sk}{S}\right] \text{ …………....….….….. (3.16)}$$

where, Q is maximum of Sk and R in the range of Sk and min is Minimum

The underlying hypothesis of homogeneity was formulated as follows:

- H_0: data are statistically homogeneous
- H_1: data are not homogeneous

3.3.5. Outlier detection and treatment

3.3.5.1. Outlier test

Outliers, observations whose values are quite different than others in the data set, often cause concern or alarm. Fitting a distribution to the data is challenging when there are outliers in the data. They can be caused by a recording error, an observation from a population not similar to that of most of the data such as a flood caused by a dam break rather than by precipitation and a rare event from a single population that is quite skewed (Helsel & Hirsch, 1992).

For the outlier analysis, the method by USWRC, (1981), also known as Bulletin 17B, is considered. If the station skew is greater than +0.4, tests for high outliers are considered first. If the station skew is less than -0.4, tests for low outliers are considered first. Where the station skew is between + 0.4, tests for both high and low outliers should be applied before eliminating any outliers from the data set. The following equations are used to detect high and low outliers.

$$X_H = \bar{X} + K_N S \text{ and } Q_H = \text{Antilog } (X_H) \text{ ………... (3.17)}$$

$$X_L = \bar{X} - K_N S \text{ and } Q_L = \text{Antilog } (X_L) \text{ ……...… (3.18)}$$

where, X_H and X_L are high and low outlier threshold in log units respectively and K_N is K value from Appendix 4 on USWRC, (1981) for sample size N. See Appendix C of this document. $\bar{X}$, S and skew coefficient are calculated as:

$$\bar{x} = \frac{\sum X}{N} \text{ ………………………………….…… (3.19)}$$

$$S = \left[\frac{(\sum X^2) - (\Sigma X)^2/_N}{(N-1)}\right]^{0.5} \text{……………………… (3.20)}$$

$$G = \frac{N^2(\sum X^3) - 3N(\sum X)(\sum X^2) + 2(\sum X)^\wedge 3}{N(N-1)(N-2)S^3} \text{……….. (3.21)}$$

where, X = logarithm of annual peak flow, N = number of items in data set, $\bar{X}$= mean logarithm, S = standard deviation of logarithms of X's and G = skew coefficient of logarithms of X's

If the logarithms of peaks in a sample are greater than X_H or less than X_L, then they are considered high and low outliers respectively. The test hypothesis interpretation is as follow:

- H_0: There is no outlier in the data
- Ha: The minimum or maximum value is an outlier

3.4. Spatial data analysis

3.4.1. LULC map preparation

HEC-RAS simulations require each cross-sectional line to carry a Manning's n value in the geometric file which is used in the model to the flow using Manning's equation. HEC-GeoRAS needs land use categories represented by polygons to extract n values. Therefore, the Sentinel-2 LULC map downloaded is converted from raster to polygon in ArcGIS and assigned n values for each land use/cover type in the attribute table.

3.4.2. Geo-referencing and Projection

Geo-referencing is the process of taking a digital image and adding geographic information to the image so that GIS or other mapping software can place the image in its appropriate real-world location. A high-resolution satellite image downloaded from Google earth is used in for easy identification of different features such as buildings, river banks, roads and ineffective flow areas. After the images are downloaded, they are imported to GIS and georeferenced based on 'control points' with assigned geographic coordinates (latitude and longitude). It is necessary to modify the coordinate system of the vector and raster datasets used. Projection is changing geographical coordinate system or unknown projection to projected coordinate system. Geographic coordinate systems indicate location using longitude and latitude based on a sphere (or spheroid) while projected coordinate systems use X and Y based on a plane. Projections manage the distortion that is inevitable when a spherical earth is viewed as a flat map. The DEM, LULC map, Landsat image, administrative boundary and river shape files are projected to WGS_1984_UTM_Zone_37N, from the geographic coordinate system (GCS_WGS_1984), which is suitable for Ethiopia.

3.5. Regional Flood Frequency Analysis of the AMS

3.5.1. The R package *lmomRFA*

Regional flood frequency analysis is used to estimate flood quantiles of gauged sites and transfer information from gauged to ungauged sites. It usually involves data preparation, homogeneous region development, regional frequency distribution fitting, distribution parameter estimation, develop regional growth curve and quantile estimation for gauged and ungauged sites. All statistical computations and graphical displays in the regional frequency analysis were made using R statistical software version 4.1.2. The R package *lmomRFA* version 3.4[11] developed by Hosking, (2022), which is IFM based on L-moments according to Hosking & Wallis (1997) procedures, was used for L-moment analysis.

3.5.2. L-moment analysis

L-moments are a function of PWMs. The data is first arranged in ascending order, and then following equations are used to calculate PWMs (Leščešen & Dolinaj, 2019):

$$\beta_r = E\{X[Fx(x)]^r\} \quad \text{(3.21)}$$

where βr is the r^{th} order PWM and $F_x(x)$ is the cumulative distribution function of X. The unbiased sample estimators (b_i) of the first four PWMs are given as:

$$\beta_0 = m = \frac{1}{n}\sum_{j=1}^{n} X_j \quad \text{(3.22)}$$

$$\beta_1 = \sum_{j=1}^{n-1}\left[\frac{n-j}{n(n-1)}\right]x_{(j)} \quad \text{(3.23)}$$

$$\beta_2 = \sum_{j=1}^{n-2}\left[\frac{(n-1)(n-j-2)}{n(n-1)(n-2)}\right]X_{(j)} \quad \text{(3.24)}$$

$$\beta_3 = \sum_{j=1}^{n-3}\left[\frac{(n-j)(n-j-1)(n-j-2)}{n(n-1)(n-2)(n-3)}\right]X_{(j)} \quad \text{(3.25)}$$

where $X_{(j)}$ represents the rank of AMS, with $X_{(1)}$ being the highest and $X_{(n)}$ the lowest value.

L-moments are more convenient to use, because they can be directly interpreted as measures of location (l_1), scale (l_2), shape (t_3) and kurtosis (t_4) of distribution functions.

Table 3.5. L-moments and L-moment ratios

L-moments	L-moment ratios
$\lambda_1 = \beta_0$	$L-CV = t_2 = \frac{\lambda_2}{\lambda_1}$
$\lambda_2 = 2\beta_1 - \beta_0$	
$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$	$L-CS = t_3 = \frac{\lambda_3}{\lambda_2}$
$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$	$L-CK = t_4 = \frac{\lambda_4}{\lambda_2}$

[11] *https://CRAN.R-project.org/package=lmomRFA*

where $L - CV$ represents Coefficient of variation, $L - CS$ (L-Coefficient of skewness) and $L - CK$ (L-Coefficient of kurtosis)

3.5.3. Tests of data homogeneity of the stations

A regional homogeneity test is used to assess whether sites in the proposed region can be regarded as a homogeneous region. If the sites in the region have at-site location, scale and shape parameters, that are similar to their corresponding regional average parameters, they have the potential to be modelled by a single frequency distribution. If the region is not homogeneous, the group of stations is regarded as heterogeneous, and as such, they are not suitable for RFFA. Discordancy measure test and test using H statistics is used to for the homogeneity test.

3.5.3.1. Discordancy (Di) measure test

The discordance measure (Di) and estimates how far a given site is from the center of the group (Sine & Ayalew, 2004). Based on L-moments, it screens out data from unusual stations whose point sample L-moments are markedly different from other stations.

Let $u_i = \left[t^{(i)}, t_3^{(i)}, t_4^{(i)}, t_5^{(i)}\right]$ be a vector containing the L-moment ratios (L-CV, L-CS, L-CK and sample 5th L-moment ratio respectively) for site *i*, the discordancy measure for site *i* is:

$$Di = \frac{1}{3} N (u_i - \bar{u})^T S^{-1} (u_i - \bar{u}) \quad \text{………..............……..} \ (3.26)$$

$$\bar{u} = \frac{1}{N} \sum_{i=1}^{N} u_i \quad \text{……………………………............……} \ (3.27)$$

$$S = \frac{1}{(N-1)} \sum_{i=1}^{N} (u_i - \bar{u})\,(u_i - \bar{u})^T \quad \text{………...............……} \ (3.28)$$

where, N is the number of stations, u_i is the vector of L-moments, u is the regional average, S is the sample covariance matrix and T denotes transposition of a vector or matrix.

Table 3.6. Critical values for the discordancy statistic Di

No. of sites	Critical value	No. of sites	Critical value
5	1.333	10	2.491
6	1.648	11	2.632
7	1.917	12	2.757
8	2.140	13	2.869
9	2.329	14	2.971
		≥ 15	3

Source: Hosking & Wallis (1997)

A site is regarded as discordant if Di exceeds the critical value given in the Table 3.6. Large values of *Di* indicate sites that are the most discordant from the group as a whole and are most worthy of investigation for the presence of data errors. If a site's Di exceeds the critical value, its data are considered to be discordant from the remaining of the regional data.

3.5.3.2. Regional homogeneity test using H statistics

Heterogeneity measure H compares the inter-site variations in sample L-moments for the group of sites with what would be expected of a homogeneous region. The L-moments test for heterogeneity fits a four-parameter Kappa distribution to the regional data set, generates a series of 500 equivalent regions' data by numerical simulation and compares the variability of the L-statistics of the actual region to those of the simulated series (Hosking & Wallis,1997). The working of the test is as follow (Viglione et al., 2007):

First, find the sample L-moment ratios pertaining to the i^{th} site: these are the L-CV, the L-CS and the L-CK defined respectively as:

$$t^{(i)} = \frac{\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{2(j-1)}{(n_i-1)} - 1\right)Y_{i,j}}{\frac{1}{n_i}\sum_{j=1}^{n_i} Y_{i,j}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(2.29)$$

$$t_3^{(i)} = \frac{\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{6(j-1)(j-2)}{(n_i-1)(n_i-2)} - \frac{6(j-1)}{(n_i-1)} + 1\right)Y_{i,j}}{\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{2(j-1)}{n_i-1} - 1\right)Y_{i,j}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.30)$$

$$t_4^{(i)} = \frac{\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{20(j-1)(j-2)(j-3)}{(n_i-1)(n_i-2)(n_i-3)} - \frac{30(j-1)(j-2)}{(n_i-1)(n_i-2)} + \frac{12(j-1)}{(n_i-1)} - 1\right)Y_{i,j}}{\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{2(j-1)}{n_i-1} - 1\right)Y_{i,j}} \ldots..(2.31)$$

where, j = let Y_{ij} be the j^{th} observation in the i^{th} sample, sorted in ascending order

Secondly, the regional averaged L-CV, L-CS, L-CK and weighted standard deviation of the at-site sample L-CVs (V), weighted proportionally to the record length at each site, are defined respectively as:

$$t^R = \frac{\sum_{i=1}^{k} n_i t^{(i)}}{\sum_{i=1}^{k} n_i}, \quad {t_3}^R = \frac{\sum_{i=1}^{k} n_i {t_3}^{(i)}}{\sum_{i=1}^{k} n_i}, \quad {t_4}^R = \frac{\sum_{i=1}^{k} n_i {t_4}^{(i)}}{\sum_{i=1}^{k} n_i} \text{ and } V = \left[\frac{\sum_{i=1}^{N} n_i (t^{(i)} - t^R)^2}{\sum_{i=1}^{N} n_i}\right]^{1/2} \ldots\ldots.. (3.32)$$

where, n_i = sample size (record length) at site i, N = number of sites within the region and k = samples of observations of the same variable at different measuring sites.

On the third step, the parameters of a four-parameter kappa distribution are fitted to the regional averaged L-moment ratios t^R, ${t_3}^R$ and ${t_4}^R$, and then generate a large number simulations (Nsim) of realizations of sets of k samples. The i[th] site sample in each set has a kappa distribution as its parent and record length equal to n_i. For each simulated homogeneous set, calculate the statistic V using equation 3.32, obtaining Nsim values. On this vector of V values, determine the mean μ_v and standard deviation σ_v of the simulated regions under the hypothesis of homogeneity. Finally, a heterogeneity measure H is computed as:

$$H = \frac{V-\mu_v}{\sigma_v} \quad \text{.................................. (3.33)}$$

Three homogeneity statistics can be employed to test the variability of three different L-statistics: H_1 associated with L – CV scatter, H_2 for the combination of L − CV and L − CS, and H_3 for the combination of L − Ck and L – CS (Saf, 2009). Regions are considered as *acceptably homogenous* if H < 1, *possibly heterogeneous* if 1 ≤ H < 2 and *definitely heterogeneous* if H ≥ 2.

3.5.4. Selection of best fit regional distribution

In the selection of best fit regional distribution, the selected distribution should be widely accepted, simple and convenient to apply, consistent, flexible or robust, and less sensitive to outliers, theoretically well grounded, and documented in a guide such as that produced by the World Meteorological Organization (WMO) (Cunnane, 1989). Five frequency distributions, namely Generalized Logistic (GLO), Generalized extreme-value (GEV), Generalized normal (GNO), Pearson type III (PE3) and Generalized pareto (GPA) were selected as a candidate distribution. The selection of best fitted distribution can be through distributions plotted with LMRD and Goodness of Fit (GoF) test called the Z statics test introduced by Hosking & Wallis (1997).

3.5.4.1. Using L-moment ratio diagram

Best fitted regional distribution can be qualitatively identified using the LMRD. This is through plotting the L-CS and L-CK of standardized flow values at each station on the LMRD, together with various theoretical distribution functions. Dividing each site's AMS by the index flood gives the standardized value of each site's flood magnitude. Here, the index flood considered for each site is an estimate of at-site mean of flood magnitudes which is usually used (Saf, 2009). The distribution to which the regional L-moment ratios computed from the sample are closer to the theoretical curve is selected as a "best-fit" (Hailegeorgis & Alfredsen, 2017). The selection of a frequency distribution based on the LMRD is not quantitative and rather subjective (Lee & Kim, 2019). LMRD can be also used to identify homogeneous regions qualitatively.

3.5.4.2. Goodness-of-fit (GoF) test

In addition to visual inspection and subjective judgment of distribution selection, statistical tests are required to confirm the appropriateness of the chosen distribution and to give a certain degree of confidence in it (Eregno, 2014). The aim of a GoF measure is to identify a distribution that fits the observed data acceptably close.

Hosking and Wallis (1997) proposed a goodness-of-fit statistic, the Z-statics, based on the difference between the L − Ck of the candidate distribution and the regional average L − Ck, weighted proportionally to the sites' record lengths. The statistic is estimated by simulating a large number (500) of kappa-distributed regions with L-moment ratios equal to the regional averages and having the same number of sites and record lengths as their real-world counterparts. For the mth simulation, the L-kurtosis of the fitted distribution is calculated, τ_4^{DIST}, and after all the simulations, the bias (B4) and standard deviation (σ_4) values are found (Saf, 2009):

$$z^{DIST} = (\tau_4^{DIST} - t_4^R + B_4)\ /\ \sigma_4 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.34)}$$

where, DIST refers to a particular distribution, t_4^R =weighted regional average L-kurtosis(t_4), and B_4, σ_4 are bias and standard deviation of t_4^R respectively obtained from simulation, defined as:

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} (t_4^{(m)} - t_4^R) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.35)}$$

$$\sigma_4 = [(N_{sim} - 1)^{-1} \{\sum_{m=1}^{N_{sim}} (t_4^{(m)} - t_4^R)^2 - N_{sim} B_4^2\}]^{1/2} \quad \ldots\ldots \text{(3.36)}$$

where, N_{sim} is the number of simulated regional data sets generated using a Kappa distribution in a similar way as for the heterogeneity statistic, m is mth simulated region obtained using the distribution and $t_4^{(m)}$ is the regional average L-kurtosis and is to be calculated for the mth simulated region. A candidate distribution is selected if $|Z^{DIST}|$ is close to zero, a reasonable criterion being $|Z^{DIST}| \leq 1.64$ which corresponds to the 90% normal quantile (Hosking & Wallis, 1997).

3.5.5. Regional growth curve and quantile estimation

Regional growth curve (RGC) is a dimensionless frequency distribution common to all sites within a homogeneous region. The sites' dimensionless flood data is computed through standardizing the values. The standardized RGC which is plotted peak flow ratio, also known as growth factor/ dimensionless index flood, against frequency of exceedance or return period *T*. Mathematically,

$$X_i = \frac{Q_i}{\bar{Q}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.37)}$$

where $\bar{Q} = {}^{1}\!/_{n} \sum_{i=1}^{N} Q_i$ is mean annual peak flow (i.e., the index flood); Q_i = annual maximum flow of year i; X_i = standardized annual flow of year i. The flow quantile for T years of return period (at-site flood quantiles), Q_T, is estimated as

$$Q_T = \bar{Q}\tilde{X}_T \quad \text{.................................} \quad (3.38)$$

where $\tilde{X}_T$ = standardized quantile flood of T years return period i.e., flood frequency curve.

In this study, a return period of 2-, 5-, 10-, 20-, 50-, 100-, 200-, and 500-year floods are estimated which will be used in flood hazard and risk analysis. Estimation of flood values for a higher return period increases the uncertainty and needs available longer flow data.

3.5.6. Performance evaluation of the frequency distribution

3.5.6.1. Quantile-Quantile (Q-Q) plots

The quantile-quantile (Q-Q) plot is a graphical technique for determining if two data sets come from populations with a common distribution by plotting the quantiles of the first data set against the quantiles of the second data set. In this plot, theoretical quantiles are plotted in the horizontal axis and sample quantiles are plotted on the vertical. If the scatter points in the plot lie in a straight line, then both the random variable have same distribution, else they have different distribution.

3.5.6.2. Accuracy of the regional growth curve

Results obtained by statistical analysis are inherently uncertain, and for the results to be maximally useful some assessment of the magnitude of uncertainty should be made. A reasonable approach is to estimate the accuracy of estimated quantiles by Monte Carlo simulation. The region used as the basis for simulation should be chosen to have the same number of sites, record length at each site and regional average L-moment ratios as the actual data. The L-moment ratios at the individual sites should be chosen to yield a region whose heterogeneity is consistent with the heterogeneity measures calculated from the data (Hosking & Wallis, 1997). For each simulation, errors in the simulated growth curve and quantiles were calculated and then the root mean square error (RMSE), and 90% error bounds estimated. The more that RMSE values come closer to 0, both the precision of the value estimated and the method are better (Boucefiane & Meddi, 2019). To approximate the RMSE of the estimators, quantiles can be averaged over all M repetitions. Therefore, the relative mean square error (RMSE) can be designated as percentages of the site i quantile estimator by:

$$R_i(F) = \left\{ \frac{1}{M} \sum_{m-1}^{M} \left[\frac{\hat{Q}_i^{(m)}(F) - Q_i(F)}{Q_i(F)} \right]^2 \right\}^{0.5} \quad \text{..................} \quad (3.39)$$

where, $Q_i(F)$ is estimate of the T-year event at-site i, $\hat{Q}_i^{(m)}(F)$ is at-site i quantile estimate

The regional average relative RMSE of the estimated quantile with N number of sites is

$$R^R(F) = \frac{1}{N}\sum_{i=1}^{N} R_i(F) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.40)}$$

3.5.7. Prediction of flood quantiles for ungauged catchments

For the prediction of flood quantiles in ungauged basin, at least two regressions curves are required. In the first curve, the mean annual peak flow, also called index flood, is plotted against some physical parameters, e.g., basin area, basin slopes, river length and the second curve is the standardized RGC. Thus, a frequency curve can be developed for any other ungauged catchment in the homogeneous region from these two curves (Obinna & Ben, 2020).

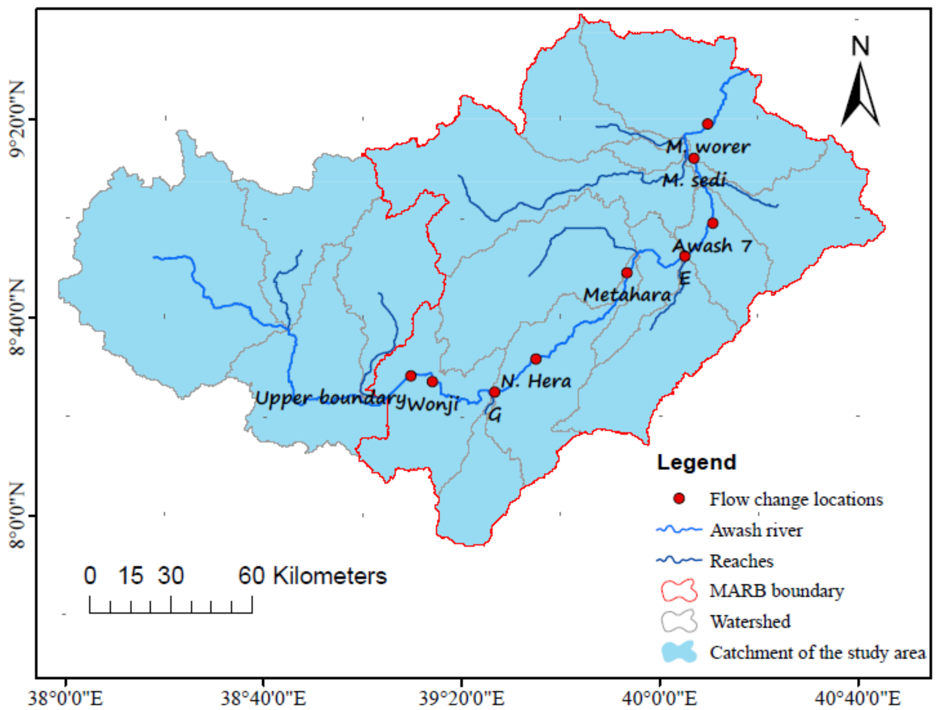


Figure 3.9. Flow change locations considered in HEC-RAS simulation

HEC-RAS needs input flow at flow change locations. Nine flow change locations are considered in which the two sites, G and E as shown in Figure 3.9, are ungauged. A relationship between the index flow ($\bar{Q}$) of gauged sites and their catchment area (A) is used to estimate $\bar{Q}$ of ungauged sites. The regression equations of A versus $\bar{Q}$ fitted to the exponential, linear, logarithmic, polynomial and power equations and R^2 values are computed.

Functions	exponential	linear	Logarithmic	polynomial	power
R^2	0.38	0.67	0.67	0.74	0.73

Table 3.7. At-site index flood and their respective catchment areas

	Mean (m^3/s)	Area (km^2)
Melka sedi	212.89	21520
Melka werer	292.79	31183
Wonji	150.92	11690
Metahara	164.70	16417
Nura Hera	225.59	14173
Abomussa	16.80	140
Awash 7 Kilo	347.77	19111

It is found that the polynomial equation has the highest R^2 value and is selected for deriving the regression equation. The plot of the regression with polynomial equation is shown in Figure 3.10. The $\bar{Q}$ (m^3/s) for the ungauged sites with A in km^2 is determined as:

$$\bar{Q} = -3 * 10^{-7}A^2 + 0.0173A + 9.6145 \quad (3.41)$$

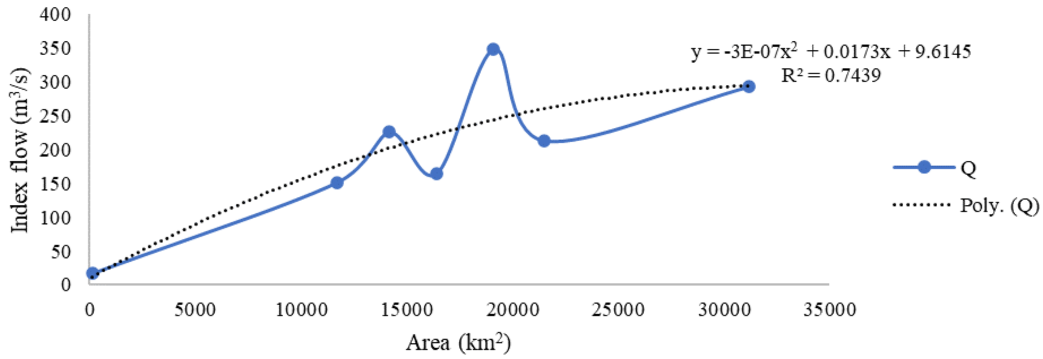


Figure 3.10. Index flood versus catchment area of gauged sites with fitted polynomial equation

Using equation 3.41, index flood for site E and G is found to be 239.05 and 189.08 respectively. Multiplying those by the estimates of growth factor (X_T) will give us quantile estimate (Q_T) for the different return periods. The estimates are used as a flow change values for the two sites during HEC-RAS processing.

3.6. Inundation Mapping

This section is intended on how to execute the actual work of flood inundation mapping using the integration of HEC-GeoRAS and HEC-RAS for the output of flood extent, depth and velocity on the river channel and its floodplain for a return period of 2-, 5-, 10-, 20-, 50-, 100-, 200- and 500-years. The inundation mapping broadly has three steps namely, pre-RAS processing, HEC-RAS processing and post-RAS processing.

3.6.1. Pre-RAS processing

Pre-RAS processing is done in a HEC-GeoRAS environment that includes creating the RAS GIS Import File and exporting GIS data to HEC-RAS. The RAS Geometry menu is for pre-processing geometric data for import into HEC-RAS. This is achieved through four steps:

- Terrain model processing
- Creating RAS layers
- Extraction of 3D spatial data
- Creating and exporting RAS GIS import files

3.6.1.1. Terrain model processing

The extraction of 3D data and their position for the created RAS layers are from TIN which is a terrain model commonly used and preferred in hydraulic modeling. The TIN is generated from ALOS PALSAR DEM of 12.5m resolution in ArcGIS (3D Analyst). TIN is preferred than DEM for hydraulic modelling as it varies in spatial resolution, well suited to represent linear features and higher computation speed.

3.6.1.2. Digitizing RAS layers

RAS layers are created after digitizing RAS GIS Import Files such as Stream Centerline, Bank Lines, Flow Path lines, Cross-sectional cut lines, Manning's n value, Ineffective flow areas and defining bridges. A georeferenced Google Earth image is used to aid in defining the location of the structures (buildings and bridges) and ineffective flow areas. The Stream Centerline, running through the center of the river, was digitized first and then Bank Lines and Flow Paths respectively. Bank lines are digitized at the right and left edge of the main channel. The Flow Path lines are created in the center of main channel, left overbank and right over bank.

Cross sections are digitized from left to right looking downstream and perpendicularly to the river's flow direction to cover the entire floodplain. We used the average bank full depth (D) to be equal to 2.8m, average bed slope (S_O) found to be 0.0029. Thus, the spacing between cross-sections (Dx) value of 141m is found. A total of 952 cross-sections were created, which is about 304 kms in length, with the average distance of 141m among them and a length between 453 to 5000 meters.

To create Bridges layer, four cross-sections are digitized at expansion zone, contraction zone and toe of the bridge fill. The needed attributes are name of the river and reach, station number, distance to the nearest section, and width as shown in Table 3.8. These attributes are entered for the three bridges namely, Awash, ELO/Awash and Awash - Ras Mekonnen WM, see section 3.2.1 for collected data of bridges.

Table 3.8. Bridges digitized in HEC-GeoRAS

Shape *	OID *	Shape_Length	BC2DID	HydroID	River	Reach	Station	USDistance	TopWidth	NodeName
Polyline Z	1	2992.247726	7106	7109	Awash River	Awash Reach	243178.7	7.4	6	Awash
Polyline Z	2	1451.731904	7107	7110	Awash River	Awash Reach	203673.6	4.18	4.4	ELO/Awash
Polyline Z	3	1741.632515	7108	7111	Awash River	Awash Reach	112764.3	7.8	8.9	Awash-Ras-ekonnen WM

Non-conveyance portion of the floodplain is defined by digitizing Ineffective Flow Areas. The impact of large buildings in the floodplain is represented by digitizing polygons of buildings as Blocked Obstructions, observed on the Google Earth aerial photograph. Areas behind bridge abutments representing contraction and expansion zones are considered as ineffective flow areas.

