

**ASSESSING WATER AVAILABILITY USING REMOTE SENSING IN SOUTH
RUKURU AND NORTH RUMPHI RIVER BASIN**

MSc (WATER RESOURCES MANAGEMENT AND DEVELOPMENT) THESIS

GOMI THOMAS GONDWE

MZUZU UNIVERSITY

JULY 2024

**ASSESSING WATER AVAILABILITY USING REMOTE SENSING IN SOUTH
RUKURU AND NORTH RUMPHI RIVER BASIN**

GOMI THOMAS GONDWE

(BSc Irrigation Engineering)

**A THESIS SUBMITTED TO THE FACULTY OF ENVIRONMENTAL SCIENCES OF MZUZU
UNIVERSITY IN FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF
MASTER OF SCIENCE DEGREE IN WATER RESOURCES MANAGEMENT AND
DEVELOPMENT**

MZUZU UNIVERSITY

JULY 2024

DECLARATION

I, Gomi Thomas Gondwe, hereby declare that this titled “*Assessing Water Availability Using Remote Sensing in South Rukuru and North Rumphu River Basin*” has been written by me and presents a record of my research work. All citations, references, and borrowed ideas have been duly acknowledged. It is being submitted in fulfillment of the requirements for the award of the Master of Science Degree (MSc) in Water Resources Management and Development at Mzuzu University. None of the present work has been submitted previously for any degree or examination in any other university.

Name of Candidate: GOMI THOMAS GONDWE

Signature _____

Date _____

CERTIFICATE OF COMPLETION

I, the undersigned, certify that this thesis is a result of the author's own work and that to the best of our knowledge, it has not been submitted for any academic qualification within Mzuzu University or elsewhere. The thesis is acceptable in form and content, and that satisfactory knowledge of the field covered herein was demonstrated by the candidate through an oral examination held on: **01st December, 2023**

Signature _____

Date _____

Major Supervisor: Dr. Brighton Austin Chunga (PhD)

Signature _____

Date _____

Co-Supervisor: Associate Professor Mavuto Tembo (PhD)

Signature _____

Date _____

Postgraduate Coordinator: Dr. Russel C. G. Chidya (PhD)

Signature _____

Date _____

Head of Department (Water and Sanitation): Dr. Brighton Austin Chunga (PhD)

DEDICATION

I dedicate this academic research to my family and friends

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisors Dr. Brighton Austin Chunga and Associate Professor Mavuto Tembo for their support and guidance rendered to me throughout the research period. I would like to extend my sincere thanks to all the lecturers and the staff members who assisted me in various activities to make this work a success.

Special thanks should go to my mother, whose love, care and desire was to see me educated up to university level. I would like to register my thanks to my father, who always reminds me that “systematic hard work and commitment lead to success”.

ABSTRACT

To effectively and efficiently assess water availability there is a need for good quality hydrological time series data. Globally there is a dense decline in hydrological data, which makes it difficult to assess water availability. Remotely acquired data can be an alternative data source. This study was conducted to assess water availability using remote sensing in South Rukuru and North Rumphi River Basin. To assess the water availability, a rainfall-runoff (soil moisture method) hydrological model was developed using the Water Evaluation and Planning (WEAP) modelling software. The WEAP inputs included the Global Land Data Assimilation System (GLDAS) climate datasets, Elevation data, and the Land Use Land Cover. The observed stream flow data was used for calibrating and validating the WEAP model. A hydrological data availability assessment was performed to check for gaps or missing data in the streamflow time series data. Although there were missing data, but with the threshold set, all stations under study had a good allowable overall data availability. The performance of the WEAP model simulation was assessed and the monthly measured and simulated streamflow statistics showed a positive strong relationship ($R^2 = 0.81$, NSE = 0.81, IA = 0.95, KGE = 0.89 and PBIAS of -3.8) at station 7G18. Similarly, there was a very good agreement between the monthly measured and simulated streamflow at station 7H3 ($R^2 = 0.94$, NSE = 0.93, IA = 0.98, KGE = 0.85 and PBIAS of -9.3). The hydrological response showed that in months at which the precipitation was at its peak in all the sub-basins, the water availability was also at peak. In general, The WEAP model has demonstrated the capability of GLDAS datasets to be used to assess water availability when there is insufficient, inconsistent or fragmented observed hydrological data.

Author Keywords: Water Availability, Remote Sensing, WEAP, Soil Moisture Method, GLDAS

ACRONYMS

AVHRR	Advanced Very High-Resolution Radiometer
CAD	Continuity in Available Data
CCI	Climate Change Initiative
DEM	Digital Elevation Model
ESA	European Space Agency
ET	Actual Evapotranspiration
ET _o	Reference Evapotranspiration
EU	European Union
FAO	Food Agriculture Organization
GES DISC	Goddard Earth Sciences Data and Information Services Centre
GIOVANNI	Geospatial Interactive Online Visualization And Analysis Infrastructure
GIS	Geographic Information System
GLDAS	Global Land Data Assimilation System
GoM	Government of Malawi
GRACE	Gravity Recovery and Climate Experiment
IA	Index of Agreement
ITCZ	Intertropical Convergence Zone
KGE	Kling and Gupta Efficiency
LCCS	Land Cover Classification System
LDA	Longest Data Availability
MOD16	MODIS Global Evapotranspiration Project
NASA	National Aeronautics and Space Administration

NetCDF	Network Common Data Form
NSE	Nash-Sutcliffe Coefficient of Efficiency
ODA	Overall Data Availability
PBIAS	Percent BIAS
R ²	Coefficient of Determination
RE	Relative Error
SDGs	Sustainable Development Goals
SEI	Stockholm Environment Institute
SRTM	Shuttle Radar Topography Mission
TRMM	Tropical Rainfall Measuring Mission
UN	United Nations
WEAP	“Water Evaluation And Planning” system
WGS84	World Geodetic System 1984
WMO	World Meteorological Organization
WRAs	Water Resource Areas
WRUs	Water Resource Units
WWAP	World Water Assessment Programme

TABLE OF CONTENTS

DECLARATION	i
CERTIFICATE OF COMPLETION	ii
DEDICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
ACRONYMS	vi
LIST OF FIGURES	xi
LIST OF TABLES	xiii
CHAPTER 1: INTRODUCTION	1
1.1. Background	1
1.2. Statement of the Problem	5
1.3 Aim of Study	6
1.3.1 Main Objective	6
1.3.2 Specific Objectives	6
1.4 Research Questions	6
1.5 Significance of the Study	6
1.6 Ethical Consideration	7
1.7 Research Thesis Outline	7
CHAPTER 2: LITERATURE REVIEW	9
2.1 Background	9
2.2 Water Resources Demand	9
2.3 The Need for Water Availability Assessments	11

2.4 Assessing Water Availability	12
2.4.1 The Water Evaluation and Planning (WEAP) Model.....	21
2.4.2 WEAP Model Building Process.....	26
2.4.3 Why WEAP Modelling Software	28
2.5 Limited and Declining Hydro-Meteorological Monitoring Networks	29
2.6 Remote Sensing Datasets	31
2.7 Previous Studies.....	33
CHAPTER 3: MATERIALS AND METHODS.....	37
3.1 Introduction	37
3.2 Study Area.....	37
3.2.1 Location and General Description	37
3.2.2 Significance of the basin.....	39
3.2.3 Basin vegetation, geology, soil and relief	39
3.2.4 Rainfall data monitoring within Basin	40
3.2.5 Hydro-metric network in the Basin	40
3.3 Research Design.....	42
3.4 Data Collection.....	42
3.4.1 WEAP Model Input Data.....	42
3.4.1.1 Remote Sensing Datasets.....	42
3.4.1.2 Surface Observations	43
3.5 Data Analysis	44
3.5.1 Observed Hydrological Data Availability Assessment.....	44

3.5.2 The WEAP Model Calibration and Validation	45
3.5.2 The WEAP Model Performance Evaluation Measures	47
3.6 Methodology Matrix	51
CHAPTER 4: RESULTS	53
4.1 Observed Hydrological Data Availability Assessment	53
4.1.1 Missing Data Proportions and Overall Data Availability	53
4.1.2 Longest Data Availability without Gaps	59
4.1.3 Continuity in Available Data	62
4.2 The Remote Sensing WEAP Model	64
4.2.1 WEAP Model: Calibration and Validation	64
4.2.2 WEAP Model Hydrological Processes	67
4.3 WEAP Model Performance Evaluation	70
CHAPTER 5: DISCUSSION	73
5.1 Observed Hydrological Data Availability Assessment	73
5.2 WEAP Model Hydrological Processes	75
5.3 WEAP Model Performance Evaluation	76
CHAPTER 6: CONCLUSION AND RECOMMENDATIONS	79
6.1 Conclusion	79
6.2 Recommendations	81
CHAPTER 7: REFERENCES	82
Appendix 1: Relevant Tables and Schematic Diagrams	92
Appendix 2: Stream Flow Data	95
Appendix 3: Research Ethics and Regulatory Approval and Permit by MZUNIREC	102

LIST OF FIGURES

Figure 1: Sketch of midsection method of computing cross-section area for discharge measurements (WMO 2010).....	14
Figure 2: Schematic diagram of the steps involved in hydrological model formulation.....	18
Figure 3: Conceptual diagram and equations incorporated in the soil moisture model (Abera Abdi & Ayenew 2021).....	23
Figure 4: Status of hydro-meteorological monitoring networks in developing countries (World Bank Group, 2018).....	29
Figure 5: South Rukuru and North Rumphi river basin showing hydro-metric stations and main river network	38
Figure 6: Hydro-metric network in the northern region of Malawi mainly in the South Rukuru River Basin (Source: Ministry of Irrigation and Water Development, 2014).	41
Figure 7: Schematic diagram for WEAP hydrologic model procedure.....	47
Figure 8: Flow chart of the study.....	52
Figure 9: Distribution of missing data in the time series stream flow data for the four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period.	55
Figure 10: Overall data availability for four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period.	57
Figure 11: Results of the overall data availability per year from 2000 to 2021 for hydrometric stations 7H3, 7G18, 7G14 and 7F1 respectively.....	58
Figure 12: Longest data availability without gaps for the four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period.	59