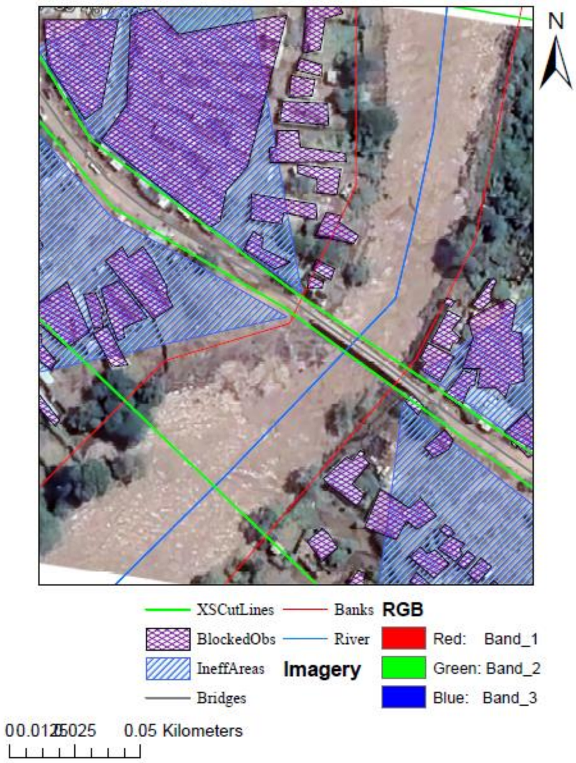


Figure 3.11. Created RAS layers

The final task before exporting the GIS data to HEC-RAS geometry file is assigning Manning's n values to individual cross-sections. The LULC map of Sentinel-2 raster data is converted to polygon and manning's n value is assigned for each land LULC type according to Chow (1989). See Appendix D for n values of the main channel and floodplains. Depending on the intersection of cross-sections with land use polygons, Manning's n values are extracted for each cross-section. The corresponding n values for each LULC type are shown in Table 3.9.

Table 3.9. Manning's n values for the different LULC classes

LULC class	n values			Description in Chow (1959)
	min.	normal	max.	
Water Bodies	0.033	0.040	0.045	Natural streams: clean, winding, some pools and shoals
Dense vegetation	0.080	0.100	0.120	Trees with heavy stand of timber, a few down trees, little undergrowth, flood stage below branches
Flooded Vegetation	0.050	0.070	0.080	heavy growth of sprouts or sluggish reaches, weedy, deep pools
Crops	0.03	0.04	0.05	mature field crops or cultivated areas
Bare Ground	0.030	0.040	0.050	cleared land with tree stumps, no sprouts
Range Land	0.025	0.035	0.050	Pasture with high grass
Built Area		0.013[12]		Artificial surfaces

3.6.1.3. Creating and exporting RAS GIS import files

The last step is to create a GIS import file for HEC-RAS so that it can import the GIS data to create the geometry file. The layers to be exported are required surface (the TIN), Required Layers (river Centerline, XSCutmLines and XSCut Lines Profiles), Optional Layers and optional tables. Once the RAS layers have been created, HEC-GeoRAS extracts the elevation values at every intersection of the RAS layers with the TIN to create the geometric data for HEC-RAS. This elevation data together with all the captured RAS layers were exported to an exchange file which will be imported into HEC-RAS during processing.

3.6.2. HEC-RAS processing

3.6.2.1. Importing and completing geometric data

i. Importing geometric data

For HEC-RAS processing, first, geometric data files exported from HEC-GeoRAS are imported to HEC-RAS from the geometry window of HEC-RAS. The cross sections were observed as station numbers in the cross-section editor window of HEC-RAS. In HEC-RAS, cross-sections are represented by numbers which shows the distance measured from d/s to u/s.

[12] (Papaioannou et al., 2018)

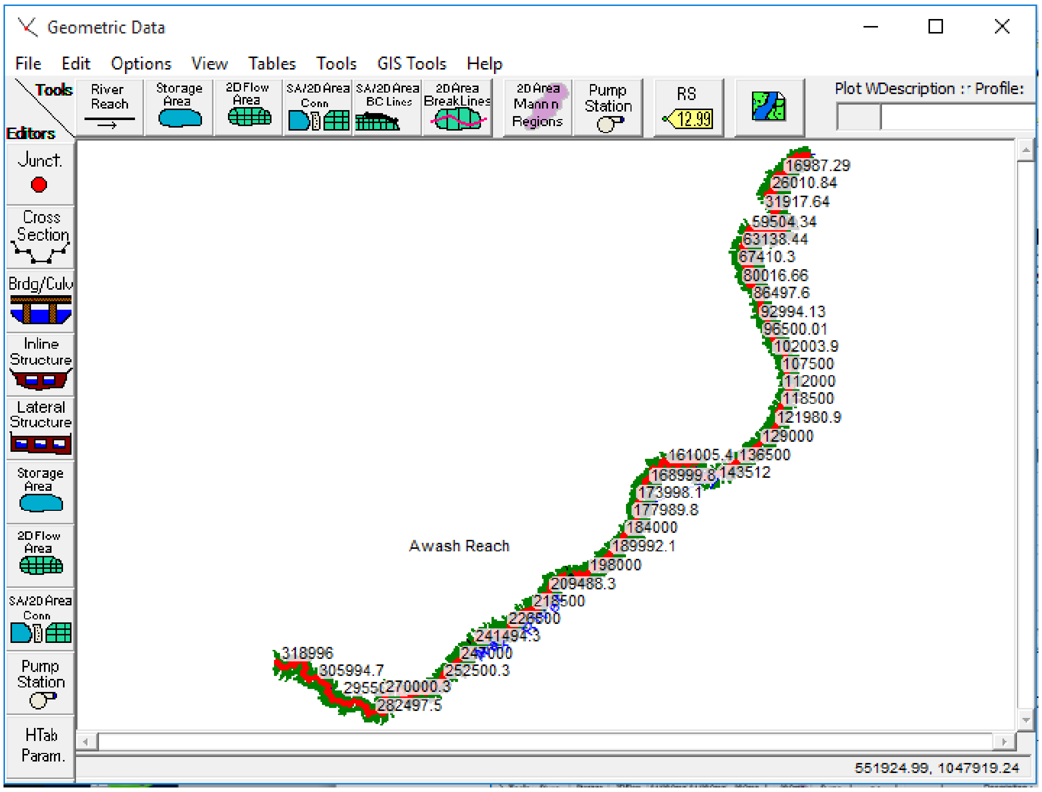


Figure 3.12. Geometric data imported to the HEC-RAS geometric editor

Graphic XS Editor tool in Geometric editor can be used to perform quality check such as moving bank stations, change the distribution of Manning’s n, add/move/delete ground points, edit structure etc. HEC-RAS has the tendency to place water into the lowest part within the cross-section profiles. Such inundations has to be manually removed (Vojtek & Vojteková, 2016). Each cross-section in HEC-RAS has information on location, elevation profile, bank locations, Manning’s n Values and distance to the next downstream cross-section as shown in Figure 3.12.

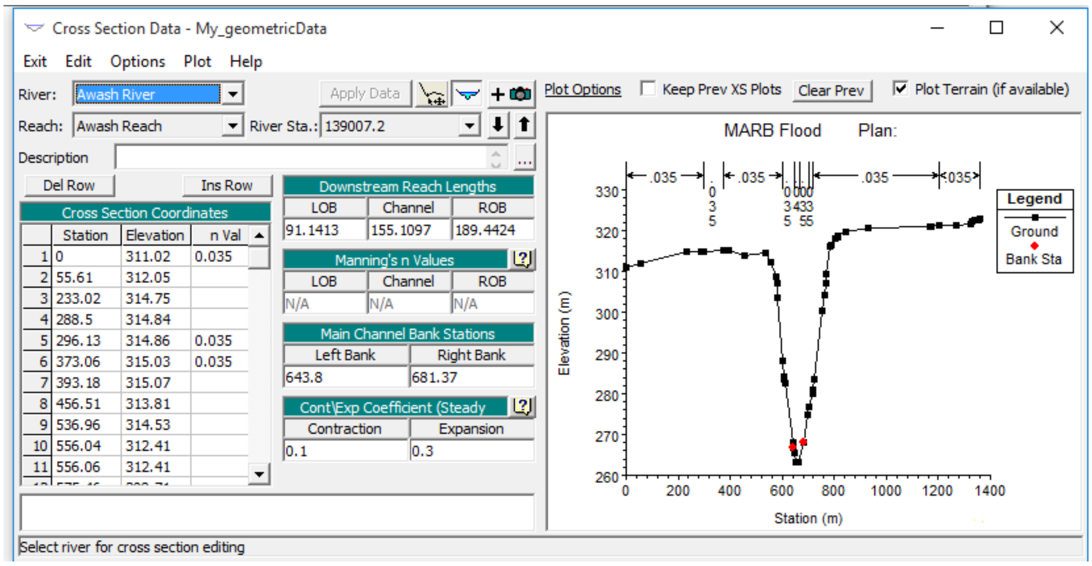


Figure 3.13. Extracted cross-sectional data

ii. Editing Bridge Structure Data

Bridge information from HEC-GeoRAS should be completed before running HEC-RAS. To add bridge to the model, the following bridge data are required (Brunner, 2020): Bridge Deck, Sloping abutments (Optional), Piers (Optional) and Bridge modeling approach. Bridge/Culvert editing button in HEC-RAS enables us editing the geometry and structural details of the bridge.

The Deck editor in the Bridge/Culvert editor is used to enter information such as distance from u/s cross-section, width of deck/roadway in the direction of flow, weir coefficient, upstream stationing, high chord and low chord, downstream stationing and weir crest shape. The high chord column defines the top elevation of the bridge, and the low chord defines the elevation of the bottom of the bridge deck. In the absence of any values for low chord, the elevations on the cross-section profile are used as the low chord values. If value for low chord is 0 then, bridge will be filled in to the ground level.

After Deck/roadway editor is completed, pier data such as number of piers, u/s and d/s XS stationing, pier centerline spacing and pier width, are entered to the Bridge Design Editor.

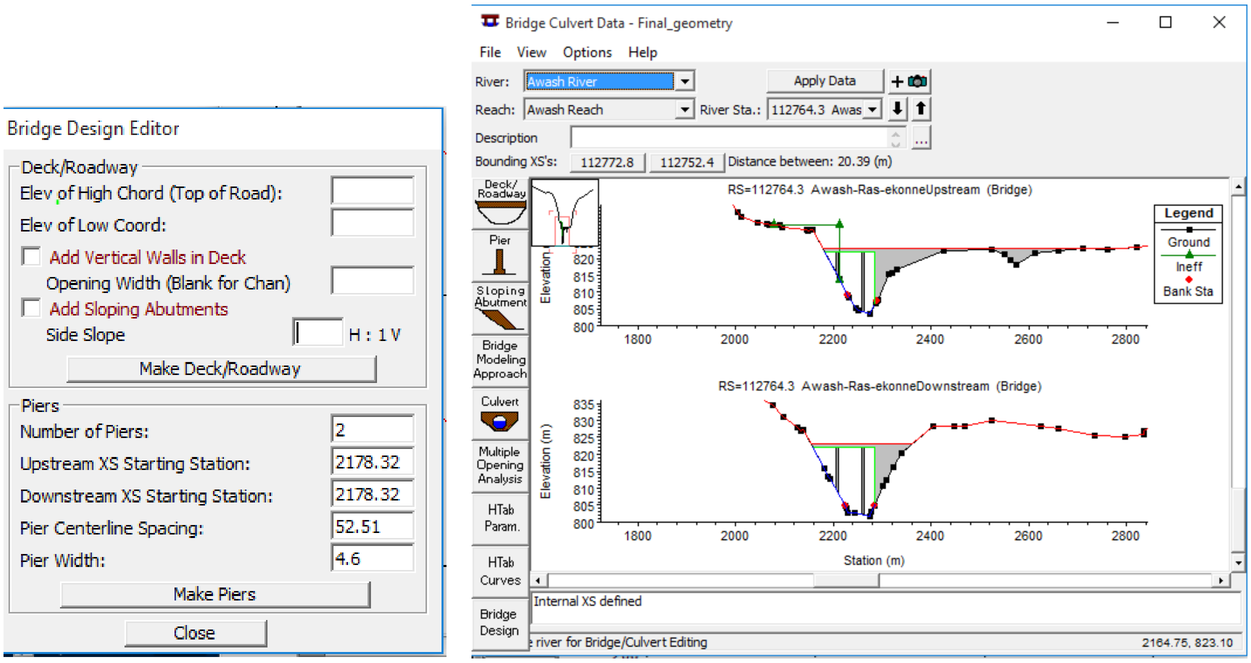


Figure 3.14. Adding piers data (Awash - Ras Mekonen WM bridge)

The Bridge modeling approach is used to define how the bridge will be modeled and to enter any coefficients. It consists defining which methods the program will use for low flow (water surface below the highest point on the low chord of the bridge opening) computations and high flow (flow at or above the maximum low chord) computations. If either the Momentum or Yarnell methods are selected, it needs a value for the pier loss coefficient while WSPRO method needs the user to enter additional information required for the method. Energy and momentum methods are in general applicable to the widest range of bridges. The energy method is selected in that it only requires only Manning’s n value for friction losses and contraction and expansion coefficients for transition losses. The three bridges are RC type and assigned n = 0.02 Chow, (1989).

iii. Expansion and contraction coefficients

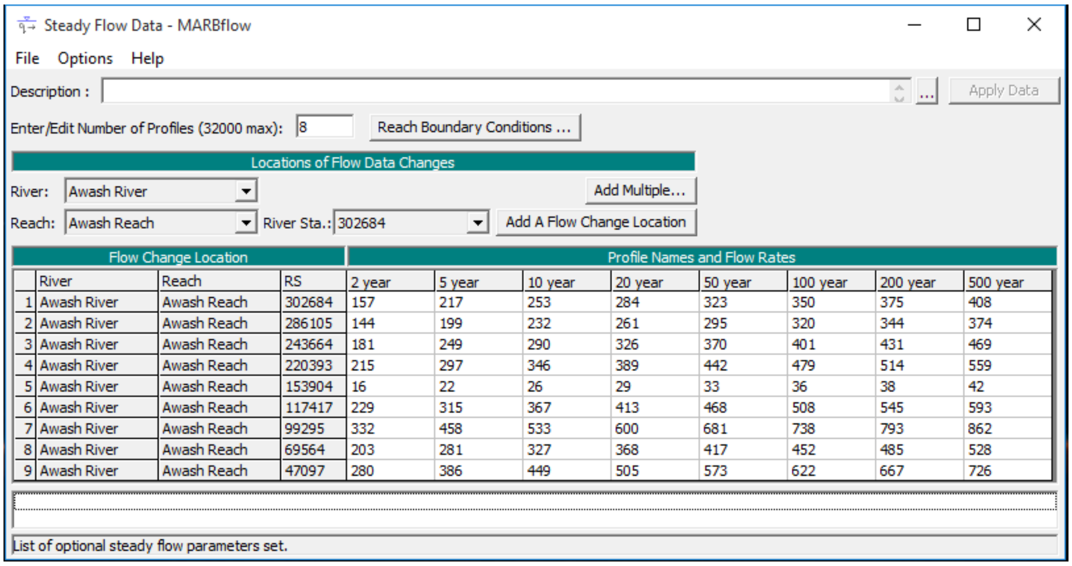
Contraction or expansion of flow due to changes in the cross-section is a common cause of energy losses within a reach (between two cross-sections). Whenever this occurs, the loss is computed from the contraction and expansion coefficients. Where the change in effective river cross-section is small and the flow is subcritical, coefficients of contraction and expansion are typically in order of 0.1 and 0.3 respectively. Expansion and contraction coefficients of 0.5 and 0.3 respectively are assigned for abrupt changes in effective cross-section area, such as at bridges, as recommended by (Brunner, 2020).

3.6.2.2. Running Steady flow Analysis

Once all of the geometric data are entered, then required steady flow data are entered for HEC-RAS simulation. Considering required model outputs, level of detail and the expected level of accuracy of the results a 1D steady state is selected for modelling. 1D steady flow models are easiest to use and understand the results. The required data to be entered are the number of profiles to be calculated, flow regime selected and boundary conditions. Sub-critical flow regime is selected which is characterized by relatively deep depths and slower velocities.

From the four boundary conditions of the 1D steady state flow, normal depth is selected as a downstream boundary condition and Q_T, estimated through RFFA, as an upstream boundary condition. Normal depth as a boundary condition needs to enter an energy slope. HEC-RAS will use that value to compute depth using Manning's equation. The energy slope can be approximated by measuring the slope of the reach downstream of the modeled reach. The river profile has an elevation difference of 1550m and a horizontal difference of 320,000m between u/s and d/s end of the river. Thus, the energy slope (S) is approximated as 0.005.

Discharge information is required at each cross-section in order to compute water surface profiles. Profiles are flow values entered into HEC-RAS simulation. Eight profiles for a return period of 2, 5-, 10-, 20-, 50-, 100-, 200- and 500-years are entered as profiles. They are estimated maximum flood discharges (Q_T) for the gauged and ungauged sites previously. At least one flow value must be entered for each reach. Once a flow value is entered at the u/s end of the reach, it is assumed that the flow remains constant until another flow is encountered with the same reach (Brunner, 2020). Thus, a flow change location of 9 sites is used to vary the flow through the river channel.



	River	Reach	RS	2 year	5 year	10 year	20 year	50 year	100 year	200 year	500 year
1	Awash River	Awash Reach	302684	157	217	253	284	323	350	375	408
2	Awash River	Awash Reach	286105	144	199	232	261	295	320	344	374
3	Awash River	Awash Reach	243664	181	249	290	326	370	401	431	469
4	Awash River	Awash Reach	220393	215	297	346	389	442	479	514	559
5	Awash River	Awash Reach	153904	16	22	26	29	33	36	38	42
6	Awash River	Awash Reach	117417	229	315	367	413	468	508	545	593
7	Awash River	Awash Reach	99295	332	458	533	600	681	738	793	862
8	Awash River	Awash Reach	69564	203	281	327	368	417	452	485	528
9	Awash River	Awash Reach	47097	280	386	449	505	573	622	667	726

Figure 3.15. Input peak flow discharges entered for simulation

After successful simulation, HEC-RAS output is exported as “RASExport.sdf” file format by selecting “Water Surfaces” and “Export Velocity Distribution” options respectively for inundation depth and velocity to be processed in GIS/HEC-GeoRAS for flood inundation and hazard mapping. The velocities are exported based on the velocity distribution computed in HEC-RAS with one point at the centroid of a flow distribution location.

3.6.2.3. Post-RAS processing

After the HEC-RAS modeling is completed, the hydraulic analysis results are imported from HEC-RAS into ArcMap to create the inundation maps. The RAS Mapping menu is for post-processing exported HEC-RAS results.

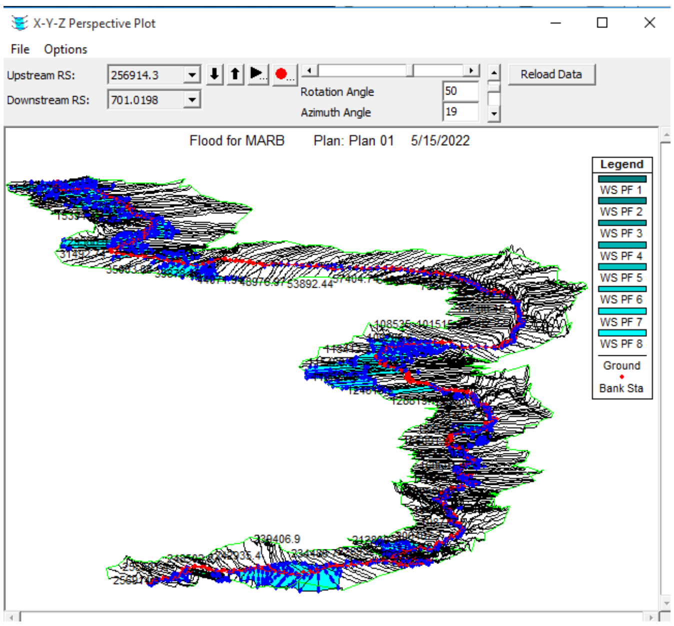


Figure 3.16. 3D perspective plot of the river system

To generate floodplain shape files, the GeoRAS extension is used to first create a water surface TIN for each of the flood events. For automatic delineation of floodplains based on the data contained in the RAS GIS output file and the original TIN layer, the water surface TIN is overlayed with the original TIN. Intersection of TIN terrain model and TIN model of water levels generates the water depth. Velocity map is generated by interpolating the detailed velocity data exported from the HEC-RAS using “Flow Distribution Simulation” option.

3.6.3. Calibration and validation of HEC-RAS

The model calibration process is required inorder to understand if the model is capable of reproducing past floods, and if it can predicting future flood events with confidence (Brunner et al., 2020). Model Validation analyzes the model capacity to simulate measured data other than those used for the calibration within standard exactness, commenced after calibration (Desalegn & Mulu, 2021).

3.6.3.1. Calibration

Usually, 1D steady flow model is calibrated by comparing water surface elevations of gauged data and any available high-water marks. Adjustments can generally be made to Manning’s n values, contraction and expansion coefficients and any hydraulic structure coefficients (weir coefficients, bridge and culvert coefficients, etc.) (Brunner et al., 2020). In this study, calibration was done by adjusting hydraulic roughness of the surface, because of the accessibility of the values, to recreate a rating curve. The rating curve was constructed by combining discharge and water level data of Awash 7 Kilo station obtained from MoWE. The land use types were the basis for estimating the roughness coefficients for river and its floodplain. For the calibration here, it is tried to adjust n values of the river channel and leaving the other’s LULC n as it is. n value for the river channel was varied between 0.033 and 0.045 as suggested in Chow, (1959).

Table 3.10. Generated rating curve

Q (m^3/s)	WSE Observed (m)	Q (m^3/s)	WSE Observed (m)	Q (m^3/s)	WSE Observed (m)
721.50	4.99	414.55	3.68	213.79	2.57
700.70	4.98	392.15	3.57	206.00	2.52
598.20	4.5	378.17	3.50	197.05	2.46
562.40	4.35	374.22	3.48	176.60	2.32
487.17	4.02	350.88	3.36	157.13	2.18
478.37	3.98	311.60	3.15	141.23	2.06
443.84	3.82	272.64	2.93	121.17	1.90
433.28	3.77	231.11	2.68	109.32	1.80
418.68	3.7	218.46	2.60	-	-

The performance of the model is examined using statistical measures like Nash-Sutcliffe model efficiency (NSE), percent bias (PBIAS), and coefficient of determination (R^2), as Kim et al., (2020) defined:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{obs\,i} - Q_{sim\,i})^2}{\sum_{i=1}^{n}(Q_{obs\,i} - Q_{sim\,i})^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.42)$$

$$PBIAS = [\frac{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{sim}\right)*100}{\sum_{i=1}^{n} Q_i^{obs}}] \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.43)$$

$$R^2 = \frac{\sum((Q_{sim(t)} - \overline{Q_{\text{sim}}})(Q_{obs(t)} - \overline{Q_{\text{obs}}}))^2}{\sum(Q_{sim(t)} - \overline{Q_{\text{sim}}})^2 \sum(Q_{obs(t)} - \overline{Q_{\text{obs}}})^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.44)$$

where, $Q_{sim(t)}$ and $Q_{obs(t)}$ are the simulated and observed values (e.g., discharges) at time step t, and $\bar{Q}_{sim}$ and $\bar{Q}_{obs}$ are the simulated and observed average discharges

Table 3.11. Performance ratings for statistics

Performance rating	NSE	PBIAS	R^2
Very good	$0.65 < NSE \leq 1.00$	$PBIAS < \pm 15$	$0.65 < R^2 \leq 1.00$
Good	$0.55 < NSE \leq 0.65$	$\pm 15 \leq PBIAS < \pm 20$	$0.55 < R^2 \leq 0.65$
Satisfactory	$0.40 < NSE \leq 0.55$	$\pm 20 \leq PBIAS < \pm 30$	$0.40 < R^2 \leq 0.55$
Unsatisfactory	$NSE \leq 0.40$	$PBIAS \geq \pm 30$	$R^2 \leq 0.40$

Source: USACE, (2018)

3.6.3.2. Validation

For the validation of HEC-RAS, flood extent of August 28, 1996 flood event from Landsat 4-5 TM image was compared with inundation extent of HEC-RAS simulation output for the flow of that day. The image of the day with little cloud cover was obtained as indicated with white color near the basin boundary as shown in Figure 3.17. This is due to the reason that Normalized Difference Water Index (NDWI) value considered the cloud cover as a water surface. But our AOI, the floodplain, in the image is found free of cloud coverage. This is one major limitation of optical sensors that they cannot penetrate through cloud cover. NDWI, having a value of [-1, 1], for the Landsat image is calculated in ArcGIS and a value greater than 0.2 is used to delineate the water surface (EOS Data Analytics inc., 2021). NDWI can be calculated as (Tamiru & Dinka, 2021):

$$NDWI = \frac{(Green - NIR)}{Green + NIR} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.45)$$

and for Landsat 4-5 TM, it is calculated as

$$NDWI = \frac{B3 - B5}{B3 + B5} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.46)$$

where B3 and B5 are bands of the Landsat image and NIR is near infrared

Available measured flow data of four stations during August 28, 1996 flood event were entered into HEC-RAS steady flow analysis and the simulated results are compared with the inundation extent of the Aug. 28, 1996 flood derived from Landsat 4-5 TM image, after it is clipped to the AOI. The four stations are Wonji, Metehara, Melka Sedi, Melka worer and Awash 7 kilo with a flow of 227.10, 214.11, 305.81, 312.51 and 473.99 m^3/s respectively. The comparison of inundation maps can be done using measure of fit, accuracy, precision, and recall (Quirós & Gagnon, 2020), agreement fit measures, under-prediction measures, and over-prediction measures (Haque et al., 2020), inundation area, average inundation width, and the F statistic (Sarhadi et al., 2012), overlapping percentage (Tamiru & Dinka, 2021) and percent of inundation area coverage with respect to the study area (Roy et al., 2021).

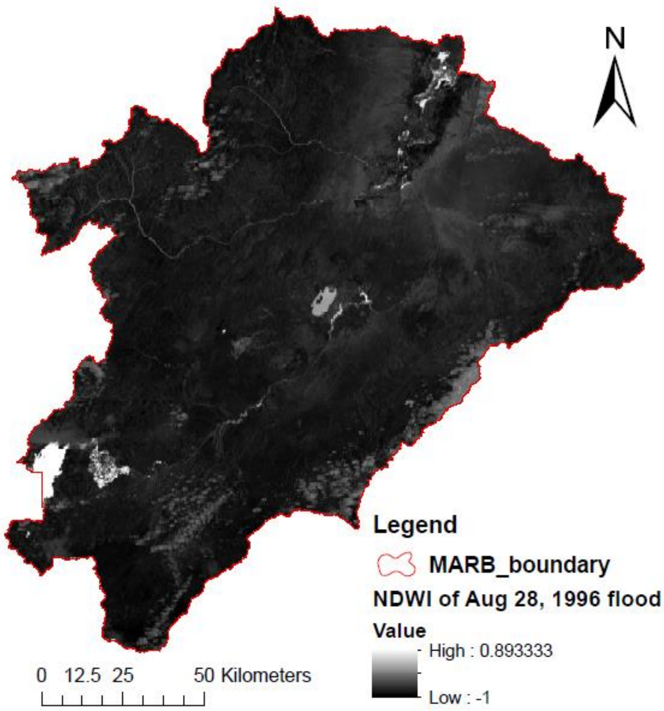


Figure 3.17. Landsat 4-5 TM used according to NDWI

Overlapping percentage and the F statistic are used in this study. First, results of NDWI $\geq$ 0.2 and HEC-RAS are converted to polygons and then their area coverage is calculated. Finally, overlapping percentage and F statistic is calculated as follows.

$$\text{Overlapping percentage (\%)} = \frac{Layer\ 1}{Layer\ 2} \times 100 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(3.47)}$$

where, Layer 1 is the area of the inundation map generated in HEC-RAS and Layer 2 is the area of water body delineated in NDWI from the flood event.

The F statistic is calculated as

$$F = \left(\frac{A_{op}}{A_o + A_p - A_{op}}\right) \times 100 \quad \text{.....................................} (3.48)$$

where, A_o is the observed area of inundation, A_p is the predicted area of inundation and A_{op} is the area that is both observed and predicted as inundated. The F statistic varies between 0 to100 showing no overlap and total overlap respectively.

3.7. Flood hazard, Vulnerability and Risk Analysis

Risk is computed as a product if hazard and vulnerability. For large floodplains, it is recommended to model flood influence at a scale of large administrative units. This study considered *kebele* (4^{th} administrative level in Ethiopia*)* and town boundaries, as a land unit and flood hazard and risk assessment were done for each *kebele* and *Woredas* (3^{rd} administrative level) for easy understanding and implementation of the results. Six *woredas* (Adama, Boset, Fentale, Awash, Amibara and Dulecha) and eighty *Kebeles* and towns, with a total area of 526004 ha, were considered in the assessment. The computation of hazard factor (H_F), vulnerability factor (V_F) and risk factor (R_F) for each land unit were made according to Tingsanchali & Karim, (2010).

3.7.1. Flood hazard Analysis

Flood hazard assessment requires determination of flood characteristics and probability of occurrence of each flood. In the present study, the probability of occurrence of each flood is determined through RFFA and the flood parameters considered are the flood extent, depth and flow velocity (Elkhrachy et al., 2021; Vojtek & Vojteková, 2016). The flood intensity (FI) is a measure for the damaging effects of the flow. A hazard index, H_I, was introduced to represent intensity of the flood hazard which is a relative scale defined in function of flow depths and flow velocities. For velocities smaller than 1m/s the value of the flood intensity is equal to the flow depth. For higher velocities the intensity is defined as the product of flow depth and velocity. Mathematically (Vojtek & Vojteková, 2016),

$$\text{FI} = \begin{Bmatrix} 0 \rightarrow d = 0m \\ d \rightarrow d > 0m, v \leq 1m/s \\ d.v \rightarrow v > 1m/s \end{Bmatrix} \quad \text{………..…………..…...……} (3.49)$$

Karim & Chowdhury, (1995) investigated three alternative scales for H_I as H_I increases linearly, the rate of increase in H_I is linear and H_I increases geometrically and observed that linear scale produces R_I (risk index) maps consistent with maps based on historicall recorded damages. So the scale of H_I based on linear increase of H_I is selected as shown in Table 3.12.

Table 3.12. Flood hazard classification

FI[13]	Hazard category	H_I	Description (Beffa, 1998)
< 0.5	Low	1	* Persons are not endangered * Limited damages of buildings are possible
0.5 to 2.0	Medium	2	* Persons outside of buildings are endangered * No abrupt destructions, damages of buildings are possible
>2.0	High	3	* Persons inside and outside of buildings are endangered * Abrupt destructions of buildings are possible

For the practical application of predicted results, the hazard is estimated for a land unit and represented by a number, the hazard factor, H_F. The H_F for each land unit was computed from the percentage area under each hazard category and corresponding hazard factor is computed using the following equation:

$$H_F(i) = \frac{\sum_{j=1}^{N_D} A(i,j) H_I(j)}{\sum_{j=1}^{N_D} A(i,j)} \qquad (3.50)$$

where, i is the land unit identification number and j represents the hazard category; N_D is the total number of hazard categories; and A(i,j) is the area under hazard category j in land unit i

3.7.2. Flood vulnerability analysis

For flood vulnerability analysis, at first, the elements at risk should be identified. This study identified elements at risk by analyzing satellite image of Sentinel-2 in the GIS environment and hence obtaining LULC map. The crop land and built-up area were selected as elements at risk. Assumptions was made that all the buildings within a particular kebele have equal weights, crop value is same for all types and population density and economic activities are uniform and proportional to built-up area.

For vulnerability computation, at first, a land use map is generated. Then, similar to the hazard index, a vulnerability index (V_I) is introduced to represent the land-use categories. The vulnerability factor, V_F, for a land unit was then computed using the following mathematical expression:

$$V_F(i) = \frac{\sum_{k-1}^{N_L} V_I(k) A(i,k)}{\sum_{k-1}^{N_L} A(i,k)} \qquad (3.51)$$

[13] Classification used in: (Beffa, 1998; Elkhrachy et al., 2021; UNDRO, 1991; Vojtek & Vojteková, 2016)

where i is a land unit identification number and k represents the land use category, N_L is the total number of land use categories; and A(i,k) is the land-area percentage for category k in cell i.

The quantification of vulnerability depends on the susceptibility of 'elements at risk' (Baky et al., 2019). The scale of V_I was established by analyzing the relative monetary value of the land use categories. LULC categories other than agricultural and built-up area are considered negligible in comparison with other land uses and a V_I of 0 was assigned for them. According to the interview conducted, it is roughly estimated that cost of built-up area is three-fold of crop land in the study area. Therefore, V_I value of 1 and 3 are given respectively for crop land and built-up area.

3.7.3. Flood risk Analysis

The magnitude of risk for a land unit was estimated by a risk factor, R_F, which is computed as a product of the hazard factor and the vulnerability factor (i.e., RF = $H_F V_F$). To establish the physical inference of those values, the scores were normalized on a scale of 0 to 100 using the following equation:

$$R_F^{New}(i) = \frac{R_F(i) - (R_F)_{min}}{(R_F)_{max} - (R_F)_{min}} 100 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (3.52)$$

The land units then categorized into low risk, moderate risk, high risk and very high risk based on R_F^{New} values having a R_F^{New} value of [1-25], (25,50], (50-75] and (75-100] respectively. This equal interval classification is used by studies like Chowdhury & Karim, (1996), Ogato et al., (2020) and Tingsanchali & Karim, (2005). It provides relatively uniform accuracy of representation of the risk zones (Tingsanchali & Karim, 2005). Land units that scored less than 1.0 were considered as risk-free zone. The reason is that a score less than one is quite insignificant on scale of 0 to 100 (Tingsanchali & Karim, 2010).

3.8. Methodology Framework

The general methodology adopted in this study follows literature review, raw data collection and analysis, flood quantile estimation using regional flood frequency analysis, hydraulic model simulation and risk mapping and assessment to reach the objective of the study. The methodology used is summarized under the flow chart below.

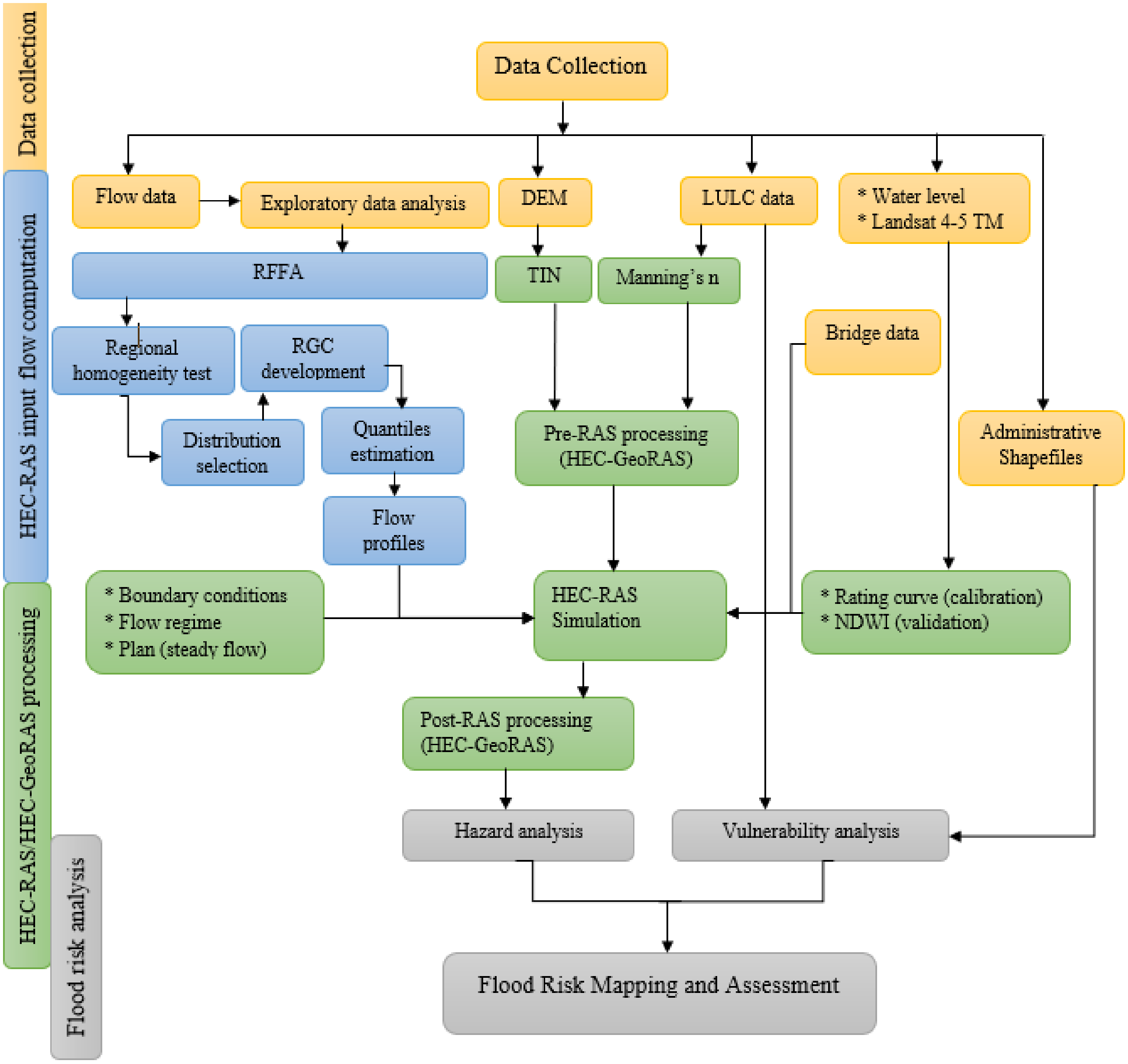


Figure 3.18. General conceptual framework of the methodology

4. RESULTS AND DISCUSSION

A methodology is applied on the third section of this document in “materials and methods” for flood risk mapping and assessment on MARB. After the collection of the required datasets, an exploratory data analysis was conducted to test if the data is enough in quality and quantity to achieve the goal of the study. Data of seven flow stations are screened for analysis with AMS which fulfils the assumptions of regional flood frequency analysis that they are independent and identically distributed (iid) based on tests of homogeneity, independence, trend and stationarity. Also, data consistency and existence of outliers were checked and the outlier found is treated according to Cunnane, (1989). Now in this part, all the necessary results have been shown and discussed towards the objective of the research.

4.1. AMS Data Quality Check Results

4.1.1. Data consistency

Double mass curve (DMC) analysis is used to check the consistency of the AMS. The assumption that a constant ratio exists between a given record and a group of records in the area may not be valid in stream flow records. If the difference in slope is low, data is considered as consistent. Using AMS data collected for this study, it is unable to conduct consistency check using double mass curve method due to a lot of reasons.

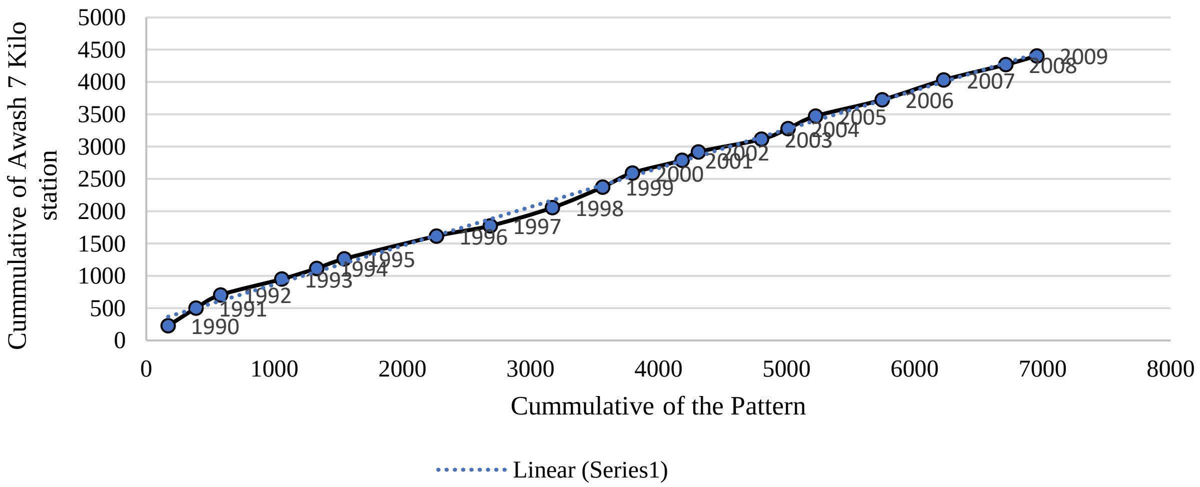


Figure 4.1. Double mass curve for Awash 7 Kilo station

According to Searcy & Hardison,(1960), stations having same length of record and period shall be used. Only stations of Wonji and Melka Sedi have full 24 years of AMS data in common. Nura Hera station have only 5 years in common and the remaining with 18 years. Poor correlation between the variables can prevent detection of inconsistencies in a record, but an increase in length of record tends to offset the effect of poor correlation (Searcy & Hardison, 1960). Some stations are poorly correlated and no long years of data is available. Searcy & Hardison, (1960) recommends to ignore any break that persists for less than 5 years and reasoned that year-to-year breaks are due to chance caused by the inherent variability in hydrologic data. In our case, frequent change in slope (less than five years) for all stations is observed and thus it is not suitable for consistency check. Example for DMC of Awash 7 Kilo station is shown in the Figure 4.1.

4.1.2. Statistical test results

The AMS data quality was also checked for iid assumptions using statistical tests of Mann-Kendall trend test, W-W test of independence and stationarity and for Buishand's homogeneity test. The Mann-Kendall trend test showed that Ho is rejected for AMS of Melka-Worer station due to an upward trend as shown in Figure 4.2 with fitted loess curve. Values of Kendall's tau statistic, two-sided p-value (p-value), Kendall Score (S), denominator (D), variance of Kendall Score (Var(S)) are listed in Table 4.1.

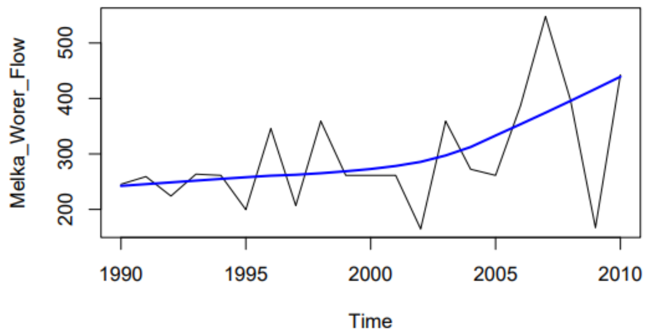


Figure 4.2. Fitting loss curve to the AMS of Melka-Worer station

Table 4.1. Mann-Kendal's trend test

Station	S	Var (S)	D	tau	Z	p-value	Ho at 5% sign. level
Wonji	28	1625.33	276	0.10	0.67	0.50	Accepted
Nura Hera	20	493.33	120	0.17	0.86	0.39	Accepted
Methara	-26	1096.67	210	-0.12	-0.75	0.57	Accepted
Melka Worer	73	1079.00	204.426	0.36	2.19	0.03	**Rejected**
Melka Sedi	-10	1625.33	276	-0.04	-0.22	0.82	Accepted
Abomusa	-1	492.33	119.499	-0.01	0	1.00	Accepted
Awash 7 kilo	36	950.00	190	0.19	1.14	0.18	Accepted

As shown in Table 4.1, for Melka Worer station, the two-sided p-value is about 0.03. This p-value is less than 0.05. Thus, we will reject the null hypothesis of the test at 5% significance level. Choosing a smaller value for Pcr like 1% leads towards accepting the null hypothesis. A similar study by (Daba et al., 2020) applied 90, 95, and 99% confidence levels to analyze the hypothesis. Thus, we accept it as no trend at 90% confidence level and pass to the next test.

Test for independence and stationarity using W-W test based on the Z test-statistic is used. The results of W-W test of independence and stationarity based on the Z-test statics are presented on Table 4.2. Because all the p-values are greater than 0.05, the null hypothesis is accepted and thus the series are independent and stationary.

Table 4.2. W-W test of independence and stationarity

Stations	n (years)	Z	P-value	H_0 at α=5%
Wonji	24	0.16	0.87	Accepted
Nura Hera	16	-1.16	0.25	Accepted
Methara	21	2.62	0.06	Accepted
Melka Sedi	24	1.61	0.11	Accepted
Melka Werer	21	0.19	0.85	Accepted
Abomusa	16	0.30	0.77	Accepted
Awash 7 kilo	20	0.10	0.92	Accepted

Homogeneity of AMS in each station is examined using Buishand's homogeneity test. For a homogenous record one may expect that the Sk's fluctuate around zero since there is no systematic pattern in the deviations of the X_i's from their average value $\bar{X}$. The null and alternate hypotheses were tested at 90%, 95% and 99% confidence interval representing 0.1, 0.05 and 0.01 degree of freedom. The plot of the result for each station is displayed in Appendix B. The results show that all stations are homogeneous at 95% confidence interval.

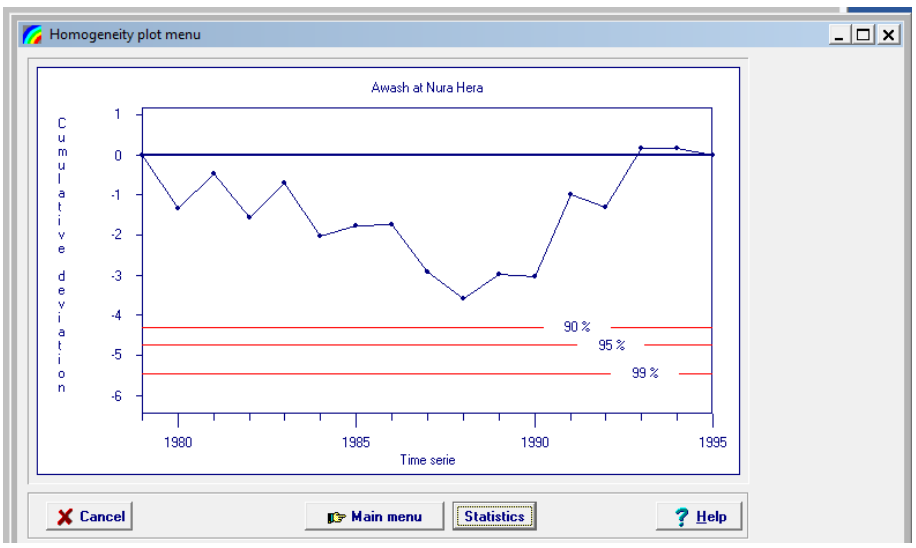


Figure 4.3. Homogeneity test for Nura Hera station

4.1.3. Outlier detection and treatment

A method used by Bulletin 17B is applied to detect low and high outliers. The outlier detection show that, the Methara station on 2004 have low outlier having a value of 32.74 m^3/s while the low threshold value estimated using outlier test is 50.33 m^3/s as shown in Table 4.3.

Table 4.3. Outlier detection

Station	$\bar{X}$	S	G	N	K_N	Q_H	Q_L	Outlier	Ho
Wonji	2.133	0.208	-0.05	24	2.467	441.00	41.77	-	Accepted
Nura Hera	2.319	0.180	-0.32	16	2.279	542.50	80.21	-	Accepted
Methara	2.100	0.200	-1.98	21	2.408	460.40	50.33	32.74	**Rejected**
Melka sedi	2.301	0.152	0.567	24	2.467	474.20	84.41	-	Accepted
Melka werer	2.446	0.134	0.295	21	2.408	588.30	132.75	-	Accepted
Abomusa	1.200	0.162	-0.893	16	2.279	37.15	6.76	-	Accepted
Awash 7 Kilo	2.496	0.206	-0.102	20	2.385	977.24	102.33	-	Accepted

The retention, modification, deletion of outliers can significantly affect the statistical parameters of the data, especially for small samples (Chow et al., 1988). Therefore, we have to consider both mathematical and hydrological considerations for their treatment. Flood peaks considered high outliers should be compared with historic flood data and flood information at nearby sites. If information is available which indicated a high outlier(s) is the maximum in an extended period of time, the outlier(s) is treated as historic flood data. If useful historic information is not available to adjust for high outliers, then they should be retained as part of the systematic record (USWRC, 1981). According to this guide, the techniques tested for handling outliers consisted of:

- keeping the value as it is;
- reducing the value to the product of the second largest event and the ratio of the second largest to eighth largest event;
- reducing the value to the product of the second largest event and the square root of that ratio; and
- discarding the value.

In the cases of low end outliers, the words largest in (b) and (c) should be changed to smallest.

In our case, comparing the historic flood data of Metehara station with nearby sites, the AMS data of the other stations shows a decreasing trend on 2004 from the previous year and and it is the minimum value Wonji station as shown in Figure 4.4. Thus, it is considered as a natural flow and retained in the series. According to Cunnane, (1989), even if retained, outliers have only a small effect if an efficient method of parameter estimation (ML or PWM) is used and this supports the retention of the above outlier as the study is through the application of RFFA based on PWMs.

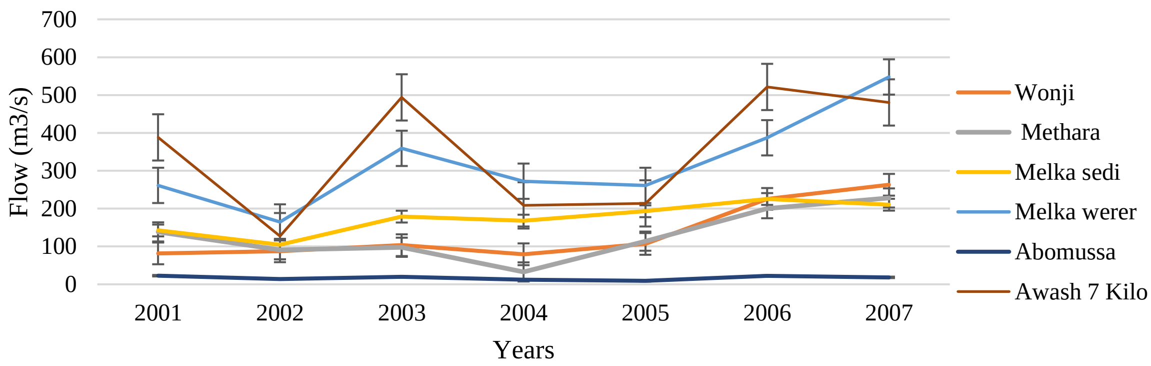


Figure 4.4. The low outlier of Metehara station 2004 in relation to other stations

4.2. Regional Flood Frequency Analysis

4.2.1. Tests of data homogeneity of the stations

The homogeneity of stations in the MARB is examined through discordancy measure test (Di) and heterogeneity measure statics (H). Values of Di are computed in terms of the L-moments and given in Table 4.4 along with recording length (n) for gauging sites of the study area.

Table 4.4. Major statistics and Discordancy of sites

Station name	n	mean (l_1)	t	t_3	t_4	t_5	Di
Wonji	24	150.921	0.265	0.140	-0.040	-0.074	1.170
Nura Hera	16	225.594	0.230	0.090	0.096	0.110	0.050
Methara	21	164.699	0.195	-0.118	0.076	-0.049	1.050
Melka Sedi	24	212.892	0.203	0.298	0.275	0.077	1.020
Melka Worer	21	292.788	0.178	0.227	0.209	0.018	1.260
Abomusa	16	16.800	0.186	-0.156	-0.021	0.034	1.190
Awash7kilo	20	347.769	0.258	0.123	-0.008	0.053	0.650

Sites with a Di greater than the critical value are considered discordant relative to the collective behavior for the proposed grouping of sites (Hosking & Wallis, 1997). It is observed that Di values of the seven gauging sites ranges from 0.05 to 1.26 which is less than the critical value 1.917. Thus, data of all the seven stream flow gauging sites are suitable for flood frequency analysis.

Heterogeneity measure, H, is used for identification of homogeneous region based on observed and simulated dispersion of L-moments for a group of sites under consideration. Regional weighted average L-moments (t_i^R) weighted proportionally to the record length at each site has gained a value of 0.222, 0.108 and 0.101 respectively for t_2^R, t_3^R and t_4^R. For heterogeneity test, a four-parameter kappa distribution is fitted to the regional data set generated from series of 500 equivalent region data by numerical simulation based on data of seven sites. The regional parameters of the kappa distribution are $\xi = 0.750$, $\alpha = 0.450$, k = 0.223 and h = 0.321. The results are shown in Table 4.5.

Table 4.5. The process of heterogeneity measure

Observed S.D. of group L-CV	0.0345
Sim. mean of S.D. of group L-CV	0.0298
Sim. S.D. of S.D. of group L-CV	0.0085
Standardized test value (**H_1**)	**0.56**
Observed Ave. of L-cv / L-skew distance	0.1241
Sim. mean of Ave. L-cv / L-skew distance	0.0821
Sim. S.D. of Ave. L-cv / L-skew distance	0.0202
Standardized test value (**H_2**)	**2.07**
Observed Ave. of L-skew/L-Kurt distance	0.1658
Sim. mean of Ave. L-skew/L-Kurt distance	0.1044
Sim. S.D. of Ave. L-skew/L-Kurt distance	0.0222
Standardized test value (**H_3**)	**2.77**

We used H_1 as a measure of regional homogeneity, as suggested by Rao and Hamed (1997). H_1 is more effective in determining homogenous regions than H_2 and H_3 (Modarres, 2006). Thus, we referred H_1 to accept or reject homogeneity. Their suggestion is most likely drawn from the fact that H_1 is calculated based on the coefficient of variation, and for realistic hydrologic regions, it has been illustrated that the coefficient of variation has a high value of regionalization (Modarres, 2008). H_1 value of 0.56 shows that the region is acceptably homogeneous. But in addition, value of H up to 3 is also acceptable (Eregno, 2014). Further, Hosking & Wallis, (1997) stated that for practical purpose, somewhat heterogeneous region can produce quantile estimates that are accurate enough. Therefore, we accepted that the study region is acceptably homogeneous.

4.2.2. Best fitted regional distribution

4.2.2.1. L-moment ratio diagram

LMRD is a widely accepted method for preliminary selection of distribution for a region although it is qualitative and subjective. Regional average L-moment statistical values (L-CS and L-CK) of the standardized flow for each station are used and plotted on LMRD along with 5 two parameter candidate distributions, (exponential (E), Gumbel (G), logistic (L), normal (N) and uniform (U)), as points and 5 three parameter distributions (GLO, GEV, GNO, PE3 and GPA) as a line curve.

But the use of GoF measure for two-parameter distributions are not recommended by Hosking & Wallis, (1997) as described in section 5.2.6. Plot of the diagram is shown in Figure 4.5.