Figure 13: Bar plot of the continuity of available data for four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period. 62

Figure 14: Results of the continuity of the available data per year from 2000 to 2021 for hydrometric stations 7H3, 7G18, 7G14 and 7F1 respectively. 64

Figure 15: Monthly observed and simulated stream flows for hydrometric stations 7H3 and 7G18 for the entire 2000 – 2021 period. 66

Figure 16: Results of the WEAP simulations: Calculated terms of the mean monthly water balance for each sub-basin (mm). Precipitation, Runoff and Evapotranspiration respectively..... 67

Figure 17: Results of water balance monthly average for hydrometric stations 7H3 and 7G18 respectively. 69

Figure 18: Visual comparison of the observed and simulated streamflow using the coefficient of determination (R^2) during both the calibration and validation periods for hydrometric station 7G18 and 7H3..... 72

LIST OF TABLES

Table 1: Summary of physical condition of hydrological stations (GoM 2014)	31
Table 2: Data availability assessment criteria for the observed hydrological data	45
Table 3: Soil and land use-related parameters that are used to re-calibrate the WEAP hydrologic model.	46
Table 4: Methodology Matrix	51
Table 5: Proportions of missing data and hydrological data availability assessment	53
Table 6: Total data values recorded for entire period from 2000 – 2021.	56
Table 7: Continuous Data Length without Gaps (DL), Number of Occurrences (NOO), Total Available Data (TAD) and Percentage of the Available Data (% TAD) for each Continuous Data Length for stations 7H3, 7G18, 7G14 and 7F1 for the entire 2000 – 2021 period	61
Table 8: Values of these initial soil and land use – related parameters (default) and values for the estimated soil and land use-related parameters.	65
Table 9: Goodness of Fit Metrics (GOF) for Hydrometric Stations 7G18 and 7H3 respectively.	70
Table 10: Comparison of WEAP hydrologic model performance evaluation with other similar studies elsewhere.....	78

CHAPTER 1: INTRODUCTION

1.1. Background

There is a growing national demand for water in Malawi and there are concerns about its availability, particularly during the dry season, mainly due to climate change, economic and population growth. The demand for water in 2010 was already greater than the supply in many Water Resource Areas (WRAs) with the situation predicted to worsen in the future (GoM 2012a). It is expected that the national water demand will increase five-fold by 2035, from 3,048 mega-liters demanded per day in 2010, to 15,643 mega-liters demanded per day in 2035 (GoM 2012b). As there are growing and competing demands for water resources, it is essential to determine whether the available water within a river basin can meet the demand of the basin (Salehie et al. 2022).

Spatiotemporal information about the availability of water resources is very important for sustainable water resource management and utilization. To monitor Sustainable Development Goal 6 (SDG 6), indicator 6.4.2, which is the level of water stress in a particular region, requires the quantification of the available total freshwater water resources (FAO and UN Water 2021). To develop effective and efficient water allocation plans, it is essential to understand and quantify the total water resources available (Speed et al. 2013). In this era in which climate change exacerbates water demand and scarcity (WWAP 2020), it is essential to understand the dynamics of water availability and how to manage it sustainably (Guug et al. 2020). It is, therefore, necessary to understand how much water is available, where it is available and when it is available in order to manage water resources sustainably.

Water availability within a river basin refers to the total available water for human and ecological needs. The total available water in a catchment is typically the discharge (runoff) which can be estimated from observed streamflow data or through hydrological modelling (Speed et al. 2013). The main technique to obtain observed streamflow data is through stream gauging where the stage height and velocity are measured at a series of points in a cross-section of a stream or by constructing a flume or weir and recording stage height (Brooks et al. 2013). Although using observed streamflow data to assess the water availability is the conventional way, hydrological modelling is often used to improve the reliability of estimates of water availability as they generate streamflow estimates and trends over long periods (Speed et al. 2013). One of the hydrological models that can be used to estimate the current status and trends of water availability in an area over a specific period of time is the rainfall-runoff soil moisture method approach in Water Evaluation and Planning (WEAP) Modelling Software (Yaykiran et al. 2019). This method takes into account a one-dimensional, 2-layer (“bucket”) soil moisture dynamic accounting system that uses empirical functions to partition water into evapotranspiration, surface runoff, sub-surface runoff (i.e., interflow), and deep percolation for a sub-catchment unit at the root zone (SEI 2022a). However, these hydrological models need ample reliable, accurate and up-to-date hydrometric information (Armanios & Fisher 2014) on all components that influence the rainfall-runoff relationship. However globally, there is a decline in the network density of operational hydro-meteorological field stations which is an obstacle to the effective development of these models (Karimi & Bastiaanssen 2014). In many regions, hydro-meteorological monitoring networks are often sparse and have large latency making it impractical to make real-time decisions on water management. In developing regions where the need for hydro-meteorological data is arguably

greatest, the hydrological monitoring networks are in a poor state and there is a lack of human capacity to monitor and collect data (Sheffield et al. 2018).

For instance, in Malawi, the water monitoring network was historically very extensive, however since the 1980s many stations have been closed due to a decline in the resources available to run the system (GoM 2011). More than 300 hydrological stations historically existed in Malawi. But now most of these hydrological monitoring stations do not have a continuous record of data and some of the stations have been abandoned or closed. Available data of daily discharge and water levels has a lot of gaps in the hydrometric databases of HYDATA and HYDSTRA. These hydrological data sets are fragmented and inconsistent. This decline in the network density of hydrological data is mainly due to low priority and limited financial resources for hydrological services. Accessibility to hydrological stations to collect data during high flows/rainy season is also another major issue. This makes it difficult to maintain and monitor hydrological stations as well as to collect and store the data. (GoM 2014).

Limited availability of hydro-meteorological data in the vast majority of river basins world-wide increases the value of alternative data sources such as remotely acquired data in hydrological modelling (Hulsman et al. 2020). Remotely acquired hydro-meteorological data are an attractive complementary source of information to in situ monitoring networks. Satellite-based sensors and earth models are now capable of making direct and indirect measurements of nearly all components of the hydrological cycle. The components include rainfall, evapotranspiration, lake and river levels, surface water, soil moisture, snow, and total water storage (Sheffield et al. 2018). With advances in information technology, several tools are being developed to assist in hydro-meteorological data acquisition and management.

For instance, the recently released Global Land Data Assimilation System (GLDAS) 2.0 version generates a variety of hydro-meteorological variables, including the following: rainfall, air temperature, wind, specific humidity, soil moisture and many more; provided at a spatial resolution of 0.25° (from 2000 to present) using the Noah model and on temporal resolution of 3 hourly and monthly (Ji et al. 2015). Making it possible to develop a rainfall-runoff model (soil moisture method) and simulate the water balance components over time.

With these recent remote sensing products now online, it is possible to assess water availability within river basins through a rainfall-runoff (soil moisture method) model. Several hydrological studies have been conducted worldwide to evaluate the use of remote sensing products for assessing water availability (Muthuwatta et al. 2010; Armanios & Fisher 2014; Oliveira et al. 2014; Moreira et al. 2019; FAO & IHE Delft 2020; Guug et al. 2020; Salehie et al. 2022). However most of these studies have been carried out in other continents other than Africa (Muthuwatta et al. 2010; Oliveira et al. 2014; Moreira et al. 2019; Guug et al. 2020; Salehie et al. 2022). In Africa particularly Malawi, only a few studies have been conducted (Stisen et al. 2008; Stisen & Sandholt 2010; Armanios & Fisher 2014; FAO & IHE Delft 2020). Most of these research studies were conducted in large river basins and used remote sensing datasets that had low spatial and temporal resolutions.

Furthermore, most of these studies developed simple conceptual water balance models to assess water availability trends. Only a few research studies (Jayantari et al. 2019) have used integrated hydrological and water resource models such as the WEAP modelling software to assess water availability trends with remotely acquired hydro-meteorological data. GLDAS datasets are widely being used and several studies have been conducted to evaluate their applicability (Ji et al. 2015; Bi et al. 2016; Zhang et al. 2018; Khasmakhi et al. 2020). However, most of these studies have

been carried out in other continents other than Africa, as such, there is a need for research studies to be conducted in Africa particularly in Malawi to evaluate the applicability of GLDAS datasets, considering the dense decline of hydro-meteorological data where GLDAS datasets can be one of the alternative sources. Moreover, few of these studies evaluated GLDAS climatic data to develop a rainfall runoff model in WEAP modelling software. Against this context, the purpose of this study was to evaluate the potential use of Global Land Data Assimilation System (GLDAS) datasets of precipitation, air temperature and wind speed in the WEAP modelling software, to develop a rainfall runoff model (soil moisture method) that will be used to assess water availability within South Rukuru and North Rumphu river basin.

1.2. Statement of the Problem

As there are growing and competing demands for water resources, it is essential to determine whether the available water within a river basin can meet the demand of the basin (Salehie et al. 2022). However, with the decline of hydro-meteorological data availability, it is difficult to determine how much water is available. Hydro-meteorological data from remote sensing products can be an alternative source. Despite this, many studies have been conducted on the potential use of remote sensing to assess water availability worldwide. Few researchers have evaluated the potential use of Global Land Data Assimilation System (GLDAS) Version 2.0 datasets in WEAP modelling software to develop a rainfall runoff model that will assess water availability. Hence, this research study was set to evaluate the potential of using GLDAS remotely acquired hydro-meteorological data to develop a rainfall runoff model in WEAP modelling software that was used to assess water availability in South Rukuru and North Rumphu Basin.

1.3 Aim of Study

1.3.1 Main Objective

The research study aimed at assessing water availability using remote sensing in South Rukuru and North Rumphu river basin.