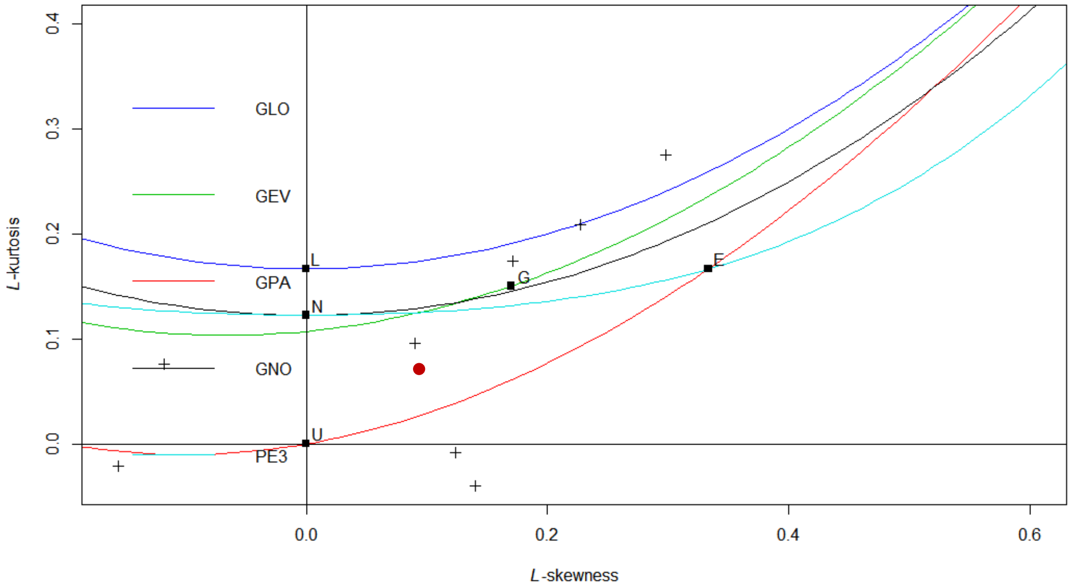


Figure 4.5. L-moment ratio diagram

In LMRDs, the data points are expected to be scattered close to the curve of a suitable distribution. It can be seen from the plot that there is wide scatter of the data points and it is almost impossible to exactly determine the suitable distribution based on scatter plot. But, observing data point defined by coordinates of the group average value of L-kurtosis and L-skewness (dark red color in Figure 4.5), is lied close to GEV, PE3 and GNO distributions and relatively far away from GLO and GPA distributions. Therefore, GEV, PE3 and GNO distributions can be accepted as suitable distributions for the study area, but the best distribution can be selected using Z-statistical test.

4.2.2.2. Goodness-of- fit values and estimated parameters

In addition to visual inspection and subjective judgment of distribution selection based on LMRD, Z-statistical test is used to evaluate GoF of GLO, GEV, GNO, PE3 and GPA distributions. The findings of the goodness-of-fit measures for 500 simulations are given in Table 4.6.

Table 4.6. Goodness-of-fit measures

Distribution type	τ_4^{DIST}	Z^{DIST}
Gen. logistics	0.1766	2.76
Gen. extreme value	0.1294	1.09
Gen. normal	0.1319	1.18
Pearson type III	0.1262	0.98
Gen. parento	0.033	-2.33

As we can see from Table 4.6, except GLO and GPA distributions, the other three distributions have a value that $|Z^{DIST}| \leq 1.64$ and thus GEV, GNO and PE3 are suitable distributions for the region. PE3 has the least Z^{DIST} value that ascertains it is the most suitable distribution for the region and therefore, PE3 is selected for estimation of flood quantiles.

Distribution parameters are estimated for the best fitted regional distribution, PE3. The computed regional weighted average L-moment values are 1, 0.222, 0.108 and 0.101 respectively for λ_1, t_2, t_3 and t_4. Using the L-moment values, parameters of location (μ), scale (σ) and shape (γ) is found to be 1, 0.401 and 0.666 respectively.

4.2.3. The RGC and estimated quantiles

As discussed in section 3.5.5, using the index flood, the annual maximum floods at each gauging station were reduced to dimensionless form ($Q_i/\bar{Q}$) before fitting the individual series to regional frequency distribution. For the homogeneous region, the model parameters derived from the best fitted distribution (PE3) are used to compute standardized quantile estimates, X_T, for the return periods T, and then used to construct the RGC (i.e., a curve showing X_T against T).

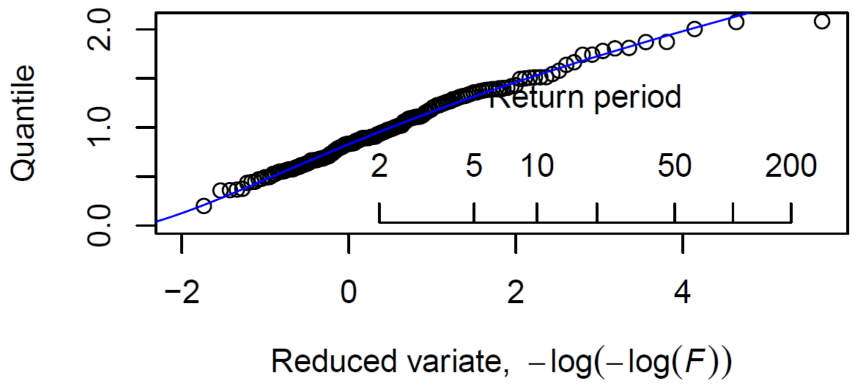


Figure 4.6. The developed RGC fitted to peak flow series

The curve fits reasonably well to the set of pooled flood peaks series in the region as shown in Figure 4.6. From the regional frequency curve, dimensionless regional quantile floods (X_T) for 2-, 5-, 10-, 20-, 50-, 100-, 200-, and 500-years of return period is listed as shown in the Table 4.7.

Table 4.7. Dimensionless regional quantile estimates

	μ = 1.000 α = 9.186			σ = 0.401 β = 0.134			γ = 0.666 ξ = -0.204		
Return period, T (year)	2	5	10	20	25	50	100	200	500
$P(x \geq X) = 1 - (1/T)$	0.5	0.8	0.9	0.95	0.96	0.98	0.99	0.995	0.998
Regional quantile (X_T)	0.96	1.32	1.53	1.73	1.73	1.96	2.12	2.28	2.48

After determining the RGC, the quantiles for different return periods are estimated by using the flood index method. The at-site quantile floods (Q_T) can be up scaled from the regional quantile flood estimates (X_T) by multiplying with the index flood of the station ($\bar{Q}$). Flow quantiles are estimated for 2-, 5-, 10-, 20-, 50-, 100-, 200-, and 500-years of return period for the seven gauged sites as shown in Table 4.8.

Table 4.8. Quantile estimates for individual sites

X_T	0.96	1.32	1.53	1.73	1.96	2.12	2.28	2.48
Stations	Q_2	Q_5	Q_{10}	Q_{20}	Q_{50}	Q_{100}	Q_{200}	Q_{500}
Wonji	144.25	198.93	231.49	260.53	295.58	320.37	344.07	374.10
Nura Hera	215.63	297.35	346.03	389.433	441.83	478.88	514.31	559.20
Methara	157.42	217.09	252.63	284.31	322.56	349.62	375.48	408.25
Melka Sedi	203.49	280.61	326.55	367.51	416.95	451.92	485.35	527.71
Melka Worer	279.85	385.92	449.10	505.43	573.43	621.52	667.50	725.76
Abomusa	16.06	22.14	25.77	29.00	32.90	35.66	38.30	41.64
Awash7 Kilo	332.40	458.39	533.43	600.34	681.10	738.23	792.84	862.04

4.2.4. Evaluating the performance of PE3 distribution

i. Quantile-quantile (Q-Q) plot

The quantile-quantile (Q-Q) plot of the observed and simulated values can determine if the two data sets can be represented with a common distribution. The Q-Q plots of the standardized observed floods versus the simulated values that generated from the best fitted regional frequency distributions (PE3) is plotted in Figure 4.7. Observing the above Q-Q plot, the points are assembled near the straight line. This implies that the frequency distributions that were chosen as a best distribution could be an appropriate regional flood model for the MARB.

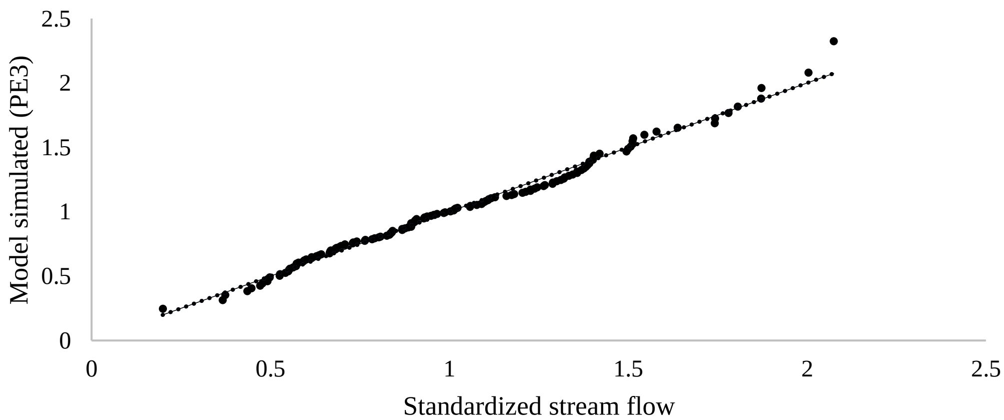


Figure 4.7. Q-Q plots of the standardized empirical discharge against randomly simulated

ii. Assessment of the accuracy of estimated quantiles

Accuracy of estimated quantiles from the RGC is validated using calculated RMSE values through Monte Carlo simulation. Quantile estimates (q(F)) are calculated for various non-exceedance probabilities (F) with 90% error bounds. For selected values of the non-exceedance probability, the results are shown in Table 4.9. Figure 4.8 shows the estimated growth curve.

Table 4.9. Accuracy measures for estimated RGC

			Error bounds	
F	q(F)	RMSE (%)	Lower	Upper
0.01	0.27	11.37	0.17	9.07
0.1	0.52	5.31	0.46	0.64
0.5	0.96	2.53	0.92	0.99
0.9	1.53	6.53	1.43	1.64
0.99	2.12	18.70	1.86	2.47
0.999	2.62	31.96	2.18	3.26

The estimated regional quantiles and their accuracy measures demonstrate that the estimated uncertainty as measured by the RMSE values and error bounds is relatively low and is acceptable as of Boucefiane & Meddi, (2019), i.e., < 20%, for return periods less than 100 years. But, for higher return periods, the uncertainty is higher and thus the width of the error bound becomes higher as shown in the Figure 4.8. This shows the need for longer time series records and more stations. Chow & Takase, (1997) suggested that we would need 1000 years of record to accurately predict a 100-year frequency event. At the lower tail of the error bound, it is found that the uncertainty of the upper bound is high. Hosking & Wallis, (1997) reasoned that that error bound by RMSE is unhelpful in the lower tail if the fitted distribution can take negative values. The results are comparable to those studies by Boucefiane & Meddi, (2019) and Malekinezhad & Zare-Garizi, (2014).

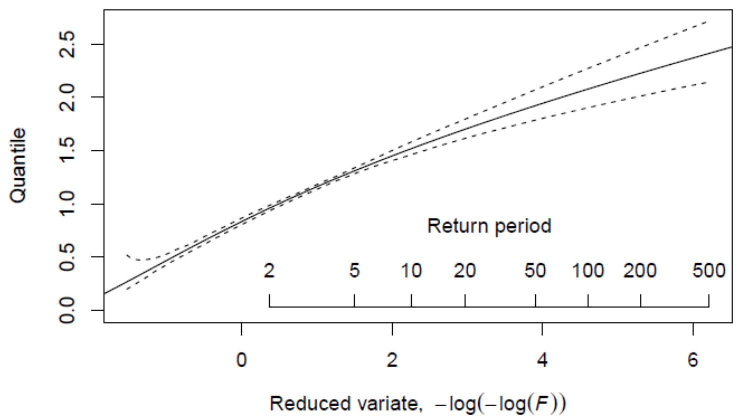


Figure 4.8. Regional quantile function fitted to PE3 distribution with 90 % error bounds

4.2.5. Estimated quantiles for ungauged catchments

A simple regression model by fitting catchment Area (A) and Index flood ($\bar{Q}$) to the polynomial equation is used to predict $\bar{Q}$ of the ungauged sites of E and G using equation 3.41. To compute quantile estimate (Q_T) of E and G for the different return periods, the values are multiplied by growth factor (X_T) as shown in Table 4.10.

Table 4.10. Quantile estimates for ungauged sites

	Return period, T, (years)				2	5	10	20	50	100	200	500
	Growth factor (X_T)				0.956	1.318	1.534	1.726	1.958	2.123	2.280	2.479
Sites	Coord. (x, y)		A (km^2)	$\bar{Q}$ (m^3/s)	Quantile estimates, Q_T (m^3/s)							
E	619721	980863	20674.08	239.05	228.53	315.07	366.70	412.60	468.06	507.50	545.04	592.61
G	549084	930351	13564.31	189.08	180.76	249.21	290.05	326.35	370.22	401.42	431.10	468.73

4.3. HEC-RAS calibration and validation results

4.3.1. Calibration results

The calibration was done so as to re-create the constructed rating curve of Awash 7 Kilo station through adjusting the Manning's roughness coefficient (n) of the river channel by varying its value between 0.033 and 0.045 as suggested in Chow, (1959). n value was varied as 0.033, 0.035, 0.040, 0.042 and 0.045 and their effect is measured through NSE, PBIAS and R^2.

Table 4.11. Observed and simulated WSE for different Q values for n = 0.042

Q (m^3/s)	**WSE Observed (m)**	**WSE Simulated (m)**	**Q (m^3/s)**	**WSE observed (m)**	**WSE Simulated (m)**
721.500	4.99	4.68	350.877	3.36	3.11
700.700	4.98	4.59	311.604	3.15	2.82
598.200	4.50	4.22	272.638	2.93	2.47
562.400	4.35	4.07	231.114	2.68	2.10
487.170	4.02	3.75	218.455	2.6	1.98
478.365	3.98	3.71	213.788	2.57	1.94
443.840	3.82	3.54	206.000	2.52	1.87
433.281	3.77	3.50	197.049	2.46	1.79
418.683	3.70	3.43	176.602	2.32	1.60
414.552	3.68	3.42	157.130	2.18	1.43
392.147	3.57	3.30	141.227	2.06	1.30
378.171	3.50	3.24	121.173	1.90	1.10
374.218	3.48	3.23	109.318	1.80	0.98

The calibration process showed that the water surface elevation increased when the Manning's roughness coefficients value increase showing consistent with the law of water flowing. Finally, 'n' value of 0.042 for the main river channel has been fixed as Manning's 'n' for all the cross sections in the basin which gives acceptable results.

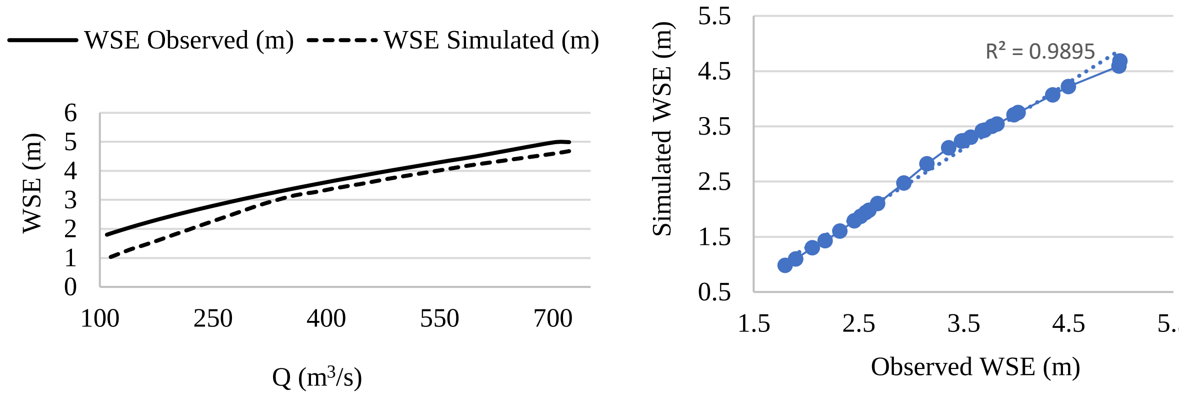


Figure 4.9. Observed vs simulated WSE of Awash 7 Kilo station

The performance examined through statistical measures of NSE, PBIAS, and R^2, using the data on Table 4.8, gives a value of 0.695 (very good), -13.80 (good) and 0.99 (very good) respectively. A close agreement between the observed and simulated values is also shown in Figure 4.9.

4.3.2. Validation results

The calibrated model was validated with a flood event of August 28, 1996 by comparing delineated water body, from Landsat 4-5 TM satellite image using NDWI value ≥ 0.2, and the inundation extent gained through the simulation of flow during the flood event.

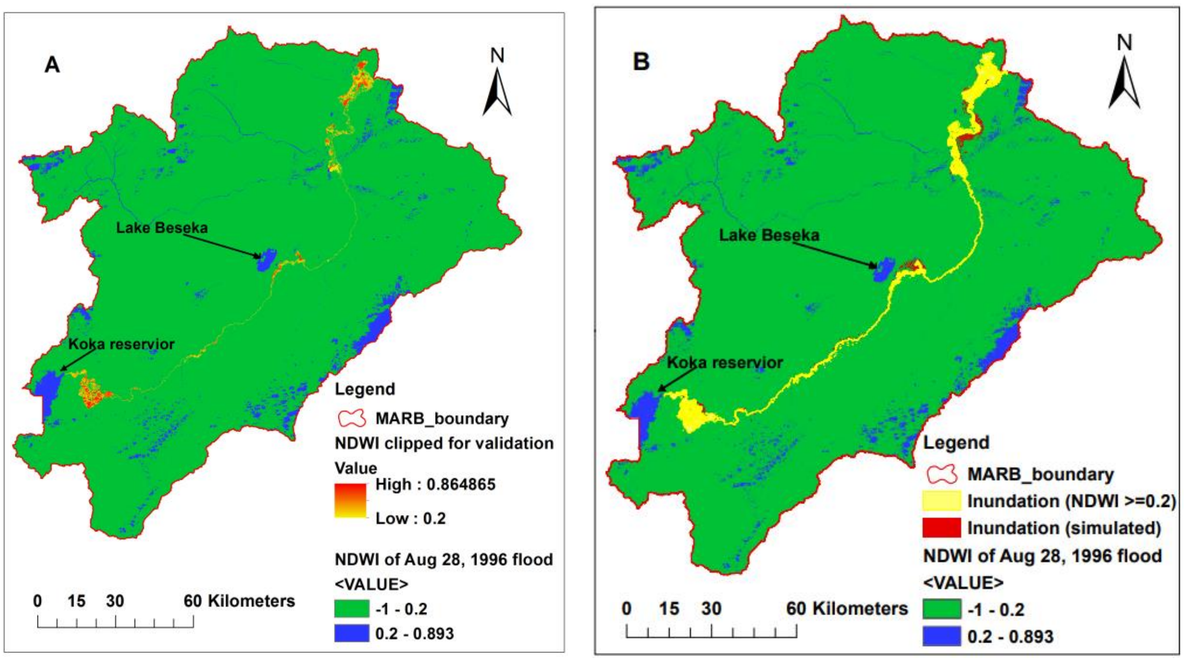


Figure 4.10. Inundation delineated based on NDWI (A) and inundation overlapping (B)

Performance assessment for inundation extent was done using the F statistic and overlapping percentage. It is found that a value of 94.74% and 86.67% respectively for overlapping percentage and F statistic. Sarhadi et al., (2012) considered F statistic value of 83.76% as relatively accurate estimation. According to Tamiru & Dinka, (2021) review, more than 85% of overlapping percentage is considered as a good agreement between the model and observed inundation extents, although Baky et al., (2019) accepted 76.5% overlapping accuracy.

Table 4.12. Performance assessment for inundation extent

	Area (ha)	Common area (ha)	Overlapping percentage (%)	F statistic (%)
Observed	18308.52	16553.74	94.74	86.67
Simulated	17345.36			

The discrepancies in the results of validation of the inundation extent come from multiple reasons. There exists inundation difference around Amibara and Dulecha *woredas* which have gentle slope. This may be attributed to a change in the river course through time. Because, there is about 20 years difference between the date of the satellite image and DEM downloaded. The two datasets have also a resolution difference in that the DEM is 12.5m while the Landsat 4-5 TM is 30m resolution. Adjustment of n value during calibration not results in exactly fitted to observed values. Thus, it will reduce the model performance during validation. NDWI value also has a great effect. Taking a threshold value of NDWI for separating flooded area and dry area is also a matter. Different literatures used variety of threshold values (Ji et al., 2009; Ji et al., 2019; Liu, 2012) and its finding is another task. The water surface produced had several islands and reservoirs due to the fact that HEC-RAS puts water on low-lying areas of the floodplain. Overall, the performance of HEC-RAS model in producing the August 28, 1996 flood inundation maps is quite promising and the simulation result can be regarded as a good agreement with inundation extent with the Landsat image.

4.4. Flood Hazard Assessment

For the preparation of flood hazard maps, three main steps namely pre-RAS, HEC-RAS and finally post-RAS processing were applied. Simulation was conducted on the calibrated and validated model to gain and analyze the inundation extent, depth and velocity under a flow of eight return periods, 2-, 5-, 10-, 20-, 50-, 100-, 200-, and 500-years. The area of inundation was computed as a product of cell size of the inundation and number of flood cells. One or more parameters can be considered in the hazard assessment depending on the characteristics of study area and floods.

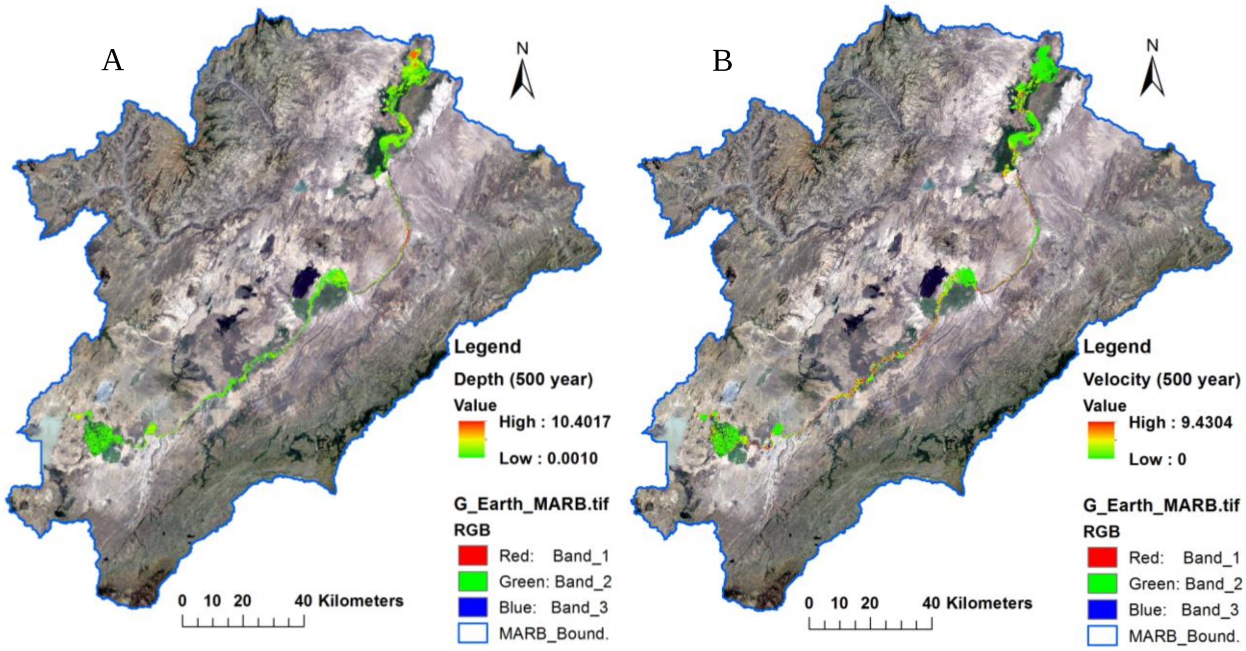
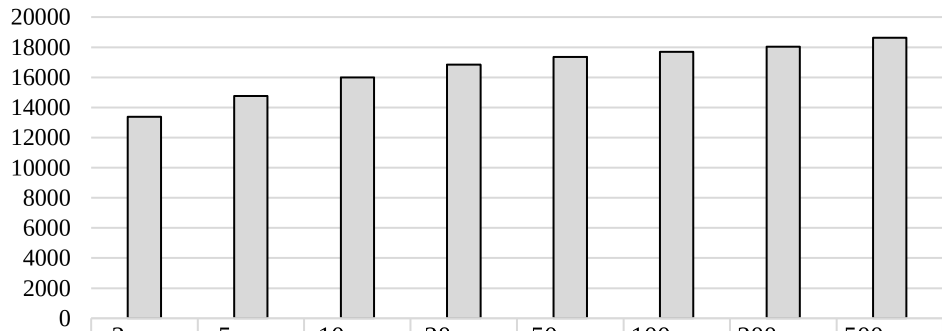


Figure 4.11. Inundation depth (A) and flow velocity map (B) for 500-year flood

The hazard mapping was done using inundation depth and velocity as hazard factors and the assessment was done on *woreda* level. Six *woredas* were identified as vulnerable to the flood hazard. These are Adama, Boset, Fentale, Awash, Amibara and Dulecha *woredas*. Inundation area for each return period for the whole study area, each *woreda* and for each hazard category is identified through overlying the *woreda* shapefiles to inundation map and the results are displayed in map and tabular forms. As indicated in the Figure 4.11, a maximum inundation depth and velocity is found to be 10.4m and 9.43m/s respectively for 500-year flood. Higher values were found towards the gorge of the river channel having low width floodplain. High inundation depth on downstream *woreda* of Dulecha is due to the inundation through a natural storage. The most u/s and d/s areas show lower velocity attributed to larger floodplain area. Maps of inundation depth and flow velocity are used for the generation of flood hazard maps combined according to the recommendation of Vojtek & Vojteková, (2016). Inundation area-return period relationship, inundation on *woredas* and the different hazard categories are discussed as follow.

i. Inundation area-return period relationship

The analysis of flood inundation areas, as shown in Figure 4.12, indicates that an increase in flood inundation with increasing discharge of flood from 2-years to 500-years return period. Similar result is found with a study by Baky et al., (2019), Chernet & Worku, (2021) and Banstola & Sapkota, (2019). It is found that the inundation due to 500-year flood is about 1.39 times the inundation by 2-year flood. As computed, a flood of 2-year return period produced 13383.23ha of flooded land, while the flooding areas of 14761.68, 15985.39, 16843.90, 17342.40, 17700.44, 18024.53, and 18636.45ha corresponds to 5-, 10-, 20-, 50-, 100-, 200- and 500-year return periods, respectively. The percentage increase diminishes as the return period increase. This shows the existence of floodplain area that may bounded by steep hills, highways and flood embankments (Tingsanchali & Karim, 2010). Highest percentage increase is found when proceeding from 2-year to 5-year with increase of 10.30% and a percentage increase of 8.29, 5.37, 2.96, 2.06, 1.83 and 3.39% respectively for 10- to 500-years of flood.



	2 year	5 year	10 year	20 year	50 year	100 year	200 year	500 year
Inundation Area (ha)	13383.23	14761.68	15985.39	16843.9	17342.4	17700.44	18024.53	18636.45

Figure 4.12. Inundation area-return period relationship

ii. Inundation area per *woreda*

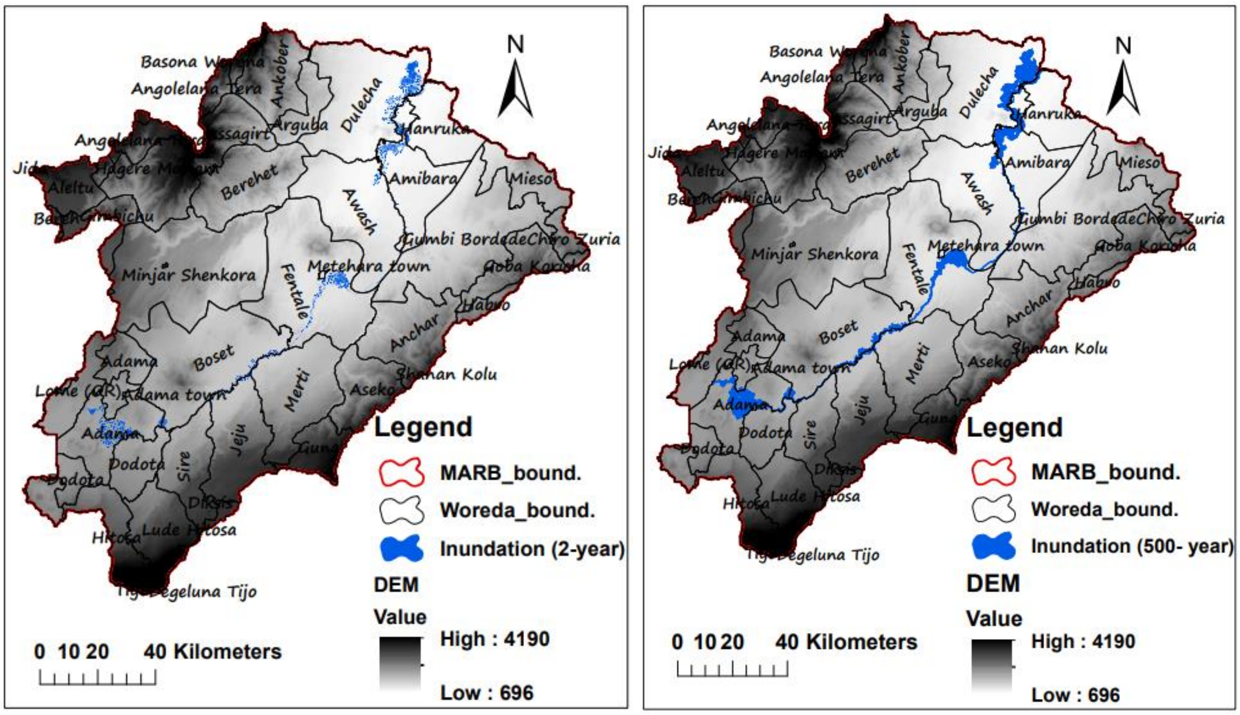
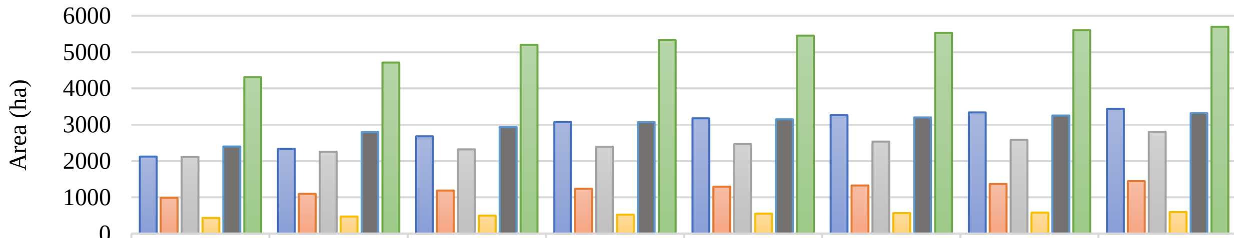


Figure 4.13. Inundation map of woredas for 2- and 500-year flood

As indicated in Figure 4.14, the *woredas* Dulecha, Adama, Fentale and Amibara show relatively higher coverage of inundation area. All woreda inundation area is increased with the increase in return period and areal percentage comparison shows, highest inundation coverage is found in

Dulecha *woreda*, due to its flat topography, covering 34.85, 35.07, 34.13, 33.89, 33.67, 33.52 and 32.93 percent of the inundated area for return periods of 2- to 500-years respectively. For all return periods, the inundation area coverage can be ranked as Dulecha, Adama, Amibara, Fentale, Boset and Awash in covering areal percentage of (32.93-34.85), (17.10-19.96), (19.14-20.45), (15.35-17.08), (7.91-8.36), (3.33-3.50) percent respectively.



	2 year	5 year	10 year	20 year	50 year	100 year	200 year	500 year
Adama	2123.85	2337.51	2681.76	3073.29	3176.11	3261.61	3339.37	3441.48
Boset	985.28	1096.51	1188.46	1236.55	1293.66	1327.49	1368.84	1446.33
Fentale	2112.36	2257.97	2322.13	2396.97	2469.74	2536.40	2581.88	2807.01
Awash	433.29	470.84	496.38	520.82	551.04	566.47	579.01	596.02
Amibara	2404.14	2797.00	2941.30	3070.56	3149.70	3204.48	3252.28	3312.19
Dulecha	4311.45	4712.68	5202.14	5336.28	5454.05	5531.83	5607.33	5698.45

Figure 4.14. Inundation areal coverage on individual woredas with their respective return period

iii. Hazard categories

Flood intensity is calculated for each 12.5m*12.5m cell using the inundation depth and velocity, ignoring the flow velocity under 1m/s. Then, the flood intensity is classified into three hazard categories as low, medium and high with FI value of < 0.5, [0.5-2] and > 2 respectively. The flood hazard is categorized based on the level of difficulties in daily life and/or damage to properties.

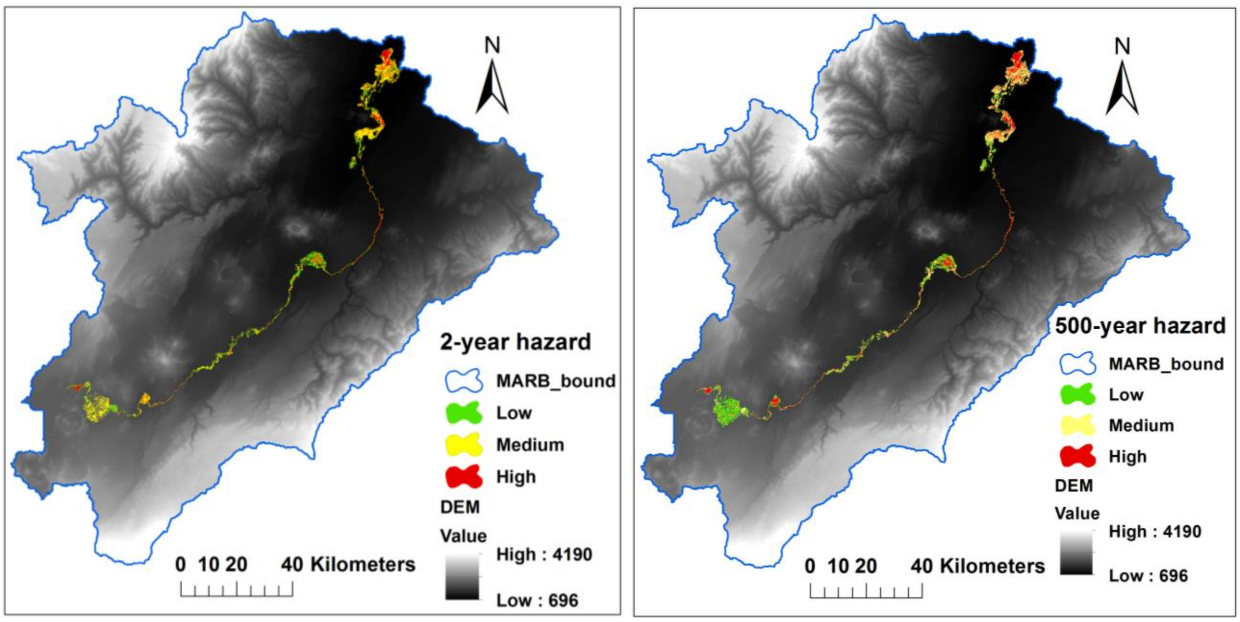


Figure 4.15. Flood hazard map for 2- and 500-year

Accordingly, the flood areas under different hazard categories were identified corresponding to each return periods. Figure 4.15 shows flood hazard categories under 2- and 500-year hazard. It is obvious from the figure that there are changes of areas in hazard category between the two return periods and most part of areas downstream are under medium and low hazard category. Also, we can observe that the area around Awash *woreda* with small floodplain area is mainly under high hazard class attributed to gorge river channel with high inundation depth and velocity. This can be also easily understood from Table 4.13 and Figure 4.16 and 17.

Table 4.13. Area coverage (ha) and percentage of hazard categories under the return periods

Return period	Hazard categories					
	Low		Medium		High	
	Area (ha)	%	Area(ha)	%	Area (ha)	%
2-year	1178.45	8.95	6505.38	49.40	5484.69	41.65
5-year	1197.91	8.23	7145.22	49.10	6210.31	42.67
10-year	2108.11	12.85	6896.00	42.03	7401.36	45.12
20-year	3019.55	17.33	6499.89	37.30	7905.95	45.37
50-year	3679.14	20.87	5940.95	33.70	8007.81	45.43
100-year	3490.30	19.57	6238.48	34.98	8104.61	45.45
200-year	2355.47	13.11	7431.56	41.38	8174.44	45.51
500-year	2657.97	14.24	7191.17	38.52	8818.89	47.24

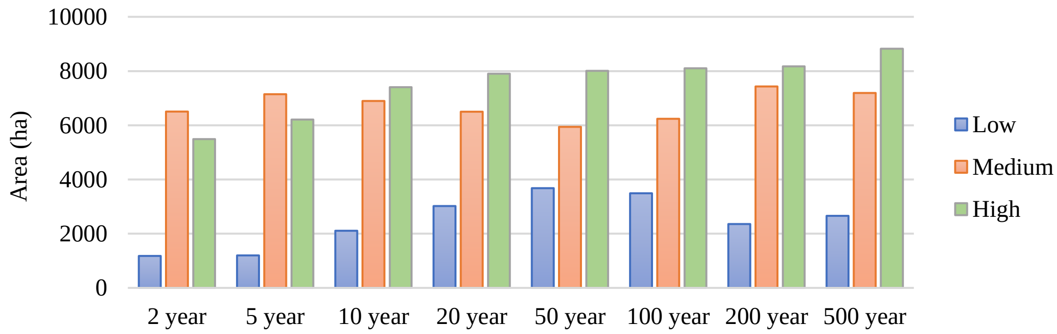


Figure 4.16. Areal coverage of the hazard categories under the different return periods

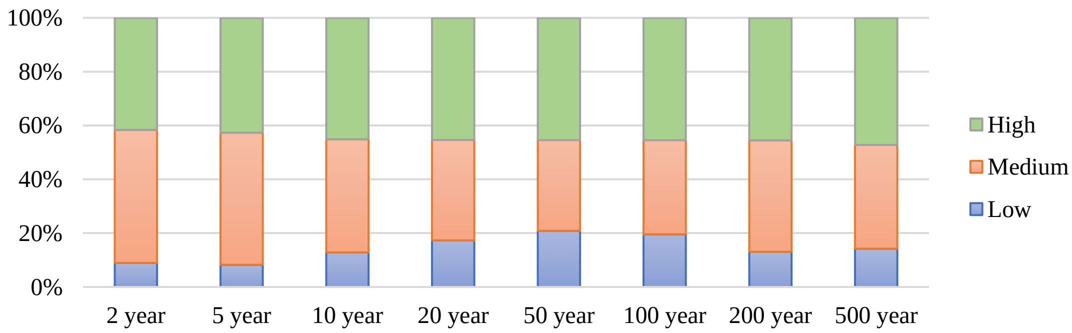


Figure 4.17. Graphical comparison of the areal percentage of individual hazard categories

It is clear from Figure 4.17 that the major part of study floodplain area is under medium and high hazard class with small percentage of low hazard coverage. For 2- and 5-year floods, medium hazard covers highest (49.40% and 49.10% respectively) and high hazard covers 45.12%, 45.37%, 45.43%, 45.45%, 45.51% and 47.24% respectively on 10-, 20-, 50-, 100-, 200- and 500-year floods. This is because the FI range in medium and high category is much wider than the low hazard category. The floodplain topography is also another factor that very low land level results higher depth as the return period increased which results in higher hazard. This can be seen in line with the results of Tingsanchali & Karim, (2005) that the flooding areas were 12%, 10%, 61% and 6%, respectively for the depth categories 0 to 0.6m, 0.6 to 1.0m, 1.0 to 3.5m and above 3.5m. It reasoned that small percentage for more than 3.5m is that the study area does not have a very low land level. Even with equal interval classification of the flood depth, the medium hazard class showed highest coverage of all other categories with only 7.50 to 9.30% coverage of very low hazard (Kim et al., 2020). This is similar with a study by (Banstola & Sapkota, 2019).

At some return periods, low and medium hazard areal coverage showed increasing and decreasing pattern as the return period increases. The decrease in coverage as the return period increases is probably due to a shift of an area from its present hazard to the higher hazard category. For low flood hazard class, with the increase in return period the inundated area increases up to 50-year return period then decreases for the next two return periods. This is probably because the inundation area shifted to medium hazard. Similarly, medium hazard area a decrease in coverage is observed on 20- and 50-year hazards due to a change to high hazard class. However, for all return periods, the high hazard class increases as the return period increases. A similar study by Baky et al., (2019) observed that the inundated area for the food class of Low decreased as compared to the food class of low at 10-year return period. Similarly, in Kim et al., (2020), the areal coverage by very low and low hazard classes showed a decreasing pattern for all return periods while it increases for very high hazard class. The percentage area of inundation increased from 8.95 to 20.87% corresponding to 2-year return period and 50-year return period, respectively, for food class Low. Whereas for medium- and high-class foods, the area of inundation increased from 33.70 to 49.40% and 41.65 to 47.24% respectively for return period of 50-year and 2-year for medium and 100-year for high hazard classes.

iv. Hazard categories coverage for each *woreda*

The flood hazard classes for each *woredas* under different return periods are presented in tabular and graphical form as shown in Table 4.14 and Figure 4.18.

Table 4.14. Hazard categories coverage per *woreda* for 2- to 500-year hazards

Woreda	2 year			5 year			10 year			20 year		
	low	medium	high	low	medium	high	low	medium	high	low	medium	high
Adama	167.31	1302.94	534.09	172.25	1433.94	566.83	681.92	1300.36	725.86	1382.34	1092.72	884.81
Boset	113.16	495.08	358.89	77.55	578.80	418.61	161.39	512.41	538.70	272.13	415.27	578.11
Fentale	624.45	742.92	808.20	73.24	1354.41	873.19	67.98	1368.27	905.17	191.47	1250.61	922.05
Awash	51.45	135.50	246.83	27.14	167.25	275.58	39.03	173.03	288.95	76.48	147.14	303.55
Amibara	120.56	1400.73	817.84	510.66	1176.39	1170.67	142.13	1606.88	1247.16	819.89	911.97	1421.77
Dulecha	42.11	2047.67	2143.92	282.67	2015.39	2289.91	899.16	1544.28	3028.64	1150.61	1309.72	3041.06

Woreda	50 year			100 year			200 year			500 year		
	low	medium	high	low	medium	high	low	medium	high	low	medium	high
Adama	1394.50	1106.36	893.41	1448.47	1101.76	909.67	1410.33	1166.55	919.80	1480.33	1094.53	1008.81
Boset	274.36	428.55	607.69	463.89	694.10	1044.81	204.81	509.59	659.50	219.84	475.52	766.84
Fentale	152.34	1302.20	939.47	154.81	1326.80	960.23	146.56	1336.73	974.31	564.11	974.22	1368.66
Awash	94.19	149.64	317.52	99.03	149.38	326.42	96.42	154.53	333.67	96.42	154.53	333.67
Amibara	549.52	1191.22	1435.53	340.05	1413.30	1442.44	257.89	1505.63	1448.61	105.00	1671.09	1455.97
Dulecha	989.31	1479.86	3047.41	938.11	1534.25	3050.72	79.36	2395.33	3053.53	53.44	2422.36	3059.55

Observing the table and figure, high and medium hazard area is mainly concentrated in Dulecha for all return periods with only a little lower for medium hazard at 10-year return period. An area of 2143.92, 2289.91, 3028.64, 3041.06, 3047.41, 3050.72, 3053.53 and 3059.55 ha of high hazard and 2047.67, 2015.39, 1544.28, 1309.72, 1479.86, 1534.25, 2395.33, 2422.36 ha of medium hazard covers Dulecha *woreda* respectively for 2-, 5-, 10-, 20-, 50-, 100-, 200- and 500-year return periods.

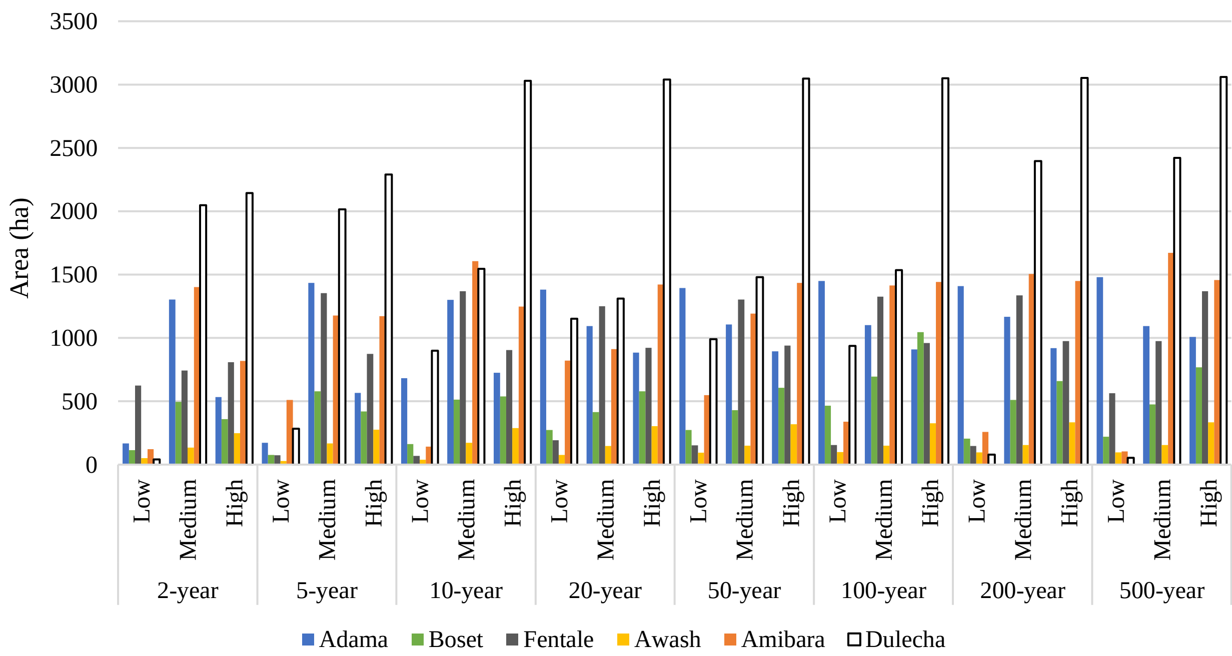


Figure 4.18. Areal coverage of hazard categories for each *woredas*

Next to Dulecha, the highest high hazard area is observed in Amibara woreda with values of 817.84, 1170.67, 1247.16, 1421.77, 1435.53, 1442.44, 1448.61, 1455.97ha respectively for 2- to 500-year hazard. This may be due to the topography that the two *woredas* have relatively flat area with gentle slope. The cumulative areal percentage of medium and high hazard for the two *woredas* cover 56.30, 47.45, 48.44, 43.33, 47.21, 47.39, 55.19, 60.27% and 60.32, 61.85, 63.49, 62.41, 61.91, 58.09, 60.93, 56.49% respectively for 2- to 500-year return periods. Ranking for areal coverage of low hazard class in *woredas* varies through the variation in return period, but, 20- to 500-year hazard of the low class, the highest coverage is found in Adama *woreda.*

4.5. Flood Vulnerability Assessment

i. Inundation of LULC types

Seven types of land use/cover types are found in the MARB. For the analysis of the inundation of land use/cover types, LULC map is clipped with inundation boundary of each return period and then vulnerable area is computed for 2- to 500-year flood as shown in Figure 4.19 and 20. From Figure 4.19, we can observe that, flood is extended towards the crop land areas to the right and left of floodplain area of the river which makes the flood vulnerability of the area higher.

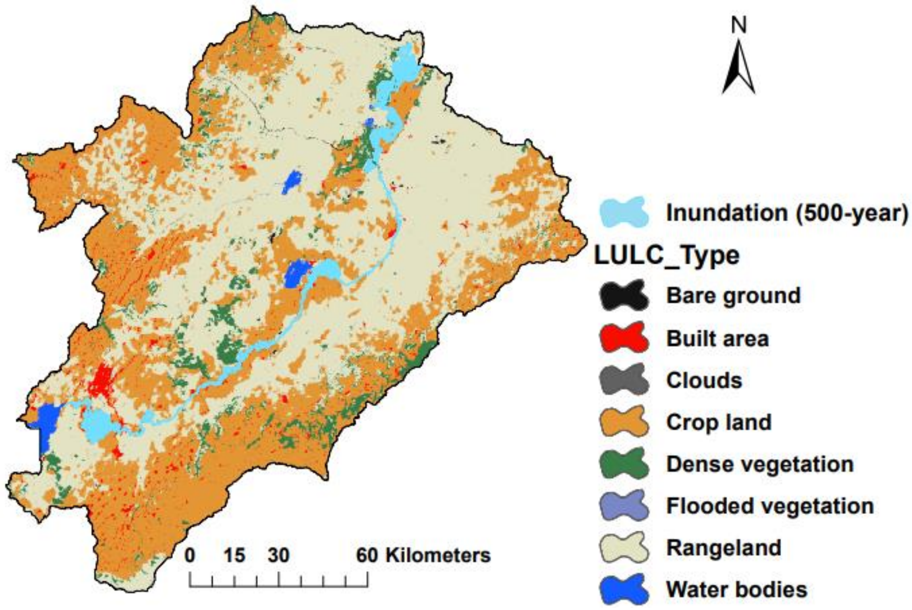
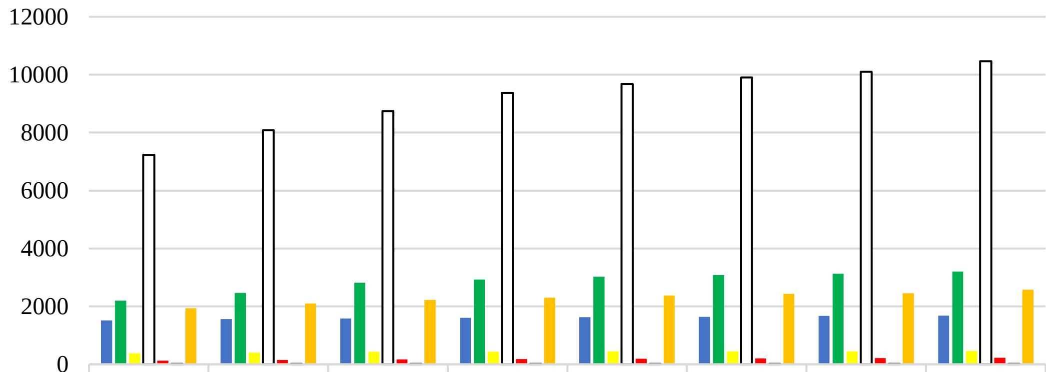


Figure 4.19. 500-year flood inundation of LULC types

Availability of dense vegetation in the downstream *woredas* of Amibara and Dulecha may leads to lower risk of flooding although it is under high flood inundation. Figure 4.20 presents the areal coverage of inundation for each land use class at different return periods of food. It is noticeable that, for all return periods, the land use/cover types are under inundation with increasing areal coverage as the return periods increases. Water bodies, dense vegetation, flooded vegetation, crop area, built area, bare ground and rangeland areas under 2-year flood are 10.2, 31.3, 17.0, 30.9, 45.6, 26.2 and 24.9% less than the inundation area by 500-year flood respectively. The crop land is the most vulnerable area with inundated area of more than half of the inundated floodplain covering areal percentage of 54.0, 54.8, 54.7, 55.7, 55.8, 55.9, 56.0 and 56.2% for 2- to 500-year return periods respectively. The other vulnerable areas of dense vegetation, rangeland, water bodies, flooded vegetation, built area and bare ground are under inundation with areal percentage of 16.4 to 17.2, 14.4 to 13.8, 11.3 to 9.0, 2.8 to 2.5, 0.9 to 1.2 and 0.1 to 0.09% respectively for a return period of 2- to 500-year inundating flood.



	2 year	5 year	10 year	20 year	50 year	100 year	200 year	500 year
Water bodies	1511.38	1556.01	1581.01	1606.87	1622.71	1633.8	1666.22	1683.47
Dense vegetation	2198.86	2459.92	2822.22	2933.23	3026.88	3086.16	3125.22	3198.8
Flooded vegetation	378.21	396.55	429.98	438.44	442.12	446.07	450.68	455.52
Crops	7236.33	8084.89	8747.28	9376.48	9678.08	9898.52	10097.7	10470.34
Built area	122.47	147.02	165.1	182.05	191.97	199.71	209.52	224.94
Bare ground	13.5	14.74	16.21	16.95	17.37	17.51	17.99	18.29
Rangeland	1929.81	2096.51	2223.49	2295.98	2374.8	2429.68	2451.9	2569.94

Figure 4.20. Inundation of LULC types under different return periods

ii. Flood vulnerability factor

Considering agricultural and built-up area as elements at risk, a vulnerability index (V_I) value is assigned as 1 and 3 respectively for crop land and built-up area based on their relative monetary value and V_I of 0 for the other types. Crop land and built-up area are usually used as elements at risk (Baky et al., 2019). But also Tingsanchali & Karim, (2005) used population density as elements at risk by assuming economic activities (agriculture) to be proportional to the population density while Chakraborty & Biswas, (2020) used landuse types and population density of the area. V_F is computed based on pixels with V_I value and area they cover for each land units considered (*kebeles* and towns) according to equation 3.51. The computed V_F for the land units showed a variation between 0.017 and 9.69. Out of the 80 land units considered, 25 land unites scored a vulnerability factor of less than 1 (48.98% of total area) and the remaining 26 and 29 land units scored 1 to 3 (27.77% of total area) and above 3 (23.26% of total area) respectively.

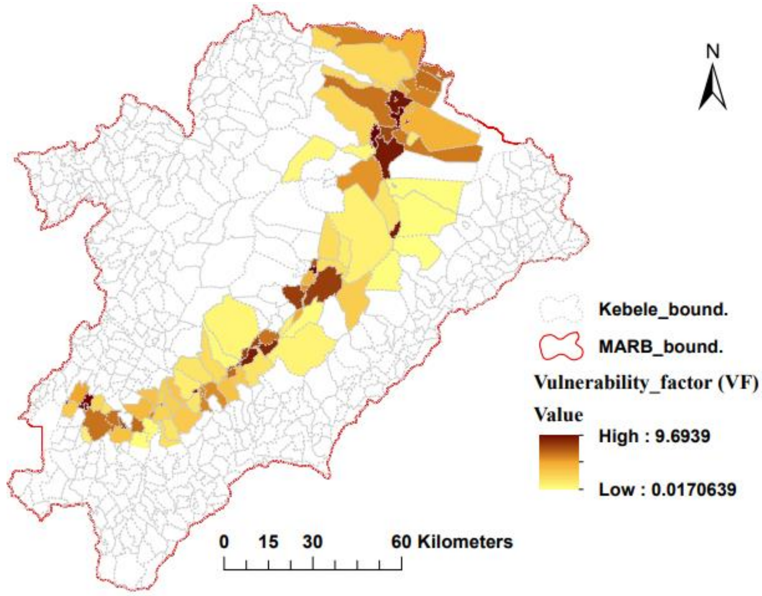


Figure 4.21. Vulnerability factor, V_F of the land units considered

4.6. Flood Risk Assessment

Risk was calculated for each land unit as a product of hazard and the vulnerability. The risk factors among the land units were found to vary from 0.05 to 26.92 for return periods of 2- to 500-years. For a systematic representation, the initial maximum risk factor of each return period was converted to 100. The other risk factors were proportionally converted to the same scale according to equation 3.52. The normalized risk factors varied from 0.24 to 100.