1.3.2 Specific Objectives

The specific objectives of the research study were to:

- a) To assess the availability of observed hydrological data within the basin.
- b) To develop a remote sensing rainfall runoff model.
- c) To conduct a performance analysis of the remote sensing rainfall runoff model.

1.4 Research Questions

- a) How complete is the observed hydrological data?
- b) Can the remote sensing rainfall runoff model simulate streamflow?
- c) Is the remote sensing rainfall runoff model good for use?

1.5 Significance of the Study

Systematically monitoring water availability (total available water) within a river basin through the use of remote sensing techniques makes it practical to devise real-time decisions of proper water allocation. The national water regulatory authorities can use this real-time data to be properly guided in monitoring the water resources, setting water allocation limits, water licensing policies and other regulatory services. The freshwater availability information helps monitor water stress levels in a particular region over a specific period of time there by tracking the sustainable development goal on water and sanitation (SDG 6) indicator 6.4.2.

Therefore, apart from water allocation planning, the policy and decision makers will be able to target interventions at regions with high water stress and sectors with high water abstraction and also ecosystem health consideration. As water availability is being monitored, hydro-meteorological information is also being monitored, this is very beneficial to the Ministry of Water and Sanitation, since many water resources management decisions require hydro-meteorological information. The research will provide skills and knowledge to the academic world on how to assess water availability with remote sensing tools for sustainable water resources management

1.6 Ethical Consideration

Ethics is an integral part of the research and its importance varies depending on the methods of inquiry. This research study sought ethical clearance from the Mzuzu University Research Ethics Committee (MZUNIREC). The MZUNIREC approved fieldwork with Protocol Reference number MZUNIREC/DOR/22/94. The study permission was also sought from the Rumphi District Water Office.

All material references have been acknowledged. This research deals directly with data sets and use of data in a predictive and confidential manner. The research makes use of legally obtained datasets and tools. The data comes from openly available data sources that the public can access. The research does not include any surveys, interviews or direct respondents where confidentiality and anonymity needs to be ensured. The research used recognized data sources, software tools that have previously been used to ensure the quality and integrity of the research outputs.

1.7 Research Thesis Outline

The thesis is organized into 6 Chapters. Chapter 1 introduces the research topic and covers the background, problem statement, aim of the study, significance of the study and ethical

consideration. Chapter 2 is a comprehensive literature review on water resources demand, the need for assessing water availability, how to assess water availability, the hydro-meteorological data situation, remote sensing products and previous studies. The chapter also identifies gaps in the literature and reviews of methods. Chapter 3 describes the study area and research design but it also includes the data collection and analysis methods and tools that were used in the study. In addition, the methodology matrix and study flow chart are also incorporated in this chapter. Chapter 4 focuses on the results of the observed hydrological data availability assessment, the remote sensing rainfall WEAP model and the WEAP model performance evaluation. Chapter 5 is a discussion of the results that were obtained for the observed hydrological data availability assessment, the remote sensing rainfall WEAP model and the WEAP model performance evaluation. Lastly, the final chapter provides the summary of the research thesis which includes the conclusion and the recommendations.

CHAPTER 2: LITERATURE REVIEW

2.1 Background

Water is an essential resource, which is increasingly becoming scarce due to rising demands. Growing economies and populations require better water resources management to cope with the rising demand for energy and food and to ensure access to safe water and adequate sanitation (García et al. 2016). While there is a broad consensus about the benefits of sustainable water resources management, putting that knowledge into practice is usually easier said than done (WWAP 2018). To be able to make good water decisions, water resource managers need systematic ways to assess and monitor changes in water availability. Accurately accounting is needed for the current water resources – where, when, and how much – water is available, as well as an illustration of the potential changes caused by seasonal, natural, and climate-induced variability, from rainfall and runoff to evaporation and transpiration (Speed et al. 2013). In light of the harsh realities of climate change, this information is needed in larger quantities over broader areas and longer time periods than ever before (García et al. 2016). This chapter will outline the water resources demand, the need for assessing water availability, how to assess water availability, the hydro-meteorological data situation, remote sensing products and previous studies. The chapter will also identify gaps in the literature and review the methods.

2.2 Water Resources Demand

Many regions around the world are expected to face either absolute or seasonal water scarcity conditions, driven by increasing competition for water between agriculture and other sectors and more variable water availability because of climate change (Greve et al. 2018). Population growth, economic development, changing consumption patterns, intensified agricultural production and

expanding cities have generated a substantial rise in water demand (Wada & Bierkens 2014). The global water demand has increased by nearly 6 times during the last 100 years and continues to grow (Wada et al. 2016). Presently the global water demand for all uses has been estimated at 4,600 km³ per year. It will increase by 20% to 30%, up to 5,500 to 6,000 km³ per year by 2050 (WWAP 2018).

Agriculture presently accounts for 70% of global water withdrawals, the vast of which is used for irrigation. The agricultural global water withdrawal will increase over time due to its intensification to match the food demand (WWAP 2018). Water utilization by the industry sector accounts for about 20% of global water withdrawals, which is dominated by energy production (approximately 75%) with the remaining 25% of industrial water withdrawals being used for manufacturing (WWAP 2014). Domestic global water use currently accounts for 10% of global water withdrawals. It is expected to increase significantly over the period 2010 to 2050 in all the world regions with the exception of Western Europe where the demand remains constant. The greatest increase in domestic demand is expected to occur in African and Asian sub-regions where it could more than triple, and it could more than double in Central and South America (WWAP 2018).

Currently, Malawi's river basins have been facing challenges in satisfying ever-growing demands for water; domestic use, irrigation, and the requirement for environmental flow to maintain ecosystem health. Consequently, competition has steadily tightened for water abstraction among many water users. Besides, the sustainability of water resources has become more challenging under potential impacts of human-induced changes such as catchment degradation and climate change (GoM 2010). Population is growing in the country, from about 13 million people in 2008 to about 17.5 million in 2018 (GoM 2020), as a result, the overall demand for water resources will

also increase and the pressure will intensify. It is expected that the national water demand is to increase five-fold by 2035, from 3,048 mega-liters demanded per day in 2010, to 15,643 mega-liters demanded per day in 2035 (GoM 2012a). This represents an average annual growth of 17 percent over 25 years. The overall water demand for human use (water supply and sanitation) is expected to triple over this period, from 717 mega-liters per day in 2010 to 2,154 mega-liters per day in 2035. Demand for water for irrigation is expected to, on average, increase by 19 percent, from daily demand of 2,632 mega-liters in 2010 to 15,364 mega-liters in 2035. With the challenged sustainability of the water resources in the dry season, water shortages occur, for instance in South Rukuru and North Rumphi river basin major rivers dry up and the basin is expected to have a shortfall of 6 mega-liters per day (GoM 2012b). It is clear that the water demand is rising country wide. In the face of this increasing water demand, it is vital to understand the location and distribution of water resources throughout the country; it is necessary to know if there is enough water available to meet demand for all uses; and if not, there is a need to identify the types of investments necessary to meet demand for sustainable water resources management.

2.3 The Need for Water Availability Assessments

Understanding how much water is available, where it is available and when it is available within a river basin (Speed et al. 2013) is essential to determine if the demand for water can be met (Salehie et al. 2022). In this era of climate change which exacerbates water demand and scarcity (WWAP 2020), it is essential to understand the dynamics of water availability and how to manage it sustainably (Guug et al. 2020). As a way to respond to the rapid growth in demand for water, developing effective and efficient water allocation plans is necessary and it is essential to understand and quantify the total water available (Speed et al. 2013). To improve the quantitative state of water resources due to existing or future water demand conditions, water availability

assessments are necessary (EC 2015). To monitor water stress which occurs when the water demand exceeds the available amount in an area over a specific period of time, it is essential to quantify the available total freshwater water resources (FAO and UN Water 2021).

To assess water availability in a particular region, there is a need for ample hydro-meteorological information (Armanios & Fisher 2014), consequently, as water availability is being assessed, hydro-meteorological information is also being monitored. A range of water resources management decisions benefit from hydro-meteorological information, that is; environmental management and protection; the design and operation of water infrastructure for hydropower, flood control, drought preparedness and mitigation, irrigation, water supply, and wastewater treatment; hydrological disaster risk management and reduction; and planning policy, and governance (García et al. 2016). In this context, effective water resources management depends on a thorough understanding of the quantity and quality of available water in the sub-basin (Lorenz & Ziegeweid 2016).

2.4 Assessing Water Availability

The current status and trends of water availability within a river basin over a specific period may be assessed from observed streamflow data or through hydrological modelling (Speed et al. 2013). With regard to this research, water availability is referred to as a total quantity of water resources within a river basin.

Stream gauging is the main technique used to obtain reliable and accurate observed streamflow data. The primary objective of the streamflow gauging network is to develop a historic record of water availability within the catchment (GoM 2011). The streamflow data referred to is primarily continuous records of discharge (runoff) at stream-gauging stations; a gauging station being a

stream site instrumented and operated to provide continuous records of discharge (WMO 2010). The general location of a gauging station has to be determined with the consideration that the location should be stable, have a sufficient depth for obtaining velocity measurements at the lowest of stream flows, and be located in a straight reach without turbulent flow in order to develop a “discharge rating curve”; which is the relation of water level at a given point in a stream to a corresponding volumetric rate of flow (Brooks et al. 2013).

There are many instruments available for observing, sensing, recording and transmitting stream stage data. Such instrumentation ranges from the simple non recording auxiliary gauges to sophisticated water level sensors (WMO 2010). However, the basic instrument most commonly used in making the measurement is the current meter which measures the stream velocity.