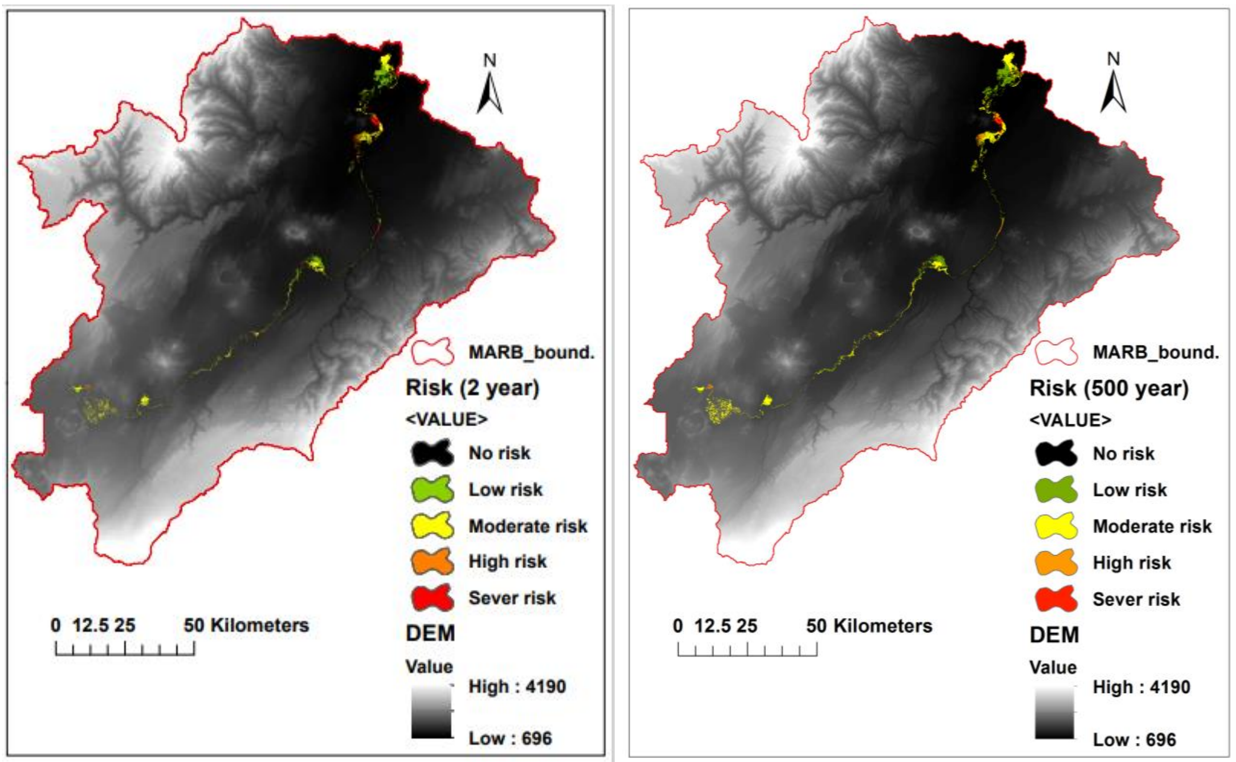


Figure 4.22. Land unit level flood risk maps for 2- and 500-year flood risk

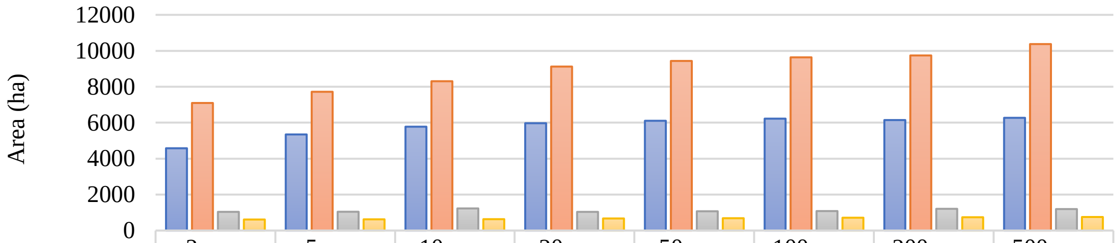
The other risk factors were proportionally converted to the same scale according to equation 3.52. The normalized risk factors varied from 0.24 to 100. Then, the land units are categorized into low, moderate-, high- and sever-risk of flood, by dividing R_F values using equal interval method (Tingsanchali & Karim, 2005). Finally, flood risks for each return period are mapped as shown in Figure 4.22.

From the figures, we can notice that most part of Dulecha is under low and moderate flood risk although it is under high flood hazard. This is because the area has low areal coverage of crop land and built-up area which leads to lower vulnerability to flood hazard. Presence of high coverage of agricultural area on Amibara woreda results in high and sever risk to be experienced on the land units in it. The high flood hazard areas of the Awash Fentale *woreda* is experienced low and moderate-risk as most of its part is covered with rangeland and contributes less for flood vulnerability. Towns along the floodplain of MARB such as Haro Adi (Fentale *woreda*), Melka Worer (Amibara *woreda*), Bole town (Boset *woreda*), Doni (Boset *woreda*), Awash Melkasa (Dodota), and Metehara (Fentale *woreda*) have experienced high and sever flood risk for the different return periods. Although Sire Robe town (Adama *woreda*), near Koka dam, showed sever-risk on the map, it is only along the river channel that does not inundate above its banks which is attributed to high inundation depth. Doni town has scored R_F of 100 for all return periods as it is a built-up and agricultural area. Generally, most gorge areas that was experiencing high flood hazard are under low- and moderate-risk.

Table 4.15. Number of land units under risk categories for each return period

Risk category	2-year	5-year	10-year	20-year	50-year	100-year	200-year	500-year
Low risk	30	30	31	29	25	23	22	22
Moderate risk	39	38	39	39	40	39	37	38
High risk	7	7	7	6	8	10	12	11
Sever risk	4	5	5	6	7	8	9	9

As shown in the Table 4.15, 80 flood vulnerable land units are identified in the study area. We can observe that the number of land units under moderate risk are the highest and the number of land units under low-risk class are in the second rank. The number of land units under each risk category showed an increasing and decreasing variations except the sever risk class, which do not show a decrease in number of land units as the return period increases. The decrease in the number of land units as the return period increased is mainly attributed to a change from on risk category to the other. Figure 4.23 shows areal coverage of the risk categories for each return periods.



	2 year	5 year	10 year	20 year	50 year	100 year	200 year	500 year
Low risk	4579.30	5347.29	5778.35	5972.36	6107.83	6226.19	6147.78	6272.15
Moderate risk	7097.49	7720.81	8308.00	9122.81	9436.28	9637.78	9742.18	10374.35
High risk	1040.45	1046.47	1231.32	1039.70	1073.71	1084.08	1209.45	1194.80
Sever risk	614.33	625.34	633.37	672.94	687.13	714.46	741.18	754.70

Figure 4.23. Flood-risk areas under floods of 2- to 500-year return periods

Figure 4.23 can be used for the comparison of area covered under the risk categories. Like the number of land units, here also the areal coverage is highest for moderate risk category followed by area under low risk category. It can be seen that, for low flood risk, flood risk area due to the 200-year flood in relation to the 100-year flood decreased by about 78ha of land. This may because the decrease in the number of land units that transformed to moderate risk due to the risk factors (R_F) were increase in to the moderate risk category under the 200-year flood. Moderate and sever risk categories showed an increasing trend for all return periods which may attributed to a risk category change of the land units from the lower return periods and increase in the areal extent of the hazard categories on the existing land units. The area under high risk category on 20- and 500-year risk decreased in relative to the preceding year flood risk, and observing the number of land units, they are decreased compared to the number of land units on 10- and 200-years of flood risk respectively.

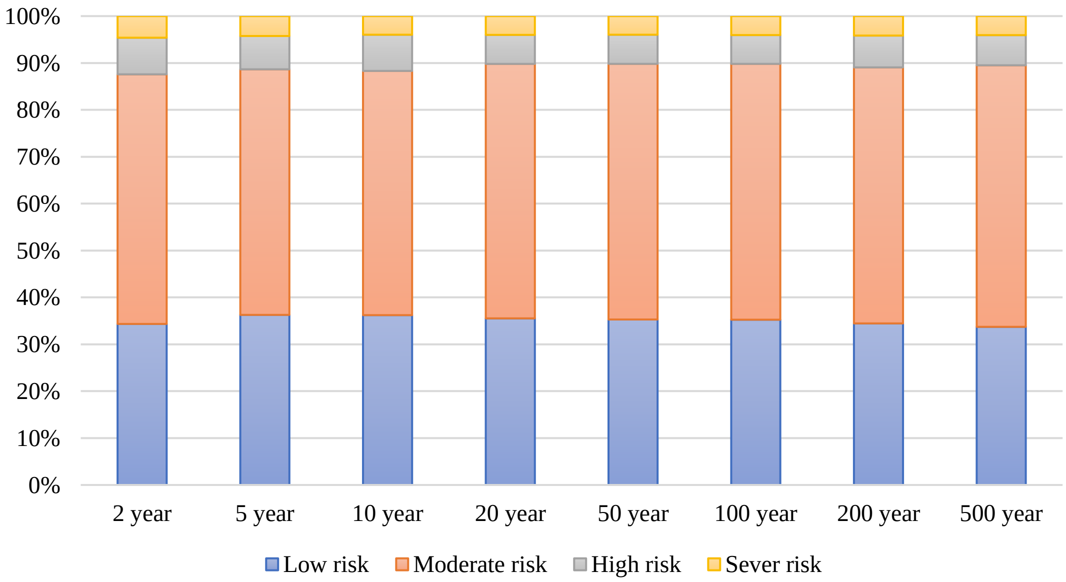


Figure 4.24. Graphical comparison of the areal percentage of individual risk categories

These trends also reported in other previous studies such as the study by Baky et al., (2019) where cropping land area covered with low and high food risk decrease with the increase in return period from 10 to 100-years and 50 to 100- years, respectively. Similarly, Tingsanchali & Karim, (2010) saw that moderate risk lands under the 50-year flood were relatively large compared to the 100- or 200-year floods and reasoned that in some high and severe risk land units under the 100-year flood, the risk factors were reduced to the moderate risk category under the 50-year flood.

The graphical comparison of risk categories by their areal coverage under each return period can be easily observed from Figure 4.24. Generally, high and sever risks are experienced on small areas covering a grand total areal percentage of 12.41, 11.34, 11.69, 10.19, 10.18, 10.18, 10.93 and 10.48 % respectively for 2-, 5-, 10-, 20-, 50-, 100-, 200- and 500-year risk and the remaining percentage is covered with low and moderate risk of flood.

A critical point in food risk assessment is that while flood hazard is the same for a given area in terms of intensity, the risk could be different depending flood vulnerability. Overall, this study addressed this fact and showed that areas of inundation increase with increase in return period of flood. These results were also reported in other previous studies such as Baky et al., (2019), Khattak et al., (2016) and Tingsanchali & Karim, (2010).

5. CONCLUSION AND RECOMMENDATION

5.1. Conclusion

This study presents a systematic approach in the preparation of hazard, vulnerability and risk maps and conducted their assessment on Middle Awash River basin with the application of 1D steady flow model (HEC-RAS), ArcGIS and HEC-GeoRAS. Datasets such as daily instantaneous streamflow data, DEM, water level data, land use/cover map, satellite image, administrative shapefiles, bridge data are used as input for analysis tools.

Exploratory data analysis performed on the instantaneous daily streamflow data of 11 gauging stations with 10 to 25 years record length. AMS of flow from seven stations are selected after initial screening and found to fulfill iid conditions and thus used for flood frequency analysis. The AMS of the stations is also considered as consistent and obtained outlier is incorporated to the dataset. For the reliable estimation of flood quantiles for gauged and ungauged stations, a RFFA based on L-moments and index flood method is applied. Based on Di and H statistics measure tests, MARB is found to be a homogeneous region. From the candidate distributions, GEV, GNO and PE3 are found suitable distributions and PE3 is selected as the best fit distribution. Regional growth curve (RGC) was developed and the quantile estimates are found to be 0.956, 1.318, 1.534, 1.726, 1.958, 2.123, 2.28, 2.479 for 2-, 5-, 10-, 20-, 50-, 100-, 200-, and 500-year return period floods respectively. Quantile estimates are computed for ungauged sites based on the RGC and Index flood ($\bar{Q}$) of ungauged sites.

Flood inundation maps are generated using flood parameters of extent, depth and velocity for the flow of 2- to 500-years. Calibration of the HEC-RAS model was made by adjusting Manning's n values, resulted a performance of very good for NSE, good for PBIAS and very good for R^2. Validation of was performed through comparing flood extent of historical flood from satellite image and HEC-RAS simulations and resulted overlapping percentage and the F statistic and gained a value of 94.74 and 86.67 respectively. This shows good performance of HEC-RAS for inundation mapping and the ability of satellite images for flood extent mapping.

Flood risk is estimated for each land units, considering boundaries of *Kebeles* and towns as land units. Six *woredas* (Adama, Boset, Fentale, Awash, Amibara and Dulecha) and eighty *Kebeles* and towns were considered in the assessment. Flood extent, depth and velocity were considered in flood hazard analysis. Flood hazard is classified as Low, Medium and High hazards. The analysis indicated a considerable increase in inundated area with an increase in the amount of peak discharge found in successive return period. Generally, the majority floodplain area of MARB is under Medium- and High-hazard class with small percentage of Low hazard area.

For the flood vulnerability analysis, crop land and built-up areas were selected as elements at risk Out of the 80 land units considered, 25 land units scored a vulnerability factor of less than 1 (48.98%) and the remaining 26 and 29 land units scored 1 to 3 (27.77%) and above 3 (23.26%) respectively. Observing the inundation extent, the crop land is the most vulnerable area with inundated area of more than half of the inundated.

The land units are categorized into low risk, moderate risk, high risk and very high risk. Towns along the floodplain of MARB such as Haro Adi (Fentale *woreda*), Melka Worer (Amibara *woreda*), Bole town (Boset *woreda*), Doni (Boset *woreda*), Awash Melkasa (Dodota *woreda*), and Metehara (Fentale *woreda*) towns have experienced high and sever flood risk. The areal coverage is highest for moderate-risk category followed by area under low-risk category. Generally, high- and sever-risks are experienced on small areas of the basin. The study addressed that flood risk of the land units can be varied even after there is no change in flood hazard intensity.

This study has shown that flood modelling is possible with the use of available data in MARB using 1D HEC-RAS software at steady state as it requires minimal data input to perform hydraulic analysis. The resulting flood hazard and risk maps hold much potential for supporting disaster mitigation decisions by providing information to identify priority areas in the risk zones of MARB which is particularly useful when there is a budget and time constraint.

5.2. Recommendations

There are recommendations that should be addressed for future studies and actions to be taken that are not included in this study due to resource and time constraints. Here are described under.

Remote sensing products: TIN generated from 12.5m DEM is used for inundation mapping. For detailed analysis, higher resolution DEM is needed. Manning's n values are assigned for each LULC type according to Chow, (1959). Conducting a field survey for the estimation of more accurate values can reduce deviations from true values. Although a wide selection of satellite images was available for the study site, only the one is correlated with the date on which a historic flood event occurred with minimal cloud cover. But this image is about 20 years difference compared to the year of DEM created. Thus, it is a good practice to develop real-time satellite imagery during a flooding event. NDWI threshold values varies in delineating water bodies and thus, it is good to study further and fix its value.

The flow data: Many gauge stations in MARB do not capture the daily flow discharge values well due to missed values (mainly stations on tributaries), flow regulation, and low record length. This will reduce the benefit from a RFFA. In the estimation of quantiles for ungauged sites, the use of catchment characteristics such as slope, altitude, infiltration capacity, stream length, drainage density, land cover etc., in addition to catchment area, will enhance the quality of the result.

Model related: 1D steady flow assumption of one-dimensional HEC-RAS model flow is not always valid in natural river channels as natural river course are turbulent, and hence it is a good practice to consider unsteady flow simulation models of 2D or even 3D. Consideration of detailed structures along the river channel and its floodplain such as levees, culverts, inline structures, obstructions, storage areas, etc., will lead to a better estimate of flood hazard and risk.

The assessment: In addition to inundation depth, extent and velocity as hazard factors, there are many factors that should be incorporated such as duration of flooding, flow velocity, timing of occurrence, rate of rise of food, etc. Another limitation is that only physical vulnerability of assets is considered while vulnerability due to food hazards encompasses physical, social, economic and environmental dimensions. In addition to the risk assessment, estimation of damages due to flood hazard with a detailed survey of field data will make future studies better. In addition, for integrated risk management the following recommendations are forwarded:

- Institutional frameworks should incorporate flood risks on their policies and plans;
- Compiling flood hazard and risk maps and identify forbidden zone, restricted zone and area for development is necessary for sustainable flood risk management;
- Watershed management plan comprising afforestation, reforestation, soil and water conservation practices should work to regulated discharge of water at downstream;
- The river training activities such as the construction of the dams, retaining walls, dykes, etc., are recommended to avoid the over bursting of the flood from the river banks;
- Early warning and communication systems are needed to facilitate information sharing;
- Land use planning can substantially reduce the vulnerability of the people and their activities.

REFERENCES

Abaya, S. W., Mandere, N., & Ewald, G. (2009). Floods and health in Gambella region, Ethiopia: A qualitative assessment of the strengths and weaknesses of coping mechanisms. *Global Health Action*, *2*(1), 2019. https://doi.org/10.3402/gha.v2i0.2019

Abdulrazzak, M., Elfeki, A., Kamis, A., Kassab, M., Alamri, N., Chaabani, A., & Noor, K. (2019). Flash flood risk assessment in urban arid environment: Case study of Taibah and Islamic universities' campuses, Medina, Kingdom of Saudi Arabia. *Geomatics, Natural Hazards and Risk*, *10*(1), 780–796.

Abebe, M. (2010). Disaster management in Ethiopia: A review of its checkered history, its transformation and some implications for a vibrant disaster management system, 1975–2008. *Journal of Sustainable Development in Africa*, *12*(4), 237–254.

Achamyeleh, K. (2003). *The Associated Programme on Flood Management; Integrated Flood Management, Case Study Ethiopia: WMO and GWP*. http://www.apfm.info/pdf/case_studies/Ethiopia.pdf

Adoleye, A. J., & Montaseri, M. (2002). Preliminary streamflow data analyses prior to water resources planning study / Analyses préliminaires des données de débit en vue d'une étude de planification des ressources en eau. *Hydrological Sciences Journal*, *47*(5), 679–692. https://doi.org/10.1080/02626660209492973

Ahrens, J. D. (1996, August 28). *Situation Report on Region 2 (Afar National Regional State)*. UNITED NATIONS DEVELOPMENT PROGRAMME Emergencies Unit for Ethiopia. https://www.africa.upenn.edu/eue_web/Awash29aug.htm

Ali, K., Bajracharya, R. M., & Koirala, H. L. (2016). A review of flood risk assessment. *International Journal of Environment, Agriculture and Biotechnology*, *1*(4), 1065–1077. http://dx.doi.org/10.22161/ijeab/1.4.62

Alzahrani, A. S. (2017). *Application of two-dimensional hydraulic modeling in riverine systems using hec-ras*. University of Dayton.

Antony, R., Rahiman, K. A., & Vishnudas, S. (2021). Flood Hazard Assessment and Flood Inundation Mapping—A Review. *Current Trends in Civil Engineering*, 209–218.

Armenakis, C., Xinheng Du, E., Natesan, S., Persad, R. A., & Zhang, Y. (2017). Flood Risk Assessment in Urban Areas Based on Spatial Analytics and Social Factors. *Geosciences*, *7*(123). https://doi.org/10.3390/geosciences7040123

Baky, M. A. A., Islam, M., & Paul, S. (2019). Flood Hazard, Vulnerability and Risk Assessment for Different Land Use Classes Using a Flow Model. *Earth Systems and Environment*, *4*(1), 225–244. https://doi.org/10.1007/s41748-019-00141-w

Banstola, P., & Sapkota, B. (2019). Flood Risk Mapping and Analysis Using Hydrodynamic Model HEC-RAS: A Case Study of Daraudi River, Chhepatar, Gorkha, Nepal. *Grassroots Journal of Natural Resources*, *2*(3), 25–44. https://doi.org/10.33002/nr2581.6853.02033

Beffa, C. (1998). *Two-dimensional modelling of flood hazards in urban areas*. 3rd International Conference on Hydroscience and Engineering, Cottbus/Berlin: Brandenbur University of Technology.

Benameur, S., Benkhaled, A., Meraghni, D., Chebana, F., & Necir, A. (2017). Complete flood frequency analysis in Abiod watershed, Biskra (Algeria). *Natural Hazards*, *86*(2), 519–534. https://doi.org/10.1007/s11069-016-2703-4

Benito, G., & Hudson, P. F. (2010). Flood hazards: The context of fluvial geomorphology. In *Geomorphological Hazards and Disaster Prevention* (pp. 111–128). Cambridge University Press. https://doi.org/10.1017/CBO9780511807527.010

Bhatt, G. D., Sinha, K., Deka, P. K., & Kumar, A. (2014). Flood hazard and risk assessment in Chamoli District, Uttarakhand using satellite remote sensing and GIS techniques. *International Journal of Innovative Research in Science, Engineering and Technology*, *3*(8), 9.

Bocchiola, D., De Michele, C., & Rosso, R. (2003). Review of recent advances in index flood estimation. *Hydrology and Earth System Sciences*, *7*(3), 283–296. https://doi.org/10.5194/hess-7-283-2003

Boucefiane, A., & Meddi, M. (2019). Regional growth curves and extreme precipitation events estimation in the steppe area of northwestern Algeria. *Atmósfera*, *32*(4), 287–303. https://www.redalyc.org/journal/565/56566341003/html/#gt6

Brook, L., & Boneya, G. (2020). Flooding in Ethiopia;Causes, Impact and Coping Mechanisms. A Review. *International Journal of Research and Analytical Reviews (IJRAR)*, *7*(3), 707–717.

Brunner, G., Savant, G., & Health, R. E. (2020). *Modeler application guidance for steady vs unsteady, and 1D vs 2D vs 3D hydraulic modelling*. U.S. Army Corps of Engineers Institute of Water Resources Engineering Center (CEIWR-HEC). https://www.erdc.usace.army.mil

Brunner, G. W. (2020). *HEC-RAS, River Analysis System Hydraulic Reference Manual Version 6.0 Beta*. US Army Corps of Engineers,Hydrologic Engineering Center. www.hec.usace.army.mil

Buishand, T. A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology*, *58*(1), 11–27. https://doi.org/10.1016/0022-1694(82)90066-X

Bulti, D. T., Mekonnen, B., & Bekele, M. (2017). Assessment of Adama city flood risk using multicriteria approach. *Ethiop J Sci Sustain Dev*, *4*(1), 6–23.

Chakraborty, S., & Biswas, S. (2020). Application of Geographic Information System and HEC-RAS in Flood Risk Mapping of a Catchment. In *Applications of Geomatics in Civil Engineering* (pp. 215–224). Springer.

Chebana, F., Ouarda, T., & Duong, T. C. (2010). Testing for stationarity, homogeneity and trend in multivariate hydrologic time series: A review. *Rep. I-271*.

Chernet, A., & Worku, S. (2021). Flood inundation mapping of Jigjiga-Town and its surrounding environment: Using GIS and HEC-RAS model. *International Journal of Environmental Protection and Policy*, *9*(2), 40. https://doi.org/10.11648/j.ijepp.20210902.14

Cheung, J. (2014). *Determining the optimal river gauge location for a flood early warning system in Uganda using HEC-RAS and AHP* [PhD Thesis]. Massachusetts Institute of Technology.

Chow, V. T. (1959). *Open Channel Hydraulics*. McGraw-Hill college. libgen.li/file.php?

Chow, V. T., & Takase, N. (1997). Design criteria for hydrologic extremes. *Journal of the Hydraulics Division*, *103*(4), 425–436. https://doi.org/10.1061/JYCEAJ.0004735

Chowdhury, J. U., & Karim, M. F. (1996). A risk-based zoning of storm surge prone area of the Ganges Tidal Plain. *J. Civil Eng*, 221–233.

Cunnane, C. (1989). Statistical distributions for flood frequency analysis. *Operational Hydrology Report (WMO)*.

Daba, M. H., Ayele, G. T., & You, S. (2020). Long-Term Homogeneity and Trends of Hydroclimatic Variables in Upper Awash River Basin, Ethiopia. *Advances in Meteorology*, *2020*.

Dahmen, E. R., & Hall, M. J. (1990). *Screening of hydrological data: Tests for stationarity and relative consistency*. ILRI.

Dalrymple, T. (1960). *Flood Frequency Analyses, Manual of Hydrology: Part 3*. https://doi.org/10.3133/wsp1543A

Davis, R. S. (2001). Flash Flood Forecast and Detection Methods. In C. A. Doswell (Ed.), *Severe Convective Storms* (pp. 481–525). American Meteorological Society. https://doi.org/10.1007/978-1-935704-06-5_12

De Risi, R., Jalayer, F., & De Paola, F. (2015). Meso-scale hazard zoning of potentially flood prone areas. *Journal of Hydrology*, *527*, 316–325. https://doi.org/10.1016/j.jhydrol.2015.04.070

Desalegn, H., & Mulu, A. (2021). Mapping flood inundation areas using GIS and HEC-RAS model at Fetam River, Upper Abbay Basin, Ethiopia. *Scientific African*, *12*, e00834.

Douglas, E. M., Vogel, R. M., & Kroll, C. N. (2000). Trends in floods and low flows in the United States: Impact of spatial correlation. *Journal of Hydrology*, *240*(1), 90–105. https://doi.org/10.1016/S0022-1694(00)00336-X

DPPC. (2006). *Joint Government and Humanitarian Partners: Flash appeal for the 2006 flood disaster in Ethiopia-Director general speech*.

Duguma, F. A., Feyessa, F. F., Demissie, T. A., & Januszkiewicz, K. (2021). Hydroclimate Trend Analysis of Upper Awash Basin, Ethiopia. *Water*, *13*(12), Article 12. https://doi.org/10.3390/w13121680

Dyhouse, G., Hatchett, J., & Benn, J. (2003). *Floodplain modeling using HEC-RAS*. Haestad Press. https://agris.fao.org/agris-search/search.do?recordID=US201300088324

Ekeu-wei, I. T., & Blackburn, G. A. (2018). Applications of Open-Access Remotely Sensed Data for Flood Modelling and Mapping in Developing Regions. *Hydrology*, *5*(39), 36. https://doi.org/10.3390/hydrology5030039

Elkhrachy, I., Pham, Q. B., Costache, R., Mohajane, M., Rahman, K. U., Shahabi, H., Linh, N. T. T., & Anh, D. T. (2021). Sentinel-1 remote sensing data and Hydrologic Engineering Centres River Analysis System two-dimensional integration for flash flood detection and modelling in New Cairo City, Egypt. *Journal of Flood Risk Management*, *14*(2), e12692.

El-Naqa, A., & Jaber, M. (2018). Floodplain Analysis using ArcGIS, HEC-GeoRAS and HEC-RAS in Attarat Um Al-Ghudran Oil Shale Concession Area, Jordan. *Journal of Civil & Environmental Engineering*, *08*. https://doi.org/10.4172/2165-784X.1000323

England, Jr., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas Jr, W. O., Veilleux, A. G., Kiang, J. E., & Mason Jr, R. R. (2019). *Guidelines for determining flood flow frequency—Bulletin 17C*. US Geological Survey.

EOS Data Analytics inc. (2021). *Normalized Difference Water Index: NDWI Formula and Clculations*. https://eos.com/make-an-analysis/ndwi/

Eregno, F. E. (2014). Regional flood frequency analysis using L-moment in the tributaries of Upper Blue Nile River, South Western Ethiopia. *Merit Journal of Engineering, Pure and Applied Sciences*, *2*(2), 12–21.

Erena, S. H., Worku, H., & De Paola, F. (2018). Flood hazard mapping using FLO-2D and local management strategies of Dire Dawa city, Ethiopia. *Journal of Hydrology: Regional Studies*, *19*, 224–239.

Gadise, T. (2020). *Regional Flood Frequency Analysis and Flood Inundation Mapping: The Case of Middle Awash River Basin, Ethiopia* [Thesis, ASTU]. http://213.55.101.20:8080/xmlui/handle/123456789/1654

Gan, T. Y., Zunic, F., Kuo, C.-C., & Strobl, T. (2012). Flood mapping of Danube River at Romania using single and multi-date ERS2-SAR images. *International Journal of Applied Earth Observation and Geoinformation*, *18*, 69–81. https://doi.org/10.1016/j.jag.2012.01.012

Garmdareh, E. S., Vafakhah, M., & Eslamian, S. S. (2018). Regional flood frequency analysis using support vector regression in arid and semi-arid regions of Iran. *Hydrological Sciences Journal*, *63*(3), 426–440. https://doi.org/10.1080/02626667.2018.1432056

Gautam, V. K., Gaurav, P. K., Murugan, P., & Annadurai, M. (2015). Assessment of Surface Water Dynamicsin Bangalore using WRI, NDWI, MNDWI,Supervised Classification and K-T Transformation. *Aquatic Procedia*, *4*, 739–746. https://doi.org/10.1016/j.aqpro.2015.02.095

Gebregiorgis, A. S., Moges, S. A., & Awulachew, S. B. (2013). Basin Regionalization for the Purpose of Water Resource Development in a Limited Data Situation: Case of Blue Nile River Basin, Ethiopia. *Journal of Hydrologic Engineering*, *18*(10), 1349–1359. https://doi.org/10.1061/(ASCE)HE.1943-5584 .0000730

Getahun, Y. S., & Gebre, S. L. (2015). Flood hazard assessment and mapping of flood inundation area of the Awash River Basin in Ethiopia using GIS and HEC-GeoRAS/HEC-RAS model. *Journal of Civil & Environmental Engineering*, *5*(4), 1.

Gezahegn, M. A., & Suryabhagavan, K. V. (2018). Developing Flood Hazard Forecasting and Early Warning System in Dire Dawa, Ethiopia. *Int. J. Adv. Multidiscip. Res*, *5*(9), 11–27.

Goodell, C. (2014, June 24). Critical Depth Primer – Kleinschmidt. *The RAS Solution*. https://www.kleinschmidtgroup.com/ras-post/critical-depth-primer/

Haan, C. (2002). *Statistical Methods in Hydrology* (2nd edn.). Iowa State Press.

Haestad. (2003). *Floodplain modeling using HEC-RAS*. Haestad Press. https://scholar.google.com/scholar_lookup?title=Floodplain+modeling+using+HEC-RAS&author=Dyhouse%2C+Gary.&publication_year=2003

Haile, D. B. (2009). *Flood risk assessment, prediction and control measures on the Middle Awash River, Ethiopia* [PhD Thesis]. Unesco-IHE.

Hailegeorgis, T. T., & Alfredsen, K. (2017). Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway. *Journal of Hydrology: Regional Studies*, 104–126. https://doi.org/10.1016/j.ejrh.2016.11.004

Hajibayov, F., Ozkul, B. D., & Terzi, F. (2017). Floodplain Modeling and Mapping Using The Geographical Information Systems (GIS) and Hec-RAS/Hec-GeoRAS Applications. Case of Edirne, Turkey. *Geo Inf Sci Res–UK (GISRUK). The University of Manchester Ukraine*, *9*, 9.

Haktanir, T., Bajabaa, S., & Masoud, M. (2013). Stochastic analyses of maximum daily rainfall series recorded at two stations across the Mediterranean Sea. *Arabian Journal of Geosciences*, *6*(10), 3943–3958. https://doi.org/10.1007/s12517-012-0652-0

Haque, M. M., Seidou, O., Mohammadian, A., & Djibo, A. G. (2020). Development of a time-varying MODIS/2D hydrodynamic model relationship between water levels and flooded areas in the Inner Niger Delta, Mali, West Africa. *Journal of Hydrology: Regional Studies*, *30*, 100703.

Helsel, D. R., & Hirsch, R. M. (1992). *Statistical methods in water resources*.

Hesarkazzazi, S., Arabzadeh, R., Hajibabaei, M., Rauch, W., Kjeldsen, T. R., Prosdocimi, I., Castellarin, A., & Sitzenfrei, R. (2021). Stationary vs non-stationary modelling of flood frequency distribution across northwest England. *Hydrological Sciences Journal*, *64*(4), 729–744, DOI. https://doi.org/10.1080/02626667.2021.1884685

Hooda, E., Hooda, B. K., & Tanwar, N. (2018). *Probability weighted moments (PWMs) and partial probability weighted moments (PPWMs) of type-II extreme value distribution*.

Hosking, J. R. M. (2005). *Research Report Fortran routines for use with the method of L-moments Version 3.04*. IBM Research Division, TJ Watson Research Center, New York.

Hosking, J. R. M. (2022). *LmomRFA: Regional flood frequency analysis using L-moments*. https://cran.r-project.org/web/packages/lmomRFA/

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis: An approach based on l-moments*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511529443

Husain, A. (2017). Estimation of design flood discharge for Kakkadavu Dam in Kariangode River Basin. *International Journal of Environmental Science & Natural Resources*, *4*(1), 555626.

Huţanu, E., Mihu-Pintilie, A., Urzica, A., Paveluc, L. E., Stoleriu, C. C., & Grozavu, A. (2020). Using 1D HEC-RAS Modeling and LiDAR Data to Improve Flood Hazard Maps Accuracy: A Case Study from Jijia Floodplain (NE Romania). *Water*, *12*(6), 1624.

Huxley, C. D. (2004). TUFLOW Testing and Validation. *Griffith University, Brisbane*.

Interagency Advisory Committee on Water Data. (1982). Guidelines for determining flood flow frequency. *Bulletin B*, *17*.

Ishak, E. H., Rahman, A., Westra, S., Sharma, A., & Kuczera, G. (2013). Evaluating the non-stationarity of Australian annual maximum flood. *Journal of Hydrology*, *494*, 134–145. https://doi.org/10.1016/j.jhydrol.2013.04.021

Jalayer, F., De Risi, R., De Paola, F., Giugni, M., Manfredi, G., Gasparini, P., Topa, M. E., Yonas, N., Yeshitela, K., & Nebebe, A. (2014). Probabilistic GIS-based method for delineation of urban flooding risk hotspots. *Natural Hazards*, *73*(2), 975–1001. https://doi.org/10.1007/s11069-014-1119-2

Jena, J., & Nath, S. (2020). An Empirical Formula for Design Flood Estimation of Un-Gauged Catchments in Brahmani Basin, Odisha. *Journal of The Institution of Engineers (India): Series A*, *101*(1), 1–6. https://doi.org/10.1007/s40030-019-00402-x

Ji, L., Zhang, L., & Wylie, B. (2009). Analysis of Dynamic Thresholds for the Normalized Difference Water Index. *Photogrammetric Engineering & Remote Sensing*, *75*, 1307–1317. https://doi.org/10.14358/PERS.75.11.1307

Ji, R. P., Yu, W. Y., Feng, R., Wu, J. W., & Zhang, Y. S. (2019). *The threshold determination methods of water body information extraction using GF-1 satellite image*. https://doi.org/10.1088/1757-899X/592/1/012088

Jie Yang, Ronald D. Townsend, & Bahram Daneshfar. (2006). Applying the HEC-RAS model and GIS techniques in river network floodplain delineation. *Can. J. Civ. Eng.*, *33*, 19–28. https://doi.org/10.1139/L05-102

Jingyi, Z., & Hall, M. J. (2004). Regional flood frequency analysis for the Gan-Ming River basin in China. *Hydrology*, *98–117*, 98–117. https://doi.org/10.1016/j.jhydrol.2004.03.018

Kachroo, R. K., Mkhandi, S. H., & Parida, B. P. (2000). Flood frequency analysis of southern Africa: I. Delineation of homogeneous regions. *Hydrological Sciences Journal*, *45*(3), 437–447. https://doi.org/10.1080/02626660009492340

Karim, M. F., & Chowdhury, J. U. (1995). Probabilistic division of storm surge flood prone area in the Ganges tidal plain. *United Nation Center for Regional Development, IFCDR, BUET, Bangladesh, Research Report, Dhaka*.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., & Brumby, S. P. (2021). Global land use / land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 4704–4707. https://doi.org/10.1109/IGARSS47720.2021.9553499

Khaing, T. W., Tantanee, S., Pratoomchai, W., & Mahavik, N. (2021). Coupling Flood Hazard with Vulnerability Map for Flood Risk Assessment: A Case Study of Nyaung-U Township in Myanmar. *Greater Mekong Subregion Academic and Research Network International Journal*, *15*, 127–138.

Khanna, R. K., Agrawal, C. K., Kumar, P., & Director, C. E. (2005). Remote sensing and GIS applications in flood management. *New Delhi: Central Water Commission*.

Khattak, M. S., Anwar, F., Saeed, T. U., Sharif, M., Sheraz, K., & Ahmed, A. (2016). Floodplain mapping using HEC-RAS and ArcGIS: A case study of Kabul River. *Arabian Journal for Science and Engineering*, *41*(4), 1375–1390.

Kim, V., Tantanee, S., & Suparta, W. (2020). Gis-based flood hazard mapping using hec-ras model: A case study of lower mekong river, cambodia. *Geographia Technica*, *15*(1).

Kisaka, M. O., Mucheru-Muna, M., Ngetich, F. K., Mugwe, J. N., Mugendi, D., & Mairura, F. (2015). Rainfall Variability, Drought Characterization, and Efficacy of Rainfall Data Reconstruction: Case of Eastern Kenya. *Advances in Meteorology*, *2015*, e380404. https://doi.org/10.1155/2015/380404

Kourgialas, N. N., & Karatzas, G. P. (2017). A national scale flood hazard mapping methodology: The case of Greece – Protection and adaptation policy approaches. *Science of The Total Environment*, *601–602*, 441–452. https://doi.org/10.1016/j.scitotenv.2017.05.197

Kuhnert, P., & Venables, B. (2005). *CSIRO Mathematical and Information Sciences An Introduction to R: Software for Statistical Modelling & Computing Course Materials and Exercises*.

Kundzewicz, Z. W. (2002). Non-structural Flood Protection and Sustainability. *Water International*, *27*(1), 3–13. https://doi.org/10.1080/02508060208686972

Lammesa, B. (2017). *Groundwater potential and Recharge zone mapping by using GIS and Remote sensing Techniques in the case of Middle Awash River Basins, Ethiopia* [Thesis, Addis Ababa Universty]. http://etd.aau.edu.et/handle/123456789/5318

Lee, D.-H., & Kim, N. W. (2019). Regional Flood Frequency Analysis for a Poorly Gauged Basin Using the Simulated Flood Data and L-Moment Method. *Water*, *11*(1717). http://dx.doi.org/10.3390/w11081717

Leščešen, I., & Dolinaj, D. (2019). Regional Flood Frequency Analysis of the Pannonian Basin. *Water*, *11*(2), Article 2. https://doi.org/10.3390/w11020193

Li, Z., Liu, H., Luo, C., & Fu, G. (2021). Assessing Surface Water Flood Risks in Urban Areas Using Machine Learning. *Water*, *13*(24), Article 24. https://doi.org/10.3390/w13243520

Likoko, B. A., Mbifo, N., Besango, L., Totiwe, T., Badjoko, D. H., Likoko, G., Botomo, D., Litemandia, N., Posho, B., & Alongo, S. (2019). Climate Change for Yangambi Forest Region, DR Congo. *J Aqua Sci Oceanography*, *1*, 203.

Lim, Y. H., & Voeller, D. L. (2009). Regional Flood Estimations in Red River Using L-Moment-Based Index-Flood and Bulletin 17B Procedures. *Journal of Hydrologic Engineering*, *14*(9), 1002–1016. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000102

Liu, W.-C., Hsieh, T.-H., & Liu, H.-M. (2021). Flood Risk Assessment in Urban Areas of Southern Taiwan. *Sustainability*, *13*(6), Article 6. https://doi.org/10.3390/su13063180

Liu, Y. (2012). *Why NDWI threshold varies in delineating water body from multitemporal images?* 4375–4378. https://doi.org/10.1109/IGARSS.2012.6350404

Lumbroso, D. (2020). Flood risk management in Africa. *Journal of Flood Risk Management*, *13*. https://doi.org/10.1111/jfr3.12612

Machiwal, D., & Jha, M. K. (2012). Methods for Time Series Analysis. In *Hydrologic Time Series Analysis: Theory and Practice* (pp. 51–84). Springer Netherlands. https://doi.org/10.1007/978-94-007-1861-6_4

Mahmoud, S. H., & Gan, T. Y. (2018). Multi-criteria approach to develop flood susceptibility maps in arid regions of Middle East. *Journal of Cleaner Production*, *196*, 216–229.

Malakeel, G. S., Abdu Rahiman, K. U., & Vishnudas, S. (2021). Flood Risk Assessment Methods—A Review. In J. Thomas, B. R. Jayalekshmi, & P. Nagarajan (Eds.), *Current Trends in Civil Engineering* (pp. 197–208). Springer. https://doi.org/10.1007/978-981-15-8151-9_19

Malekinezhad, H., & Zare-Garizi, A. (2014). Regional frequency analysis of daily rainfall extremes using L-moments approach. *Atmósfera*, *27*(4), 411–427. https://doi.org/10.1016/S0187-6236(14)70039-6

Mamo, S., Berhanu, B., & Melesse, A. M. (2019). Historical flood events and hydrological extremes in Ethiopia. In *Extreme Hydrology and Climate Variability* (pp. 379–384). Elsevier.

Maskong, H., Jothityangkoon, C., & Hirunteeyakul, C. (2019). Flood hazard mapping using on-site surveyed flood map, HEC-RAS V. 5 and GIS Tool: A case study of nakhon ratchasima municipality, Thailand. *International Journal of GEOMATE*, *16*(54), 1–8.

Mfwango, L. H., Salim, C. J., & Kazumba, S. (2018). Estimation of Missing River Flow Data for Hydrologic Analysis: The Case of Great Ruaha River Catchment. *Hydrology Current Research*, *9*(2), 299. https://doi.org/10.41172/2157-7587.1000299

Miller, J. D., & Hutchins, M. (2017). The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom. *Journal of Hydrology: Regional Studies*, *12*, 345–362. https://doi.org/10.1016/j.ejrh.2017.06.006

Mitiku, A. B., Woldyohanes, B., Mekonnen, S. A., Mekonnen, K., & Tadesse, A. (2020). The need for integrated flood management approach from social, economic and environmental perspectives: The case of Upper Awash River Basin. *International Journal of Water Resources and Environmental Engineering*, *12*(3), 57–70. https://doi.org/10.5897/IJWREE2020.0935

Modarres, R. (2006). Regional precipitation climates of Iran. *Journal of Hydrology*, *45*, 13–27. http://www.jstor.com/stable/43944936

Modarres, R. (2008). Regional Frequency Distribution Type of Low Flow in North of Iran by L-moments. *Water Resources Management*, *22*(7), 823–841. https://doi.org/10.1007/s11269-007-9194-8

MoWIE. (2015). *The project for ground water resources assessment in the middle Awash river basin in the federal democratic republic of Ethiopia: Final report executive summary*. https://vdocuments.mx/reader/full/the-project-for-groundwater-resources-the-middle-awash-river-basin-final-report

Nama, A. L., & Tesfaye, F. (2020). Flood Hazard Analysis on Dulecha Woreda Located in Middle Awash River Basin. *International Journal of Water Resources Engineering*, *6*(1), Article 1. https://doi.org/10.37628/jwre.v6i1.592

Namara, W. G., Damisse, T. A., & Tufa, F. G. (2021). Application of HEC-RAS and HEC-GeoRAS model for Flood Inundation Mapping, the case of Awash Bello Flood Plain, Upper Awash River Basin, Oromiya Regional State, Ethiopia. *Modeling Earth Systems and Environment*. https://doi.org/10.1007/s40808-021-01166-9

Nanesa, K. (2021). Awash River's the Ongoing Irrigation Practices, Future Projects and its Impacts on the Environment of Awash River Basin. *Irrigation and Drainage Systems Engineering*, *10*(6).

NDRMC. (2018). *Federal Democratic Republic of Ethiopia National Disaster Risk Management Commission, Early Warning and Emergency Response Directorate*. https://reliefweb.int/report/ethiopia/federal-democratic-republic-ethiopia-national-disaster-risk-management-commission

NDRMC. (2020). *National Disaster Risk Management Commission (NDRMC) Flood Alert # 1*. http://www.ndrmc.gov.et/

Nitheshnirmal, S., Thilagaraj, P., Rahaman, S. A., & Jegankumar, R. (2019). Erosion risk assessment through morphometric indices for prioritisation of Arjuna watershed using ALOS-PALSAR DEM. *Modeling Earth Systems and Environment*, *5*(3), 907–924. https://doi.org/10.1007/s40808-019-00578-y

Nogherotto, R., Fantini, A., Raffaele, F., Di Sante, F., Dottori, F., Coppola, E., & Giorgi, F. (2022). A combined hydrological and hydraulic modelling approach for the flood hazard mapping of the Po river basin. *Journal of Flood Risk Management*, *15*(1), e12755. https://doi.org/10.1111/jfr3.12755

Noto, L. V., & La Loggia, G. (2009). Use of L-Moments Approach for Regional Flood Frequency Analysis in Sicily, Italy. *Water Resources Management*, *23*(11), 2207–2229. https://doi.org/10.1007/s11269-008-9378-x

Obinna, A. I., & Ben, E. N. (2020). Regional Flood Frequency Analysis using Dimensionless Index Flood Method. *Civil Engineering Journal*, *6*(12), 2425–2436. http://dx.doi.org/10.28991/cej-2020-03091627

OCHA. (2020, September). *Gov't providing humanitarian assisstance for 600,000 flood affected persons*. https://reliefweb.int/report/ethiopia/gov-t-providing-humanitarian-assisstance-6000000-flood-affected-persons

Ogato, G. S., Bantider, A., Abebe, K., & Geneletti, D. (2020). Geographic information system (GIS)-Based multicriteria analysis of flooding hazard and risk in Ambo Town and its watershed, West shoa zone, oromia regional State, Ethiopia. *Journal of Hydrology: Regional Studies*, *27*, 100659.

Okoli, K. (2019). *Design flood estimation under uncertainty* [PhD Thesis, Acta Universitatis Upsaliensis]. http://urn.kb.se/resolce?urn=urn:nbn:se:uu:diva-390362

Ologunorisa, T. E., & Abawua, M. J. (2005). Flood Risk Assessment: A Review. *Journal of Applied Sciences and Environmental Management*, *9*(1), 57–63. http://www.bioline.org.br/abstract?ja05010

Ongdas, N., Akiyanova, F., Karakulov, Y., Muratbayeva, A., & Zinabdin, N. (2020). Application of HEC-RAS (2D) for Flood Hazard Maps Generation for Yesil (Ishim) River in Kazakhstan. *Water*, *12*(10), Article 10. https://doi.org/10.3390/w12102672

Papaioannou, G., Efstratiadis, A., Vasiliades, L., Loukas, A., Papalexiou, S. M., Koukouvinos, A., Tsoukalas, I., & Kossieris, P. (2018). An Operational Method for Flood Directive Implementation in Ungauged Urban Areas. *Hydrology, 5*, 24. https://doi.org/10.3390/hydrology5020024

Parsian, S., Amani, M., Moghimi, A., Ghorbanian, A., & Mahdavi, S. (2021). Flood Hazard Mapping Using Fuzzy Logic, Analytical Hierarchy Process, and Multi-Source Geospatial Datasets. *Remote Sensing*, *13*(23), Article 23. https://doi.org/10.3390/rs13234761

Pathan, A. I., & Agnihotri, P. G. (2019). A Combined Approach For 1-D Hydrodynamic Flood Modeling By using Arc-Gis, Hec-Georas, Hec-Ras Interface—A Case Study On Purna River Of Navsari City, Gujarat. *International Journal of Recent Technology and Engineering (IJRTE)*, *8*(1).

Pojani, E., & Keçi, J. (2016). *Disaster Risk Management in the Western Balkans: A comprehensive approach on technical and economic perspectives*. University of Novi Sad, Faculty of Technical Sciences,.

Prinos, P. (2008). *Review of Flood Hazard Mapping*. FLOODsite. www.floodsite.net

Quesada-Román, A., Ballesteros-Cánovas, J. A., Granados-Bolaños, S., Birkel, C., & Stoffel, M. (2022). Improving regional flood risk assessment using flood frequency and dendrogeomorphic analyses in mountain catchments impacted by tropical cyclones. *Geomorphology*, *396*, 108000. https://doi.org/10.1016/j.geomorph.2021.108000

Quirós, E., & Gagnon, A. S. (2020). Validation of flood risk maps using Open source optical and radar satellite imagery. *Transactions in GIS*, *24*(5), 1208–1226. https://doi.org/10.1111/tgis.12637

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. https://www.R-project.org/. https://www.R-project.org/

Raes, D., Willems, P., & Gbaguidi, F. (2006). RAINBOW–A software package for hydrometeorological frequency analysis and testing the homogeneity of historical data sets. *Proceedings of the 4th International Workshop on Sustainable Management of Marginal Drylands. Islamabad, Pakistan*, *2731*, 12.

Rahman, Md. R., & Thakur, P. K. (2018). Detecting, mapping and analysing of flood water propagation using synthetic aperture radar (SAR) satellite data and GIS: A case study from the Kendrapara

District of Orissa State of India. *The Egyptian Journal of Remote Sensing and Space Science*, *21*, S37–S41. https://doi.org/10.1016/j.ejrs.2017.10.002

Rahmati, O., Zeinivand, H., & Besharat, M. (2016). Flood hazard zoning in Yasooj region, Iran, using GIS and multi-criteria decision analysis. *Geomatics, Natural Hazards and Risk*, *7*(3), 1000–1017. https://doi.org/10.1080/19475705.2015.1045043

Rangari, V. A., Sridhar, V., Umamahesh, N. V., & Patel, A. K. (2019). Floodplain mapping and management of urban catchment using HEC-RAS: A case study of Hyderabad City. *Journal of the Institution of Engineers (India): Series A*, *100*(1), 49–63.

Rao, A. R., & Hamed, K. H. (2000). *Flood Frequency Analysis* (1st ed.). CRC Press. https://doi.org/10.1201/9780429128813

Rezende de Souzaa, G., Merwade, V., Coutinho de Oliveira, L. F., Violab, M. R., & Fariasb, M. de S. (2021). Regional flood frequency analysis and uncertainties: Maximum streamflow estimates in ungauged basins in the region of Lavras, MG, Brazil. *Catena*.

Romali, N. S., & Yusop, Z. (2017). Frequency analysis of annual maximum flood for Segamat River. *MATEC Web of Conferences*, *103*, 04003.

Roy, B., Khan, M., Islam, A. K. M. S., Khan, U., & Mohammed, K. (2021). Integrated flood risk assessment of the Arial Khan River under changing climate using IPCC AR5 risk framework. *Journal of Water and Climate Change*, *12*(7). https://doi.org/10.2166/wcc.2021.341

Saf, B. (2009). Regional Flood Frequency Analysis Using L-Moments for the West Mediterranean Region of Turkey. *Water Resources Management*, *23*(3), 531–551. https://doi.org/10.1007/s11269-008-9287-z

San, Z. M. L. T., Zin, W. W., Kawasaki, A., Acierto, R. A., & Oo, T. Z. (2020). Developing Flood Inundation Map Using RRI and SOBEK Models: A Case Study of the Bago River Basin, Myanmar. *Journal of Disaster Research*, *15*(3), 277–287. https://doi.org/10.20965/jdr.2020.p0277

Sarhadi, A., Soltani, S., & Modarres, R. (2012). Probabilistic flood inundation mapping of ungauged rivers: Linking GIS techniques and frequency analysis. *Journal of Hydrology*, 68–86. http://dx.doi.org/10.1016/j.jhydrol.2012.06.039

Searcy, J. K., & Hardison, C. H. (1960). *Double-mass corves*. US Government Printing Office.

Selvanathan, S., & Dymond, R. L. (2010). FloodwayGIS: An ArcGIS Visualization Environment to Remodel a Floodway. *Transactions in GIS*, *14*(5), 671–688. https://doi.org/10.1111/j.1467-9671.2010.01225.x

Shah, M. a. r., Rahman, A., & Chowdhury, S. h. (2018). Challenges for achieving sustainable flood risk management. *Journal of Flood Risk Management*, *11*(S1), S352–S358. https://doi.org/10.1111/jfr3.12211

Sharkey, J. (2014). Investigating Instabilities with HEC-RAS Unsteady Flow Modeling for Regulated Rivers at Low Flow Stages. *Masters Theses*. https://trace.tennessee.edu/utk_gradthes/3183

Sholihah, Q., Kuncoro, W., Wahyuni, S., Puni Suwandi, S., & Dwi Feditasari, E. (2020). The analysis of the causes of flood disasters and their impacts in the perspective of environmental law. *IOP Conference Series: Earth and Environmental Science*, *437*(1), 012056. https://doi.org/10.1088/1755-1315/437/1/012056

Shustikova, I., Domeneghetti, A., Neal, J. C., Bates, P., & Castellarin, A. (2019). Comparing 2D capabilities of HEC-RAS and LISFLOOD-FP on complex topography. *Hydrological Sciences Journal*, *64*(14), 1769–1782. https://doi.org/10.1080/02626667.2019.1671982

Sine, A., & Ayalew, S. (2004). *Identification and delination of hydrological homogeneous regions—The case of Blue Nile River basin*. *4*, 59–72.

Sinshaw, B. G., Moges, M. A., kinde Tefera, A., Atanaw, S. B., Fenta, H. M., Bazezew, H. A., Eshete, D. G., & Legese, K. G. (2018). *Analysis of regional flood frequency and its socio-economic impact in Ethiopia*.

Sohl, T., & Sleeter, B. (2012). Role of remote sensing for land-use land-cover change modelling. In *Remote sensing of land use and land cover: Principles and applications* (pp. 225–239).

Stedinger, J. R. (2000). Flood Frequency Analysis and Statistical Estimation of Flood Risk. *Cambridge University Press*, 334–358. https://doi.org/10.1017/CBO9780511529412.013

Stone, R. J. (2001). Changing seasonal rainfall patterns in Trinidad: Myth or reality? *West Indian J Eng*, *23*(2), 9–16.

Tadese, M. T., Kumar, L., Koech, R., & Zemadim, B. (2019). Hydro-Climatic Variability: A Characterisation and Trend Study of the Awash River Basin, Ethiopia. *Hydrology*, *6*(2), Article 2. https://doi.org/10.3390/hydrology6020035

Tamang, N. B., & Tamrakar, N. K. (2017). Use of GIS and HEC-RAS tools in the flood hazard mapping of an ungauged river: A case study of the Malekhu Khola, central Nepal. *International Journal Of Advanced Research in Engineering & Management (IJAREM)*, *03*(07), 57–64.

Tamiru, H., & Dinka, M. O. (2021). Application of ANN and HEC-RAS model for flood inundation mapping in lower Baro Akobo River Basin, Ethiopia. *Journal of Hydrology: Regional Studies*, *36*, 100855. https://doi.org/10.1016/j.ejrh.2021.100855

Tazin, T. (2018). Flood Hazard Mapping of Dharla River Floodplain Using HEC-RAS 1D/2D Coupled Model. *M. Sc. Engg. Thesis, Bangladesh University of Engineering and Technology (BUET), Dhaka*.

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F. W., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling & Software*, *90*, 201–216. https://doi.org/10.1016/j.envsoft.2017.01.006

Tessema, S. M., Setegn, S. G., & Mörtberg, U. (2015). Watershed modeling as a tool for sustainable water resources management: SWAT model application in the Awash River basin, Ethiopia. In *Sustainability of Integrated Water Resources Management* (pp. 579–606). Springer.

Thoummalangsy, S., Tuankrua, V., & Pukngam, S. (2019). Flood Risk Mapping Using HEC-RAS and GIS Technique: Case of the Xe Bangfai Floodplain, Khammoune Province, Lao PDR. *Thai Environmental Engineering Journal*, *33*(3), Article 3. https://so05.tci-thaijo.org/index.php/teej/article/view/202311

Tingsanchali, T., & Karim, F. (2010). Flood-hazard assessment and risk-based zoning of a tropical flood plain: Case study of the Yom River, Thailand. *Hydrological Sciences Journal*, *55*(2), 145–161. https://doi.org/10.1080/02626660903545987

Tingsanchali, T., & Karim, M. F. (2005). Flood hazard and risk analysis in the southwest region of Bangladesh. *Hydrological Processes*, *19*(10), 2055–2069. https://doi.org/10.1002/hyp.5666

Tramblay, Y., Villarini, G., & Zhang, W. (2020). Observed changes in flood hazard in Africa. *Environmental Research Letters*, *15*(10), 1040b5. https://doi.org/10.1088/1748-9326/abb90b

UNDP. (1999). *Ethiopia Situation Report for period Aug 1999—Ethiopia*. ReliefWeb. https://reliefweb.int/report/ethiopia/ethiopia-situation-report-period-aug-1999

UNDRO. (1991). *Mitigating natural disasters: Phenomena, effects and options: A manual for policy makers and planners*. Office of the United Nations Disaster Relief Coordinator.