The stream channel cross section is divided into vertical subsections. Area for each subsection is calculated by measuring the average width and depth. The velocity of flow at a point is proportional to the rate of rotation of the rotor during a fixed period of time (Dobriyal et al. 2017). The total discharge for a current meter measurement is the summation of the products of the partial areas of the stream cross section and their respective average velocities. This computation is expressed by the equation below:

$$Q = \sum_i^n q_i = \sum_i^n v_i a_i = \sum_i^n v_i (w_i d_i) \quad (1)$$

where Q = Total Stream Discharge (m^3/s),
 q_i = Discharge in subsection i (m^3/s),
 v_i = Mean flow velocity normal to subsection i (m/s),
 a_i = Cross-sectional area of subsection i (m^2)
 w_i = Width of subsection i (m)
 d_i = Depth of subsection i (m)
 n = Number of subsections dividing the stream cross-section

Figure 1 showing the midsection method of computing cross-section area for discharge measurements.

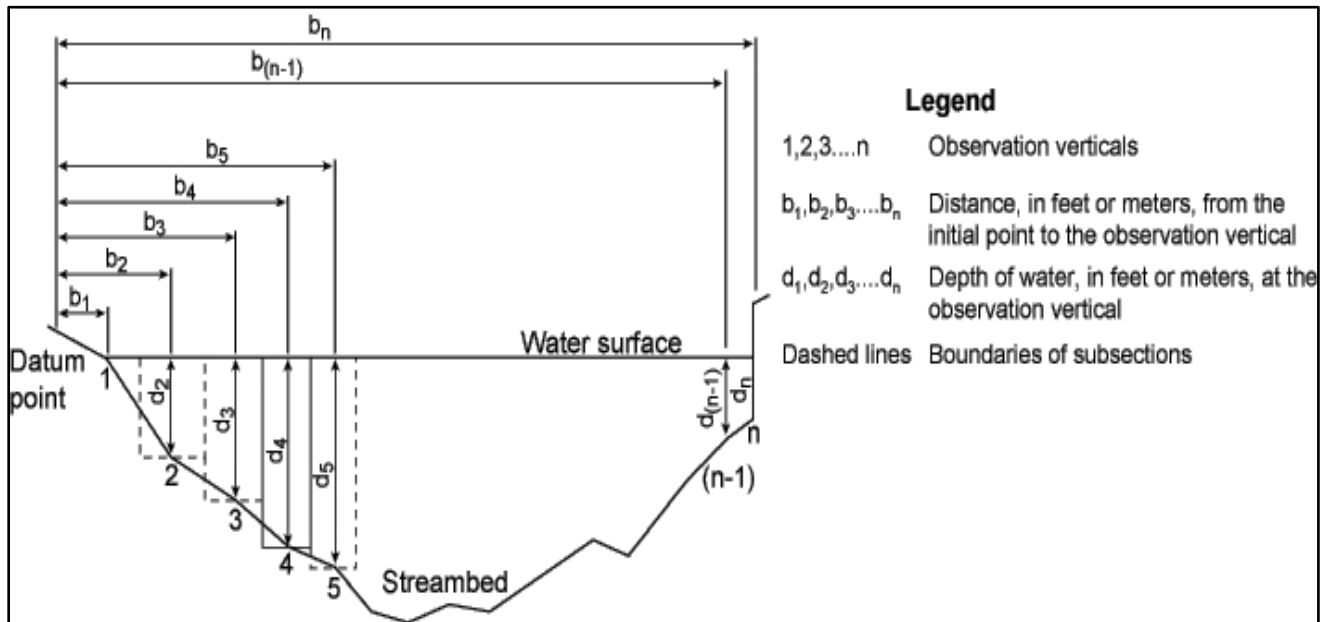


Figure 1: Sketch of midsection method of computing cross-section area for discharge measurements (WMO 2010)

In Malawi, hydrological observation of water level and discharge (streamflow) is managed by the Surface Water Division of the Water Department under the Ministry of Water and Sanitation. There are 12 hydrometric districts in the country. Each hydrometric district is managed by the corresponding responsible district water office. For instance, the South Rukuru and North Rumphi river basin falls into two hydrometric districts; Rumphi and Mzimba. It is also under the responsibility of the Rumphi and Mzuzu water offices. It is the responsibility of district staff to oversee the operations and maintenance of the gauging stations (GoM 2011).

Discharge measurement has been and is being carried out by staff of the district water offices. Gauge readers are assigned to operational stations to perform daily observation and data recording.

The frequency of water level observation is twice a day (basically 8:00 AM and 4:00 PM) and data recording is in triplicate hardcopy format. The district officer collects the records from gauge readers basically on a monthly basis. However, this is sometimes restrained by logistical problems of unavailability of staff, vehicles or fuel. A copy of the record is filed at the District Office, a second copy is sent to the Regional Office, and the third copy is sent to the Ministry's head office in Lilongwe. It takes about two to three months for the data to reach the head office. The collected recorded sheets are then input into the hydrological database of HYDSTRA in the head office. Simple maintenance work like clearance of the site is done by gauge readers, but in cases of serious rehabilitation/replacement, the maintenance work is done by the district staff (GoM 2011).

However, remuneration to gauge readers is not enough to ensure their commitment to the extent that many gauges are now not even being read. In addition, maintenance of the hydrological stations is also constrained by unavailability of staff, vehicles or fuel. Some of the hydrological stations have been vandalized or damaged by floods. The process of collecting and storing data is time consuming, making it difficult to make real time decisions (GoM 2014). This shows that the hydrological data sets are fragmented and inconsistent.

Although using observed streamflow data to assess the water availability is the conventional way, it can be costly, time-consuming and frequently dangerous (Costa et al. 2006), and therefore hydrological modelling is often used to improve the reliability of estimates of water availability as they generate streamflow estimates and trends over long periods (Speed et al. 2013). While hydrological models can generate historical, current or natural streamflow records, they can also be used to generate future yields or scenario modelling. They also provide a chance to estimate some variables that are difficult to measure in the field, for instance the evapotranspiration and total water storage changes (Abera Abdi & Ayenew 2021).

Hydrological modeling can be referred to as the representation of real hydrologic features and systems with models, mathematical analogues, and computer simulations (Anees et al. 2016). A model can be defined as a simplified representation of a phenomenon or a process. Hydrological models are simplified mathematical representations of a real-world system that is used to enhance the understanding and prediction of hydrologic processes (Solomatine & Wagener 2011). In the recent decade, hydrologic models have played a very crucial role in water resources management decisions such as depicting the input-output properties in an existing basin, estimating runoff for ungauged catchments, predicting impacts of changes in the basin, coupling the hydrology with geochemistry or coupling the hydrology and meteorology (García et al. 2016).

Developing a hydrological model requires the application of proper procedures of model identification, calibration, validation, uncertainty assessment and performance evaluation (Kumambala 2010). In developing a hydrological model, the first step in the process is model-building (how does a model come about), then the modeling protocol (a procedure used to make the model a success). The model-building process requires four stages. The first stage is the perceptual model, where the understanding of the system to be modeled in the modeler's head. This perceptual model forms the basis of the conceptual model. This conceptual model is a formalization of the perceptual model through the definition of system boundaries, inputs - states - outputs, connections of system components. Once a conceptual model has been derived, it is then translated into a mathematical model in the form of input - output equations. Finally, the mathematical model has to be implemented as computer code so that the equation can be solved in a computational model (Solomatine & Wagener 2011).

At this stage, a modeling protocol is used to apply the model. Modelling protocol focuses on understanding the behaviour and performance of the model. It is the verification of the model

appropriateness and performance (Solomatine & Wagener 2011). A modeling protocol, at its simplest level involves calibration process followed by the validation process, sensitivity and uncertainty analysis, lastly the performance evaluation is conducted (Donigian 2002). This is the crucial phase where the model is assessed whether it can reasonably represent the watershed behaviour (checking how well the model represents the watershed). Below is figure 2 showing the schematic diagram of the steps involved in hydrological model formulation;

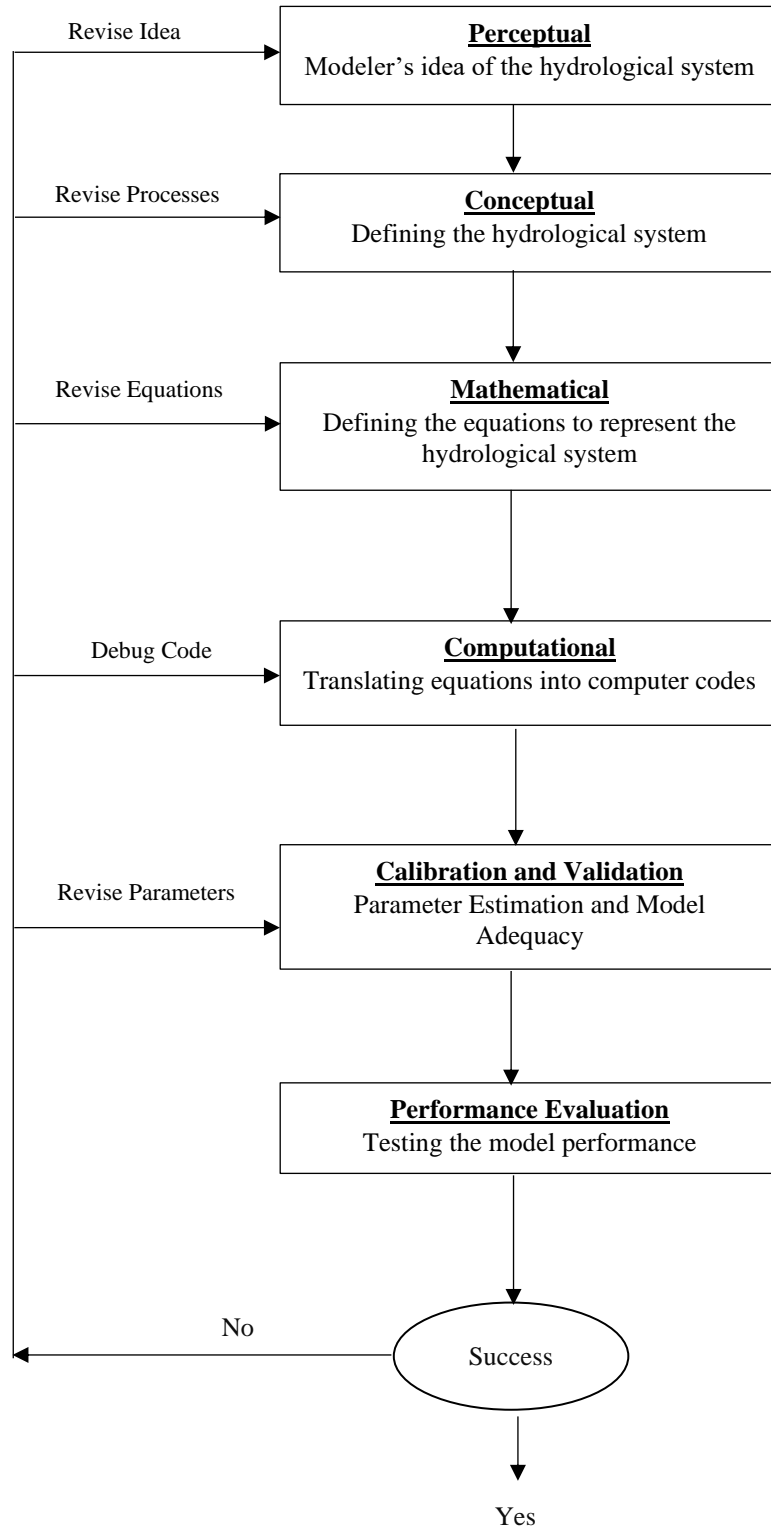


Figure 2: Schematic diagram of the steps involved in hydrological model formulation (Kumambala 2010).

After the computational model has been derived, the modelling protocol is applied. The first step in modelling protocol is to estimate the parameters of the model. The parameters of the watershed model, in general, cannot be determined directly from physical characteristics of the catchments, and hence the parameters values must be estimated using observed data. However, with fragmented and inconsistent hydrological data it can be difficult to estimate parameters. If this is the case, remotely acquired hydrological data can be a complementary source of information. The process of parameter estimation is known as model calibration. The main purpose of the model calibration is to obtain a parameter set for a catchment which gives the best possible fit between the simulated and observed flows (Jazim 2006).

The broad approaches to watershed model calibration are; a) trial and error parameter optimization which is based on the selection of model parameters by trial and checking the goodness of fit of the estimated output (Kumambala 2010) and b) automatic parameter optimization where model parameters are estimated automatically with the goal seeking algorithms (Madsen 2000; Madsen, Wilson & Ammentorp 2002; Madsen 2003).

It would be an easy task if the calibrated parameters obtained through the application of optimization techniques can be generalized as optimal parameters. But, in watershed hydrological modeling, another procedure is required to verify the acceptability of this calibration. This procedure is called model validation. In the model validation process, the parameters resulting from the calibration of a certain hydrological dataset should be inserted in the model as input to another set of hydrological data, then the resulting output should be reasonably close to the observed measurements for the model to be a success. The model is considered as valid only when the performance is satisfactory in both calibration and validation (Jazim 2006). The tendency has

been to use most of the available data in calibration and to confine the validation period to a couple of years (Madsen et al. 2002; Madsen 2003).

It is desirable that a model should represent as closely as possible the physical process occurring within the catchment. It is also essential that a model should accurately represent the transformation of the input into output. To assess how the model is performing (to check if it is good or fit for use) different objective functions are used (Jazim 2006). Some of the commonly used objective functions are; the sum of squares of differences between the observed and estimated discharges, with the summation taken over the whole calibration period, as shown in the equation below;

$$F = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{2}$$

where F = An index of residual error,

y = Measured output,

\hat{y} = Model output $i = 1, 2, 3, \dots, n$

The quantity F is an index of residual error, which reflects the extent to which a model is successful in reproducing the observed discharges. The value of F should be close to zero for a good simulation. The second objective function that is well known is the relative error (RE) in volume between the observed and simulated outputs (Jazim 2006), which is defined by:

$$RE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{\sum_{i=1}^n y_i} \times 100 \tag{3}$$

where RE = Relative error in volume

y = Measured output,

\hat{y} = Model output $i = 1, 2, 3, \dots, n$

For a good match between the total observed and simulated volume, RE should be close to zero. The other objective function that is commonly used is the Nash-Sutcliffe Coefficient of Efficiency (Nash & Sutcliffe 1970) by defining the model ‘efficiency’ R^2 ;

$$R^2 = \frac{F_o - F}{F_o} \quad (4)$$

Where F_o is the initial variance given by the following equation

$$F_o = \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad (5)$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^n y_i \quad (6)$$

where \bar{y} is the mean in the calibration period and N is the number of data points.

Nowadays, various hydrological models have been developed across the world, each model having its own unique characteristics and developed for different purposes. Some of the models include; MODFLOW, SWAT (Soil and Water Assessment Tool), MIKE SHE, HSPF (Hydrological Simulation Program – Fortran) Model, HEC HMS, WEAP (Water Evaluation and Planning), AWBM (Australian Water Balance Model), HBV hydrology model, TOPMODEL. For this research study, the WEAP (Water Evaluation and Planning) model an integrated water resources model was used.

2.4.1 The Water Evaluation and Planning (WEAP) Model

WEAP is a software tool for integrated water resources planning, developed by the Stockholm Environment Institute (SEI). It provides a comprehensive, flexible and user-friendly framework

for water resources management. It is a water resources modeling system that includes options to simulate both the natural rainfall-runoff processes and the management of an implemented water system (Yates et al. 2009). A growing number of water specialists and environmentalists are finding WEAP to be an important addition to their toolbox of models, databases, spreadsheets, and other software (SEI 2005). The model is lumped, spatially continuous with areas configured as sub-catchments that cover the entire river basin (Ingol-Blanco & McKinney 2013). WEAP offers five different methods to simulate watershed hydrological processes (evapotranspiration, runoff, and infiltration) and water demand (environmental flow, irrigation, domestic uses etc.) (Yaykiran et al. 2019). These methods are;

- Rainfall Runoff
- Simplified Coefficient Approach
- Soil Moisture Method
- MABIA Method
- Plant Growth Method

In this research study, WEAP's Soil Moisture Method to estimate the rainfall-runoff processes at the sub-basin level was used. This method takes into account a one-dimensional, 2-layer ("bucket") soil moisture dynamic accounting system that uses empirical functions to partition water into evapotranspiration, surface runoff, sub-surface runoff (i.e., interflow), and deep percolation for a sub-catchment unit at the root zone (SEI 2022a). Figure 3 shows a conceptual diagram and equations incorporated into the soil moisture model.

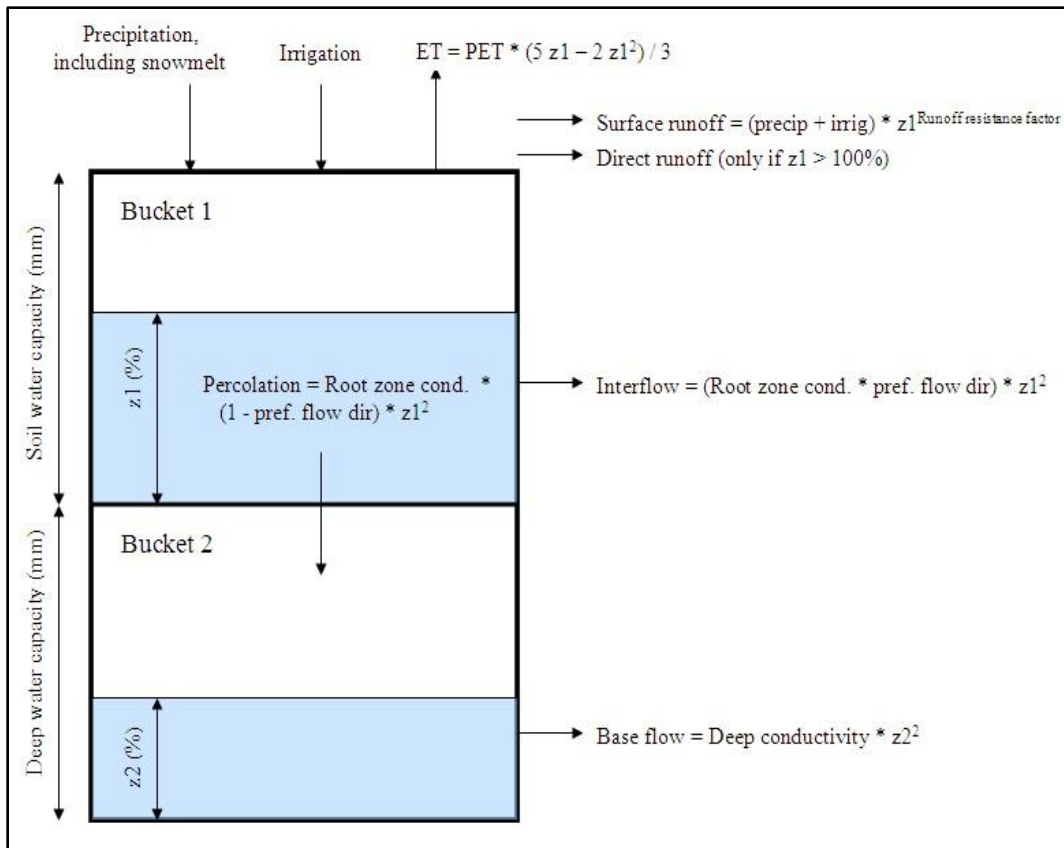


Figure 3: Conceptual diagram and equations incorporated in the soil moisture model (Abera Abdi & Ayenew 2021).