USACE. (2018). *Analyzing Flood Risk for Forecast Informed Reservoir Operations in the Russian River Watershed Using HEC-WAT.* Institute for Water Resources Hydrologic Engineering Center: Institute for Water Resources Hydrologic Engineering Center.

USWRC, U. S. W. R. C., Hydrology Comittee. (1981). *Guidelines for determining flood flow frequency*. U.S. Gov. Print. Off.

Ven Chow, David Maidment, & Larry Mays. (1988). *Applied Hydrology*. McGraw-Hill Science/Engineering/Math.

Viglione, A., Laio, F., & Claps, P. (2007). A comparison of homogeneity tests for regional frequency analysis. *Water Resources Research*, *43*(3). https://doi.org/10.1029/2006WR005095

Vojtek, M., Petroselli, A., Vojteková, J., & Asgharinia, S. (2019). Flood inundation mapping in small and ungauged basins: Sensitivity analysis using the EBA4SUB and HEC-RAS modeling approach. *Hydrology Research*, *50*(4), 1002–1019.

Vojtek, M., & Vojteková, J. (2016). Flood hazard and flood risk assessment at the local spatial scale: A case study. *Geomatics, Natural Hazards and Risk, 7*(6), 1973–1992. https://doi.org/10.1080/19475705.2016.1166874

Wondafrash, W. (2021). *Evaluation of existing water conveyance, storage and control structures: A case of amibara irrigation scheme, Ethiopia* [Thesis, ASTU]. http://213.55.101.20:8080/xmlui/handle/123456789/1800

Wondim, Y. K. (2016). Flood Hazard and Risk Assessment Using GIS and Remote Sensing in Lower Awash Sub-basin, Ethiopia. *Journal of Environment and Earth Science*, *6*(9).

Wright, D. B. (2015). *Methods in Flood Hazard and Risk Assessment*. World Bank. https://doi.org/10/25136493/methods-flood-hazard-risk-assessment

Zoka, M., Psomiadis, E., & Dercas, N. (2018). The Complementary Use of Optical and SAR Data in Monitoring Flood Events and Their Effects. *Proceedings*, *2*(11), Article 11. https://doi.org/10.3390/proceedings2110644

APPENDIX

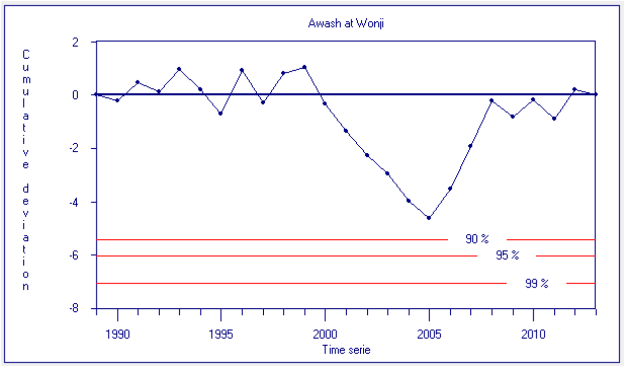
Appendix A: Distributions and their cumulative distribution functions (Hosking, 2005)

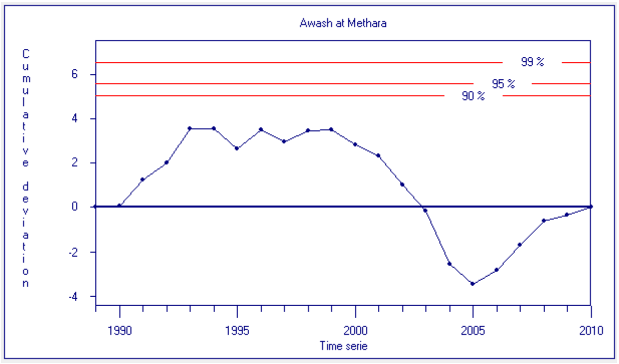
Distribution	Code	Number of parameters	Parameters	$F(x),\ x(F)$
Exponential	EXP	2	$\xi\ \alpha$	$F = 1 - \exp\{-(x-\xi)/\alpha\}$ $x = \xi - \alpha \log(1-F)$
Gamma	GAM	2	$\alpha\ \beta$	$F = G(x/\beta, \alpha)$ $x(F)$ not explicitly defined
Generalized extreme-value	GEV	3	$\xi\ \alpha\ k$	$F = \exp[-\{1-k(x-\xi)/\alpha\}^{1/k}]$ $x = \xi + \alpha\{1-(-\log F)^k\}/k$
Generalized logistic	GLO	3	$\xi\ \alpha\ k$	$F = 1/[1+\{1-k(x-\xi)/\alpha\}^{1/k}]$ $x = \xi + \alpha[1-\{(1-F)/F\}^k]/k$
Generalized Normal	GNO	3	$\xi\ \alpha\ k$	$F = \Phi[-k^{-1}\log\{1-k(x-\xi)/\alpha\}]$ $x(F)$ not explicitly defined
Generalized Pareto	GPA	3	$\xi\ \alpha\ k$	$F = 1-\{1-k(x-\xi)/\alpha\}^{1/k}$ $x = \xi + \alpha\{1-(1-F)^k\}/k$
Gumbel	GUM	2	$\xi\ \alpha$	$F = \exp[-\exp\{-(x-\xi)/\alpha\}]$ $x = \xi + \alpha \log(-\log F)$
Kappa	KAP	4	$\xi\ \alpha\ k\ h$	$F = [1-h\{1-k(x-\xi)/\alpha\}^{1/k}]^{1/h}$ $x = \xi + \alpha[1-\{(1-F)^h/h\}^k]/k$
Normal	NOR	2	$\mu\ \sigma$	$F = \Phi\{(x-\mu)/\sigma\}$ $x(F)$ not explicitly defined
Pearson type III	PE3	3	$\mu\ \sigma\ \gamma$	$F = G\big((x-\mu+2\sigma/\gamma)/\lvert\tfrac{1}{2}\sigma\gamma\rvert, 4/\gamma^2\big), \quad \gamma > 0$ $F = 1-G\big(-(x-\mu+2\sigma/\gamma)/\lvert\tfrac{1}{2}\sigma\gamma\rvert, 4/\gamma^2\big), \ \gamma<0$ $x(F)$ not explicitly defined
Wakeby	WAK	5	$\xi\ \alpha\ \beta\ \gamma\ \delta$	$F(x)$ not explicitly defined $x = \xi + \frac{\alpha}{\beta}\{1-(1-F)^\beta\} - \frac{\gamma}{\delta}\{1-(1-F)^{-\delta}\}$

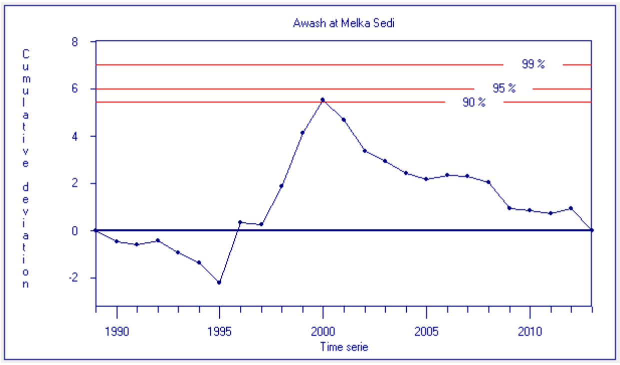
$G(x,\alpha) = \{\Gamma(\alpha)\}^{-1}\int_0^x t^{\alpha-1}e^{-t}\,dt$ is the incomplete gamma integral.

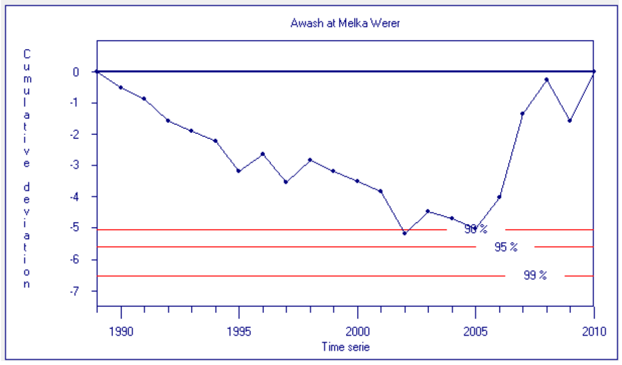
$\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^x \exp(-t^2/2)\,dt$ is the standard Normal cumulative distribution function.

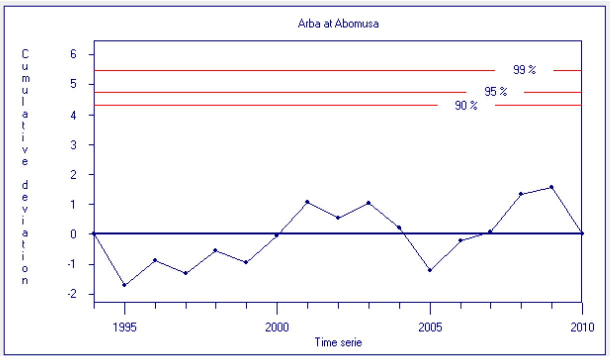
Appendix B: W-W homogeneity test results

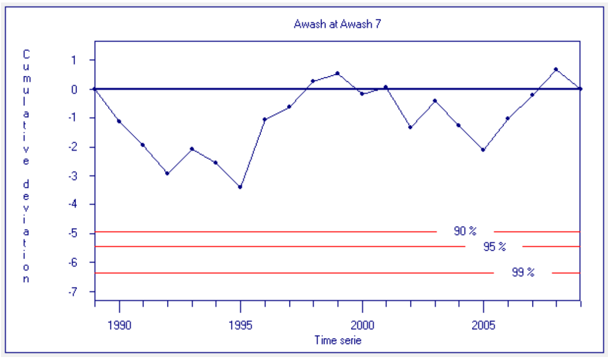


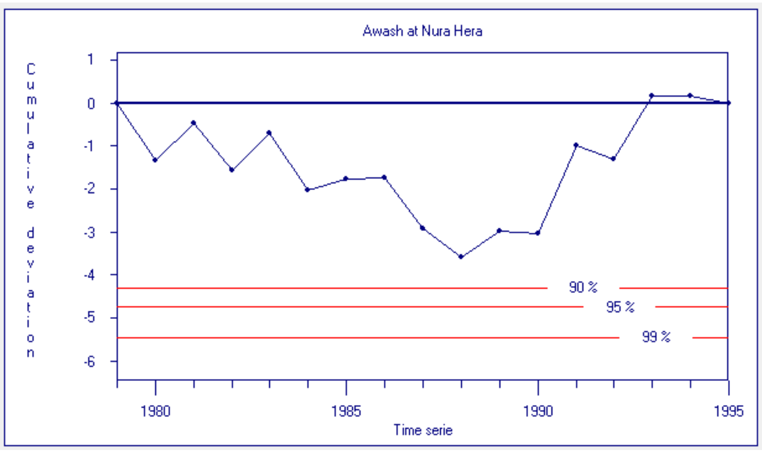












Appendix C: 10% significance level K_N values (USWRC, 1981)

Sample size	K_N value	Sample size	K_N value	Sample size	K_N value	Sample size	K_N value
10	2.036	45	2.727	80	2.940	115	3.064
11	2.088	46	2.736	81	2.945	116	3.067
12	2.134	47	2.744	82	2.949	117	3.070
13	2.175	48	2.753	83	2.953	118	3.073
14	2.213	49	2.760	84	2.957	119	3.075
15	2.247	50	2.768	85	2.961	120	3.078
16	2.279	51	2.775	86	2.966	121	3.081
17	2.309	52	2.783	87	2.970	122	3.083
18	2.335	53	2.790	88	2.973	123	3.086
19	2.361	54	2.798	89	2.977	124	3.089
20	2.385	55	2.804	90	2.981	125	3.092
21	2.408	56	2.811	91	2.984	126	3.095
22	2.429	57	2.818	92	2.989	127	3.097
23	2.448	58	2.824	93	2.993	128	3.100
24	2.467	59	2.831	94	2.996	129	3.102
25	2.486	60	2.837	95	3.000	130	3.104
26	2.502	61	2.842	96	3.003	131	3.107
27	2.519	62	2.849	97	3.006	132	3.109
28	2.534	63	2.854	98	3.011	133	3.112
29	2.549	64	2.860	99	3.014	134	3.114
30	2.563	65	2.866	100	3.017	135	3.116
31	2.577	66	2.871	101	3.021	136	3.119
32	2.591	67	2.877	102	3.024	137	3.122
33	2.604	68	2.883	103	3.027	138	3.124
34	2.616	69	2.888	104	3.030	139	3.126
35	2.628	70	2.893	105	3.033	140	3.129
36	2.639	71	2.897	106	3.037	141	3.131
37	2.650	72	2.903	107	3.040	142	3.133
38	2.661	73	2.908	108	3.043	143	3.135
39	2.671	74	2.912	109	3.046	144	3.138
40	2.682	75	2.917	110	3.049	145	3.140
41	2.692	76	2.922	111	3.052	146	3.142
42	2.700	77	2.927	112	3.055	147	3.144
43	2.710	78	2.931	113	3.058	148	3.146
44	2.719	79	2.935	114	3.061	149	3.148

Appendix D: Manning's n values (Chow, 1959)

Type of Channel and Description	Minimum	Normal	Maximum
Natural streams - minor streams (top width at floodstage < 100 ft)			
Main Channels			
a. clean, straight, full stage, no rifts or deep pools	0.025	0.030	0.033
b. same as above, but more stones and weeds	0.030	0.035	0.040
c. clean, winding, some pools and shoals	0.033	0.040	0.045
d. same as above, but some weeds and stones	0.035	0.045	0.050
e. same as above, lower stages, more ineffective slopes and sections	0.040	0.048	0.055
f. same as "d" with more stones	0.045	0.050	0.060
g. sluggish reaches, weedy, deep pools	0.050	0.070	0.080
h. very weedy reaches, deep pools, or floodways with heavy stand of timber and underbrush	0.075	0.100	0.150
Floodplains			
a. Pasture, no brush			
1.short grass	0.025	0.030	0.035
2. high grass	0.030	0.035	0.050
b. Cultivated areas			
1. no crop	0.020	0.030	0.040
2. mature row crops	0.025	0.035	0.045
3. mature field crops	0.030	0.040	0.050
c. Brush			
1. scattered brush, heavy weeds	0.035	0.050	0.070
2. light brush and trees, in winter	0.035	0.050	0.060
3. light brush and trees, in summer	0.040	0.060	0.080
4. medium to dense brush, in winter	0.045	0.070	0.110
5. medium to dense brush, in summer	0.070	0.100	0.160
d. Trees			
1. dense willows, summer, straight	0.110	0.150	0.200
2. cleared land with tree stumps, no sprouts	0.030	0.040	0.050
3. same as above, but with heavy growth of sprouts	0.050	0.060	0.080
4. heavy stand of timber, a few down trees, little undergrowth, flood stage below branches	0.080	0.100	0.120
5. same as 4. with flood stage reaching branches	0.100	0.120	0.160

Appendix E: Description of Sentinel-2 LULC data

1. Water: Areas where water was predominantly present throughout the year; may not cover areas with sporadic or ephemeral water; contains little to no sparse vegetation, no rock outcrop nor built up features like docks; examples: rivers, ponds, lakes, oceans, flooded salt plains.

2. Trees: Any significant clustering of tall (~15-m or higher) dense vegetation, typically with a closed or dense canopy; examples: wooded vegetation, clusters of dense tall vegetation within savannas, plantations, swamp or mangroves.

3. Grass: Open areas covered in homogenous grasses with little to no taller vegetation; wild cereals and grasses with no obvious human plotting (i.e., not a plotted field); e.g., natural meadows, fields with sparse to no tree cover, open savanna with few to no trees, parks/golf courses/lawns, pastures.

4. Flooded vegetation: Areas of any type of vegetation with obvious intermixing of water throughout a majority of the year; seasonally flooded area that is a mix of grass/shrub/trees/bare ground; examples: flooded mangroves, emergent vegetation, rice paddies and other heavily irrigated and inundated agriculture.

5. Crops: Human planted/plotted cereals, grasses, and crops not at tree height; examples: corn, wheat, soy, fallow plots of structured land.

6. Scrub/shrub: Mix of small clusters of plants or single plants dispersed on a landscape that shows exposed soil or rock; scrub-filled clearings within dense forests that are clearly not taller than trees; examples: moderate to sparse cover of bushes, shrubs and tufts of grass, savannas with very sparse grasses, trees or other plants

7. Built Area: Human made structures; major road and rail networks; large homogenous impervious surfaces including parking structures, office buildings and residential housing; examples: houses, dense villages / towns / cities, paved roads, asphalt.

8. Bare ground: Areas of rock or soil with very sparse to no vegetation for the entire year; large areas of sand and deserts with no to little vegetation; examples: exposed rock or soil, desert and sand dunes, dry salt flats/pans, dried lake beds, mines.

9. Snow/Ice: Large homogenous areas of permanent snow or ice, typically only in mountain areas or highest latitudes; examples: glaciers, permanent snowpack, snow fields.

10. Clouds: No land cover information due to persistent cloud cover.

11. Rangeland: Open areas covered in homogenous grasses with little to no taller vegetation; wild cereals and grasses with no obvious human plotting; examples: natural meadows and fields with sparse to no tree cover, open savanna with few to no trees, parks/golf courses/lawns, pastures. Mix of small clusters of plants or single plants dispersed on a landscape that shows exposed soil or rock; scrub-filled clearings within dense forests that are clearly not taller than trees; examples: moderate to sparse cover of bushes, shrubs and tufts of grass, savannas with very sparse grasses, trees or other plants.

Appendix F. Land Unit level H_F and V_F values

			Hazard Factor, HF, for return periods								
	Land units	Woreda	2- year	5-year	10-year	20- year	50-year	100-year	200- year	500- year	VF
1	Achamo Gulo Weniz	Jeju	2.41	2.417	2.423	2.341	2.355	2.363	2.37	2.362	1.529
2	Adulala Boku	Adama	2	1.975	1.302	1.302	1.302	1.302	1.302	1.302	0.758
3	Alaga Dore	Jeju	1.952	2.08	2.078	1.977	1.949	1.972	2.018	2.039	1.723
4	Alge	Fentale	1.758	2.039	2.061	1.983	1.99	2.011	2.029	1.963	2.859
5	Ararso Bero	Boset	2.743	2.799	2.784	2.774	2.833	2.851	2.856	2.805	1.625
6	Asebe Hari	Dulecha	2.691	2.691	2.537	2.537	2.537	2.537	2.697	2.697	2.169
7	Asoba	Amibara	2.379	2.379	2.182	2.182	2.182	2.182	2.463	2.463	3.338
8	Awash 7 killo town	Fentale	2.931	2.932	2.834	2.935	2.935	2.931	2.931	2.931	4.799
9	Awash Arba	Amibara	2.497	2.558	2.572	2.554	2.459	2.458	2.566	2.577	0.091
10	Awash B ishola	Dodota	2.014	2.123	2.103	1.996	1.996	1.995	1.966	1.993	1.73
11	Awash biherawi park	Fentale	2.731	2.776	2.8	2.78	2.805	2.805	2.804	2.79	0.539
12	Awash Melkasa Town	Adama	2.425	2.35	2.368	2.34	2.337	2.273	2.113	2.021	4.728
13	Awash Sheleko	Amibara	2.178	2.37	2.374	2.341	2.326	2.327	2.327	2.488	4.836
14	Bati Germama	Adama	2.747	2.813	2.82	2.817	2.82	2.82	2.819	2.82	1.548
15	Batu Degaga	Adama	2.408	2.406	2.383	2.384	2.387	2.387	2.573	2.55	1.958
16	Bedul Ale	Amibara	2.225	2.208	2	1.982	1.982	1.993	2.203	2.332	3.15
17	Boku Kurabo	Fentale	1.921	1.609	1.64	1.588	1.616	1.622	1.627	1.641	1.178
18	Bole Town	Boset	2.007	2.235	2.239	2.124	2.196	2.22	2.214	2.252	9.694
19	Bolhamo	Dulecha	2.421	2.362	2.331	2.292	2.33	2.332	2.522	2.522	0.082
20	Bonta	Amibara	2.318	2.309	2.109	2.102	2.106	2.121	2.425	2.425	3.025
21	Bose Doche	Boset	1.995	1.686	1.864	1.666	1.691	1.699	1.772	1.799	0.212
22	Buri	Amibara	2.278	2.411	2.246	2.233	2.227	2.23	2.408	2.449	3.392
23	Burtali	Dulecha	2.053	2.084	2.194	2.135	2.161	2.154	2.144	2.132	1.1
24	Choka Alem Tena	Adama	2.088	2.465	1.843	1.649	1.649	1.649	1.649	1.649	0.469
25	Degaga 01	Boset	2.422	2.422	2.412	2.412	2.412	2.412	2.59	2.648	3.038
26	Dere Qiletu	Dodota	2.014	2.216	2.239	2.115	2.102	2.119	2.131	2.154	0.069
27	Diduba	Fentale	2.755	2.8	2.822	2.803	2.813	2.817	2.825	2.816	0.64
28	Diho	Fentale	1.997	2.153	2.081	1.987	1.932	1.933	1.962	1.98	2.793
29	Dire Degaga	Boset	2.485	1.454	2.332	2.332	2.332	2.332	2.568	2.476	1.655
30	Doni Town	Boset	2.43	2.571	2.579	2.562	2.561	2.564	2.584	2.603	7.62
31	Ebesta Ouduga	Sire	2	2.05	2.05	1.44	2.04	1.7	2.2	2.314	0.875
32	Eibile	Amibara	2.135	1.977	2.294	1.962	2.273	2.264	2.258	2.252	5.201
33	Gara Dima	Fentale	2.301	2.452	2.499	2.447	2.468	2.498	2.534	2.577	0.403
34	Gederana Kube	Fentale	2.311	2.41	2.459	2.381	2.437	2.455	2.472	2.511	2.395
35	Gelcha	Fentale	2.085	2.386	2.388	2.371	2.379	2.378	2.383	2.232	0.87
36	Gerbo Hafe	Dulecha	2.02	2.02	1.555	1.555	1.555	1.555	2.103	2.12	3.014
37	Gola	Fentale	1.794	2.156	2.163	1.955	2.079	2.13	2.125	2.09	2.135
38	Golegota	Merti	2.018	2.213	2.112	2.021	2.036	2.044	2.059	2.178	4.538
39	Gumibi	Meiso	2.834	2.795	2.823	2.821	2.845	2.788	2.783	2.773	4.798
40	Halay Sumale	Amibara	2.215	2.126	2.365	2.116	2.192	2.368	2.396	2.396	3.983
41	Haro Adi Town	Fentale	1.834	2.026	2.113	1.91	1.884	1.723	1.741	1.767	7.832
42	Hugub	Dulecha	2.236	2.116	2.267	2.036	2.137	2.225	2.252	2.311	3.251
43	Huruta Dore	Jeju	1.9	2.193	2.234	2.129	2.153	2.103	2.094	2.087	2.869
44	Kebena	Fentale	1.353	2	2	1.412	1.919	1.895	1.85	1.773	0.204
45	Kobo	Fentale	1.685	2.217	2.217	2.217	2.217	2.217	2.217	1.768	0.96
46	Koloba Biqa	Sire	2.649	2.75	2.778	2.773	2.75	2.75	2.76	2.804	1.752
47	koloba Bli	Sire	2.67	2.792	2.757	2.733	2.797	2.8	2.847	2.871	0.735
48	Koma Gidiyaro	Amibara	2.248	1.901	2.321	1.902	2.319	2.322	2.322	2.322	2.119
49	Kuriftu Eda	Adama	2.06	2.195	2.203	2.09	2.069	2.058	2.029	2.032	5.058
50	Kurkura	Amibara	2.835	2.853	2.877	2.873	2.872	2.851	2.857	2.858	0.076
51	Ledi Fate	Fentale	2.384	2.564	2.602	2.581	2.583	2.584	2.604	2.651	1.42
52	Megacha	Sire	2.692	2.759	2.77	2.729	2.739	2.74	2.762	2.781	0.708
53	Melka Sedi Town	Amibara	1.964	1.978	2.105	1.824	1.998	2.031	2.047	2.062	4.747
54	Melka Werer Town	Amibara	2.604	2.604	2.607	2.555	2.555	2.555	2.558	2.695	6.645
55	Merti Sikuar	Fentale	2.1	2.359	2.361	2.332	2.339	2.343	2.349	2.348	4.142
56	Metehara Town	Fentale	2.943	2.893	2.91	2.945	2.946	2.876	2.853	2.75	4.695
57	Nura Hase	Boset	1.912	2.103	2.112	2.011	2.044	2.074	2.122	2.147	0.969
58	Nura Hera 01	Boset	2.182	2.285	2.287	2.173	2.175	2.184	2.209	2.226	4.054
59	Nura Hera 02	Boset	2.245	2.282	2.204	2.137	2.143	2.146	2.154	2.294	2.256
60	Qawa Hara Mirqasa	Boset	1.775	1.985	1.976	1.801	1.769	1.779	1.789	1.783	0.275
61	Qechachule Guja	Boset	2.465	2.538	2.481	2.455	2.515	2.533	2.633	2.682	1.211
62	Qoboluto Tumsa	Adama	2.695	2.782	2.787	2.782	2.786	2.787	2.788	2.789	2.551
63	Qombe Gugisa	Boset	2.457	2.606	2.641	2.587	2.635	2.656	2.686	2.669	0.48
64	Qoro	Dodota	2.412	2.362	2.352	2.229	2.281	2.324	2.457	2.399	0.161
65	Sama Yearbto ader Akababi	Anchar	2.74	2.754	2.73	2.734	2.7	2.724	2.724	2.97	0.014
66	Sar Areda	Boset	2.677	2.741	2.76	2.746	2.754	2.752	2.751	2.766	1.331
67	Sarana Weba	Fentale	2.511	2.61	2.62	2.558	2.577	2.591	2.595	2.645	0.128
68	Seganto	Dulecha	3	1.375	2	1.375	2	2	2	2	1.182
69	Serkamo	Amibara	2.676	2.676	2.676	2.688	2.688	2.688	2.688	2.779	2.161
70	Sifa Bote	Boset	1.998	2.113	2.229	2.085	2.13	2.164	2.197	2.223	0.697
71	Sire Robe Town	Adama	2.684	2.724	2.753	2.725	2.76	2.771	2.768	2.772	5.668
72	Tebela Erisha Limat	Jeju	2.708	2.711	2.774	2.691	2.769	2.785	2.788	2.744	2.971
73	Techadola	Merti	1.854	2.18	2.218	2.018	2.072	2.074	2.082	2.079	0.732
74	Ulaga Melka Aba	Adama	2.259	2.373	2.418	2.348	2.357	2.388	2.406	2.443	3.366
75	Unknown kebele	Jeju	1.816	1.89	1.924	1.736	1.796	1.804	1.835	1.854	3.404
76	Waqemiya Tiyo	Adama	1.81	2.026	1.93	1.699	1.709	1.703	1.705	1.71	3.288
77	Wershqona	Merti	2.189	2.239	2.233	2.063	2.058	2.051	2.054	2.056	2.706
78	Wetro Deno	Merti	2.564	2.617	2.66	2.65	2.67	2.676	2.683	2.732	0.346
79	Wonji Gefersa Town	Adama	2.45	2.439	2.404	2.258	2.252	2.073	2.086	2.099	6.665
80	Wonji Shewa Sugar Cane Pla	Adama	2.003	1.909	1.709	1.562	1.562	1.556	1.555	1.556	3.295