The model considers the movement of water through two vertical soil layers. The first layer represents water retained near the surface, which is available to plant roots; the second layer is deeper and water from this layer can be transmitted as base flow or groundwater recharge (Figure 3). For each sub catchment, the model computes the water balance due to inflows, outflows, and storage change in each layer. A basin is subdivided into a number of sub basins with different fractional land use or soil-type areas, the mathematical formulation to compute the storage change in the first layers is expressed in terms of a water balance as follows (Abera Abdi & Ayenew 2021):

$$Rd_j \frac{dz_{1,j}}{dt} = P_e(t) - PET(t)K_{c,j}(t) \left\{ \frac{5z_{1,j} - 2z_{1,j}^2}{3} \right\} - P_e(t)z_{1,j}^{RRF_j} - f_j k_{s,j} z_{1,j}^2 - (1 - f_j) k_{s,j} z_{1,j}^2 \quad (7)$$

where $Z_{1,j} \in [0, 1]$ is relative soil water storage, a fraction of the total effective water storage in the root zone layer land use j [dimensionless], j is land use and land cover unit (e.g., cultivated land, forest land, shrub) and each sub basin has N fraction of land use and land cover units, Rd_j is soil water holding capacity of land use j [mm], P_e is effective precipitation [mm], $ET_0(t)$ is reference evapotranspiration [mm/day], $K_{c,j}$ is crop coefficient land use j ; RRF_j is runoff resistance factor of land use and land cover j , $P_e(t)z_{1,j}^{RRF_j}$ is the surface runoff; $f_j k_{s,j} z_{1,j}^2$ is interflow from the first layer land use j , f_j is partitioning coefficient related to the land cover type, soil, and topography for the area j , that divides flow into horizontal f_j and vertical $(1-f_j)$ flows, and $K_{s,j}$ is saturated hydraulic conductivity of the root zone layer of land use j [mm/time].

Total surface runoff (RT) from each sub-catchment at time t is given as Eq. (8);

$$RT(t) = \sum_{j=1}^N A_j P_e(t) z_{1,j}^{RRF_j} \quad (8)$$

where A_j is area of the land use and land cover unit j .

The total interflow (IF), for each sub-catchment at time t , is given as Eq. (9);

$$IF(t) = \sum_{j=1}^N A_j f_j K_{s,j} z_{1,j}^2 \quad (9)$$

When an alluvial aquifer is introduced into the model and a runoff/infiltration link is established between the watershed unit and the groundwater node and groundwater recharge (R) (volume/time) to the aquifer is computed as follows:

$$R = \sum_{j=1}^N A_j(1 - f_j)K_{s,j}Z_{1,j}^2 \quad (10)$$

For applications where no return flow link is created from a catchment to a groundwater node base flow emanating from the second bucket is computed as;

$$S_{max} \frac{dz_2}{dt} = \left\{ \sum_{j=1}^N A_j(1 - f_j)K_{s,j}Z_{1,j}^2 \right\} - K_{s,2}Z_{2,j}^2 \quad (11)$$

where the inflow to this storage, S_{max} is the deep percolation from the upper storage given in Eq. (7) and $K_{s,2}$ is the saturated conductivity of the lower storage (mm/time), which is given as a single value for the sub basin, and therefore, does not include a subscript, j .

The reference evapotranspiration (ET_o) is estimated using the Penman-Monteith equation, modified for a standardized crop of grass, 0.12 m in height and with a surface resistance of 69 s/m.

The Penman-Monteith method to estimate ET_o is expressed as:

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (12)$$

where, ET_o is the reference evapotranspiration (mm/day), R_n is the net radiation at the crop surface (MJ/m² day), G is soil heat flux density (MJ/m² day), T is mean daily air temperature at 2 m height

(°C), u_2 is the wind speed at 2 m height (m/s), e_s is the saturation vapor pressure (kPa), e_a is the actual vapor pressure (kPa), $e_s - e_a$ is saturation vapor pressure deficit (kPa), Δ is slope vapor pressure curve (kPa/°C), and γ is the psychrometric constant (kPa/°C).

Actual evapotranspiration (ET) is also estimated using reference evapotranspiration (ET) (mm/day), crop coefficient (K_c), and soil water level in the modeling unit root zone given by Eq. (13).

$$ET = ET_o \times K_c (5Z_1 - Z_1^2) / 3 \quad (13)$$

A model for the South Rukuru and North Rumphi river basin has been built using the Water Evaluation and Planning (WEAP) modelling software.

2.4.2 WEAP Model Building Process

WEAP modelling software nowadays can automatically delineate catchments and rivers (using digital elevation data), calculate land area (disaggregated by elevation band and land cover), download historical climate data for each catchment (by elevation band) and create a climate summary background map layer (Jayantari et al. 2019). To develop a hydrological model in WEAP, the first step in the process is to delineate the catchment. Area boundaries should be set, a GIS schematic vector map of the river basin showing the catchment area and river network should be added with WGS84 projection, and then the automatic delineation mode should be activated. A panel will appear on the right side of the schematic, with various information and settings. The Digital Elevation Model data will be downloaded automatically. After DEM data has been downloaded, the catchment area is delineated automatically. Over the schematic, the following information can be seen; latitude and longitude, elevation, flow direction and flow accumulation.

In addition to delineating the catchment area, then this tool will also produce land cover branches since parameters for evapotranspiration, infiltration and runoff vary by land cover. The WEAP catchments are typically subdivided by land cover category. With the automatic delineation mode tool, Princeton historical climate dataset can be obtained for 1948 – 2010 at a 0.25-degree spatial resolution. Because climate often varies with elevation, the WEAP model subdivides the catchment into elevation bands, e.g., 0-500m, 500-1000m, 1000-1500m; so that climate data can be loaded separately for each elevation. Using the digital elevation data, WEAP automatically creates branches underneath each catchment according to the range of elevations in the catchment and calculates the area in each elevation band (SEI 2022b).

Basins and catchments can also be created. One basin can contain one or more catchments. A basin in the WEAP model is regarded as an area that drains to an outlet point. To create a basin, click the left mouse button anywhere within the area of the basin. The "Create New Basin" window will appear, and you can give the name for the Basin, initial Catchment and River that will be created. When a basin has been created, a catchment will also be created, initially occupying the entire basin. Then it can be subdivided into as many catchments within the basin. Each catchment has an associated river that runs through it and into which it drains. When a catchment is created, WEAP automatically creates a river object (blue line) based on the path through the catchment with the largest "flow accumulation". Lastly, the stream flow gauge element should be added, to enter the observed streamflow data so that it will be used to compare with the simulated streamflow (SEI 2022b).

2.4.3 Why WEAP Modelling Software

Numerous criteria can be adopted for choosing the “right” hydrological model for the project requirements and needs (Kumambala 2010). The WEAP model was chosen for the following reasons; Unlike other models, the WEAP model is an integrated hydrological and water resource model that is capable of scenario analyses of the water demand and supply balancing in a friendly approach giving a wide range of model results in a simplified manner. WEAP sets the demand and supply side of the equation in an organized way. Water use patterns, efficiencies, reuse, and allocation are placed on the demand side, while the supply side includes streamflow, groundwater, and water transfers (Tena et al. 2019).

The model is also a scalable tool, and it can be updated at any time. This allows for future improvement of the model results. With the WEAP Version 2019.2 software, WEAP can automatically download global data sets of elevation, land cover and climate as needed and then delineate catchments and rivers (using digital elevation data), calculate land area (disaggregated by elevation band and land cover) and create a climate summary background map layer. This simplifies the process of setting up and modelling catchment hydrology (Jayantari et al. 2019). Lastly, remotely acquired hydro-meteorological data can be easily loaded in WEAP. In summary, WEAP is practical for water resources planning and management. As a database, WEAP provides a system for maintaining water demand and supply information. As a forecasting tool, WEAP simulates water demand, supply, runoff, stream flows, storage, pollution generation, treatment and discharge, and instream water quality. As a policy analysis tool, WEAP evaluates a full range of water development and management options and takes into account multiple and competing uses of water systems (SEI 2005).

2.5 Limited and Declining Hydro-Meteorological Monitoring Networks

To assess water availability in a particular region, there is a need for ample hydro-meteorological information (Armanios & Fisher 2014). Ground-based (in situ) observation networks are fundamental but provide infrequent or sparse information over small areas and at a high cost. Particularly in developing countries, such hydro-meteorological networks have deteriorated over time, at present providing only limited hydro-meteorological information for (García et al. 2016) making sound, evidence-based water resource management decisions (Karimi & Bastiaanssen 2014). Water observation and hydro-metric monitoring networks are in decline and inadequately funded. Water resources monitoring can be costly, often requiring rather large instrumentation that can literally be ‘washed away’ during extreme flow events. Water observation networks provide incomplete and incompatible data on surface and groundwater quality and quantity (WWAP 2016). In many regions, hydro-meteorological monitoring networks are often sparse and have large latency and so are impractical for real-time decision-making. Figure 4, shows the status of hydro-meteorological monitoring networks in developing countries.

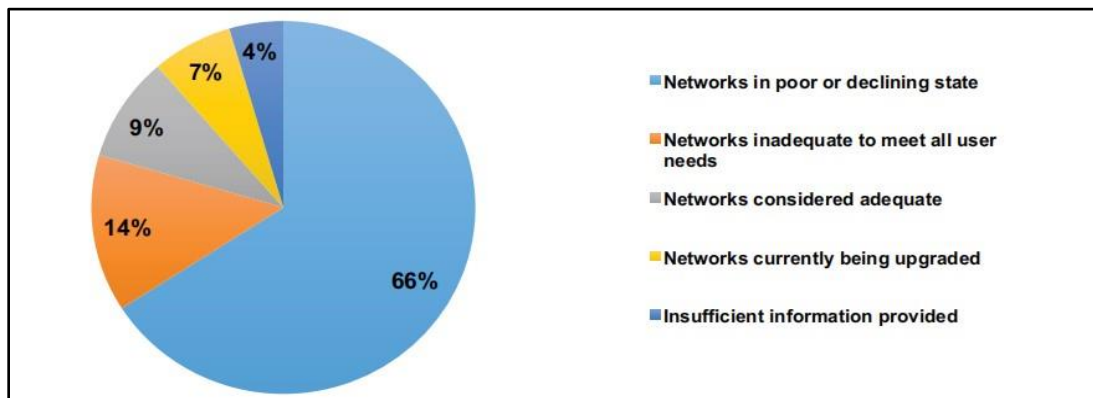


Figure 4: Status of hydro-meteorological monitoring networks in developing countries (World Bank Group, 2018).

In developing regions where the need for hydro-meteorological data is arguably greatest, the hydrological monitoring networks are in poor state and have the most deficient infrastructure for gathering data needed to monitor and predict them (Sheffield et al. 2018). Within the Southern Africa region, many governments have only a limited ability to collect the data needed for long-term water resources management, particularly good quality hydro-meteorological time series, such as rainfall and river flow. In addition, there are few skilled people and a high turnover of qualified staff and a lack of the funds, equipment and facilities for hydro-meteorological monitoring (Houghton-Carr & Fry 2006).

For instance, in Malawi, the major challenge is the decline in the network density of hydro-meteorological data. In the 1980s there were about 800 rainfall stations, but there are only between 100 and 200 operational rainfall stations at present. More than 300 hydrological stations historically existed in Malawi, and the number of operational stations historically has changed. According to GoM (2014), 136 stations are operational and 94 stations are closed, the rest have been abandoned. Among operational stations, stations with acceptable operational conditions are only about 40% (Table 1). There is a lot of data missing in the hydrometric databases of HYDATA and HYDSTRA for daily discharge and water levels. This decline in the network density of hydrological data is mainly due to low priority and limited financial resources for hydrological services. This makes it difficult to maintain and monitor hydrological stations as well as to collect and store the data. To be able to access the hydrological stations to collect data during high flows/rainy season is also another major issue in hydrological observation (GoM 2014). To this end, this research study assessed the availability of historical hydrological time series data from the hydrometric stations within the river basin.

Table 1: Summary of physical condition of hydrological stations (GoM 2014)

WRA		Open				Total Open	Closed	Total: Open & Closed
No.	Name	Good	Average	Poor	N.A			
1	Shire	17	6	1	1	25	23	48
2	Lake Chilwa	7	1	1	2	11	9	20
3	South West Lakeshore	5	0	0	2	7	3	10
4	Linthipe	12	2	4	0	18	2	20
5	Bua	6	0	2	0	8	3	11
6	Dwangwa	6	0	2	0	8	6	14
7	South Rukuru / North Rumphi	9	4	1	2	16	8	24
8	North Rukuru	0	1	0	2	3	1	4
9	Songwe / Lufira	1	1	5	2	9	7	16
10	South East Lakeshore	0	0	0	0	0	0	0
11	Lake Chiuta	1	0	0	0	1	1	2
12	Likoma Island	0	0	0	0	0	0	0
13	Chizumulu Island	0	0	0	1	1	0	1
14	Ruo	7	3	1	1	12	11	23
15	Nkhota-kota Lakeshore	0	0	4	1	5	1	6
16	Nkhata-bay Lakeshore	3	0	5	1	9	13	22
17	Karonga Lakeshore	0	1	1	1	3	6	9
Total		74	19	27	16	136	94	230

2.6 Remote Sensing Datasets

With limited and often declining hydro-meteorological monitoring networks, remotely sensed hydro-meteorological data are an attractive complementary source of information to in situ monitoring networks (Karimi & Bastiaanssen 2014). Satellite-based remote sensing methods are nowadays increasingly being used to capture the spatial variation in the hydro-meteorological and

catchment characteristics, resulting in significant improvement in the hydrologic modelling (Kumar & Reshmidevi 2013). Remote sensing enables coverage over large areas and spans of time without heavy field personnel requirements, and its accessibility, reliability, and accuracy have improved dramatically in recent years (García et al. 2016). Satellite-based sensors and earth models are now capable of making direct and indirect measurements of nearly all components of the hydrological cycle (Sheffield et al. 2018).

For instance, NASA Global Land Data Assimilation System version 2 (GLDAS-2.0) is an important tool for hydrology and water resources due to the provision of hydrological parameters on a global scale (Khasmakhi et al. 2020). It is a tool that merges the satellite- and ground-based observational data products, using advanced land surface modeling and data assimilation techniques in order to provide optimal simulations of global land surface states and fluxes in near-real time (Bi et al. 2016). It generates a variety of hydro-meteorological variables, including: rainfall, air temperature, wind, specific humidity and soil moisture (Ji et al. 2015), making it possible to develop a rainfall-runoff model (soil moisture method) and simulate the water balance components over time.

While these remote sensing products are being used in various research fields, the validity or “ground truth” of these products has to be taken into account when considering their potential use. Remote sensing estimations are prone to several sources of uncertainty, which can significantly affect the quality of the variables to be forecasted (García et al. 2016). It includes issues related to latency (how readily available are datasets in near real time); resolution both temporal and spatial; legacy (the length and continuity of long-term records), and how well remote sensing data represents observed data. It is, therefore, necessary that before using these remote sensing datasets, they have to be validated and calibrated with ground data (Sheffield et al. 2018).

2.7 Previous Studies

Several hydrological studies have been conducted worldwide to evaluate the use of remote sensing products for assessing water availability (Muthuwatta et al. 2010; Armanios & Fisher 2014; Oliveira et al. 2014; Moreira et al. 2019; FAO & IHE Delft 2020; Guug et al. 2020; Salehie et al. 2022). Moreira et al. (2019) investigated the potential of assessing the water availability using a water balance model in South America, at multiple scales, using satellite remote sensing precipitation and evapotranspiration datasets, terrestrial water storage changes from Gravity Recovery and Climate Experiment (GRACE), and discharge measurements. They concluded that there is indeed a great potential for assessing terrestrial water availability and understanding their spatial and temporal variability at multiple scales but the challenge remains of large uncertainties in remote sensing estimations.

Armanios & Fisher (2014) also conducted an exploratory study to assess the feasibility of using an entire satellite remote sensing based hydrologic budget model to measure water availability in the Rufiji basin, Tanzania. They concluded that the hydrologic budget model indicated a good performance and it was recommended that future researchers have to use newly up-to-date remote sensing tools that can provide near-real-time retrievals and also at good spatial and temporal resolution to a scale closer to that of interest of a water resource manager.

Oliveira et al. (2014) assessed the water balance of the Brazilian Cerrado based on remotely sensed estimates of precipitation (TRMM), evapotranspiration (MOD16), and terrestrial water storage (GRACE). The model demonstrated very good performance. They concluded that, although complete water budget closure from remote sensing remains a significant challenge due to

uncertainties in the data, it provides a useful way to evaluate the water availability trends and major water balance components.

But, most of these studies were carried out in the other continents other than Africa (Muthuwatta et al. 2010; Oliveira et al. 2014; Moreira et al. 2019; Guug et al. 2020; Salehie et al. 2022). In Africa particularly Malawi, only a few studies have been conducted (Stisen et al. 2008; Stisen & Sandholt 2010; Armanios & Fisher 2014; FAO & IHE Delft 2020). Most of these research studies were conducted in large river basins and used remote sensing datasets that had low spatial and temporal resolutions. Furthermore, most of these studies developed simple conceptual water balance models to assess water availability trends. Only a few research studies have used integrated hydrological and water resource models to assess water availability trends with remote sensing products. To this effect, the water evaluation and planning (WEAP) modelling software was selected to be used in the research study.

The WEAP modelling software has been used widely to develop hydrological models. Tena et al. (2019) conducted a hydrological modelling and water resources assessment of Chongwe River Basin, Zambia using WEAP Model. The performance of the WEAP model simulation was assessed statistically indicating a satisfactory model fit and result. Ingol-Blanco & McKinney (2013) also conducted a study in Rio Conchos Basin, Mexico of developing a hydrological model using WEAP, it was concluded that the model indicated a good performance. Abera Abdi & Ayenew (2021) conducted an evaluation study of the WEAP model in simulating sub basin hydrology in the Central Rift Valley basin, Ethiopia. The model has demonstrated a very good agreement between measured and simulated streamflow at the outlet of the sub basin. Bhave et al. (2020) conducted a climate change sensitivity analysis for Lake Malawi Basin with the WEAP model, and used the soil moisture method in WEAP to simulate the lake hydrology. It was

concluded that the model shows reasonable ability to capture the observed stream flows and lake levels. However, these studies did not use remotely sensed hydro-meteorological data.

But only a few researchers have used remotely sensed hydro-meteorological data in WEAP modelling software. Jayantari et al. (2019) used satellite data in the WEAP model to evaluate the water availability in Unda River Basin, Indonesia. It was concluded that the remotely sensed hydro-meteorological data can be used to facilitate the evaluation of water availability trends in a river basin when the availability of the in-situ hydro-meteorological data is insufficient. However, they did not use near-real-time retrievals to simulate real-time hydrological parameters that are necessary to make real-time water resources management decisions. The remotely acquired climate data was from the Princeton historical climate dataset of monthly temperature, precipitation and wind speed for 1948 – 2010, at a 0.25° spatial resolution. To this regard, the research study assessed the use of near real-time retrievals of GLDAS climate datasets.

Several studies have been conducted to investigate the applicability of GLDAS datasets. Ji et al. (2015) evaluated the GLDAS air temperature data products and the study demonstrated that the air temperature estimates are generally accurate, but caution should be taken when the data are used in mountainous areas or places with sparse weather stations. Bi et al. (2016) compared soil moisture in GLDAS model simulations and in situ observations over the Tibetan Plateau, China but the results showed poor agreement with the observed surface soil moisture. Khasmakhi et al. (2020) simulated river discharge in ungauged catchments by forcing GLDAS products into a hydrological model in Polroud Basin, Iran. The results clearly indicated that the forcing of GLDAS data into the hydrological model showed promising results with acceptable correlation with observed data. It was concluded that GLDAS datasets would be a good alternative in ungauged or poorly gauged watersheds. Zhang et al. (2018) conducted a study to evaluate the applicability of

GLDAS data and assess the climate and hydrological change in Yarlung Zangbo river basin, China. It was concluded that GLDAS data have a good applicability within the basin and to some extent it was recommended that GLDAS data can be a good supplement of climate and hydrology data to data limited areas.

However, most of these studies have been carried out in other continents other than Africa, therefore there is also a need for research studies to be conducted in Africa particularly in Malawi to evaluate the applicability of GLDAS datasets, since it is in this region where there is a dense decline of hydro-meteorological data, where GLDAS datasets can be one of the alternative sources. Furthermore, few researchers have seriously used and evaluated GLDAS climatic data to develop a rainfall runoff model in WEAP modelling software. To this context, the research study evaluated the potential of using Global Land Data Assimilation System (GLDAS) datasets of precipitation, air temperature and wind speed in the WEAP modelling software, to develop a rainfall runoff model (soil moisture method) that was used to assess water availability within South Rukuru and North Rumphu river basin.

CHAPTER 3: MATERIALS AND METHODS

3.1 Introduction

The chapter aims to describe the South Rukuru and North Rumphi River Basin in terms of its significance, hydrology, geology, land cover and land use. It also outlines the research design, data collection and analysis methods, the methodology matrix and the study flow chart.

3.2 Study Area

3.2.1 Location and General Description

This research study was conducted in the South Rukuru and North Rumphi River Basin which is located in the Rumphi and Mzimba districts of the Northern Region of Malawi. Malawi is divided into 17 water resource areas (WRAs) based on the river basins and further divided into water resource units (WRUs). There are 5 WRAs in the Northern Region. South Rukuru and North Rumphi River Basin is WRA 7 which is further divided into 8 WRUs (GoM 2014).

It is the largest water resource area that drains directly into Lake Malawi covering an area of 12,705 km². The major river in the basin is South Rukuru which originates from the Viphya Mountains and drains into Lake Malawi. The main tributaries are the Kasitu River, Runyina River, Rumphi River and Mzimba River. The North Rumphi River forms the other part of the river basin with a catchment area of 712 km². The North Rumphi River originates from the eastern side of the Nyika Plateau and drains into Lake Malawi (Kumambala 2010). Figure 5 shows the hydro-metric stations and main river network of South Rukuru and North Rumphi river basin.

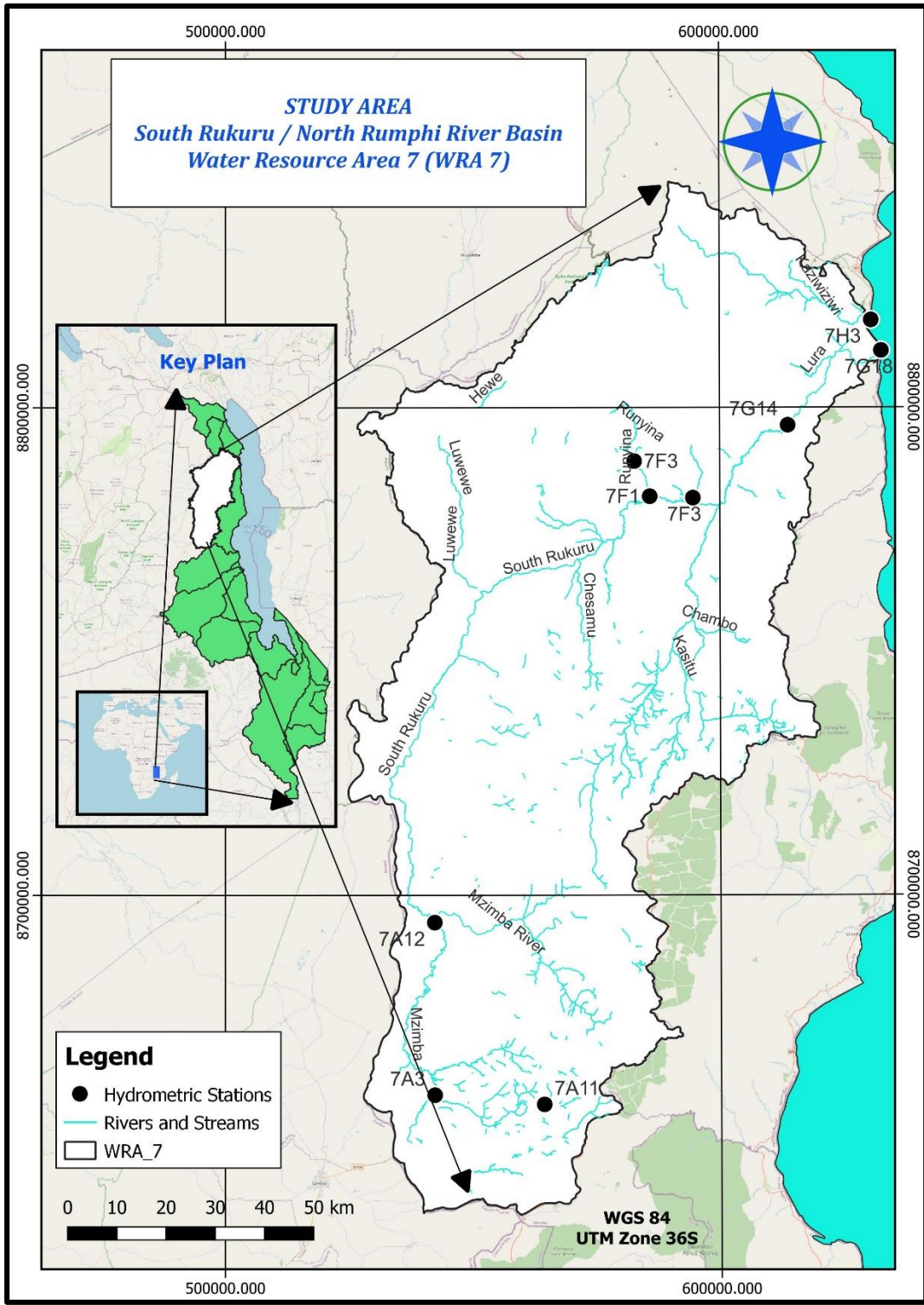


Figure 5: South Rukuru and North Rumpi river basin showing hydro-metric stations and main river network

3.2.2 Significance of the basin

South Rukuru and North Rumphu river basin has been chosen for the study for two main reasons. Firstly, there is still some consistent and reliable ground data compared to other catchments (Kumambala 2010). Secondly, the basin's water resources are critical to socio-economic activities such as crop and livestock production, fish farming and manufacturing industries. The basin also provides water for domestic use in rural households, peri-urban areas, and towns. The water resources within the basin also sustain the biodiversity and ecosystems of the Nyika National Park and Vwaza Marsh Wildlife Reserve.

3.2.3 Basin vegetation, geology, soil and relief

This river basin has much of its catchment area covered with woodlands. Open woodland of *Brachystegia* and *Julbernardia* types form the major part of the catchment's vegetation. Rolling grassland and scrub-lands generally cover most of the highland areas of Nyika and Viphya mountains (Kumambala 2010). The river basin is mainly made up of igneous and metamorphic rocks. The low plateau areas consist of schists, quartzites, gneisses and granulites while the Nyika Plateau and parts of Viphya mountains have granite. The Vwaza marshes have hydromorphic and ferralitic soils. The soils on high altitude areas of Viphya Mountains and Nyika plateau are mostly humic ferralitic soils. The ferralitic with ferrisols and lithosols soils are generally found on the low hill and slopes of the upper and south Rukuru valley (GoM 2005).

The basin has a varied relief which can be divided into a number of natural regions. The middle reach of the basin contains Vwaza marsh and Nkhamanga plain with scattered hills which lie between 1000 and 1100 m above sea level. The southern part of the basin has flat plains and with broad dambo and swamps which lies between 1070 and 1830 m above sea level. On the eastern

side of the basin lies the Viphya Mountains which is the source of the main tributaries such as Mzimba and Kasitu. The northern part of the basin is the Nyika Plateau (2000 – 2600 m amsl) where Runyina and Rumphu rivers originate. From the confluence with Rumphu river the South Rukuru river passes through the Rukuru - Kasitu valley. Beyond Kasitu valley the river passes through a meandering gorge over a series of rapids before discharging into Lake Malawi (Kumambala 2010).

3.2.4 Rainfall data monitoring within Basin

The basin had as many as 90 rainfall stations. But most of these stations now do not have a continuous record of data and some of the stations have been abandoned. Currently, there are only a few stations with continuous climate data. The basin receives low rainfall except for the highland areas of Viphya and Nyika Plateau. The upper and middle part of the basin receives 800 - 900 mm of rainfall annually. The Vwaza marsh receives lower rainfall in the range of 650 – 700 mm, while the highland areas receive annual rainfall of 1200 to 1600 mm. The only climate station available in the basin with a full complete record of climate data is Mzuzu station which is on the eastern side of the basin (Kumambala 2010).

3.2.5 Hydro-metric network in the Basin

Hydro-metric network in the basin started in 1953 with three gauging stations, one at Zombwe on Lunyangwa river, Edundu on the Kasitu river and Phoka court on the North Rumphu. For easy identification and effective data management, the gauging stations are provided with a unique numbering style. For instance, station **7G18**; the **7** stands for the Water Resource Area (WRA) in which the station is located, the **G** stands for the Water Resource Unit (WRU) in which it is located and **18** is the gauge/station number. In later years the network expanded progressively up to 27

stations. However, some of the stations have been closed and only a few stations have up-to-date records. Of the current stations 7G18 and 7G14 on South Rukuru river, 7H3 on North Rumpi, 7F1 on Runyina, and 7F2 on Rumpi River have up-to-date flow records. Station 7G14 at Phwezi is the oldest important station on South Rukuru River. In 1977 the station was upgraded by installing an autographic recorder. In 1985 another station was set up at Mlowe Bridge (7G18). Station 7G18 provides lake inflows from South Rukuru River (Kumambala, 2010). Figure 6 shows part of the hydro-metric network in the northern region of Malawi mainly in the South Rukuru River Basin.

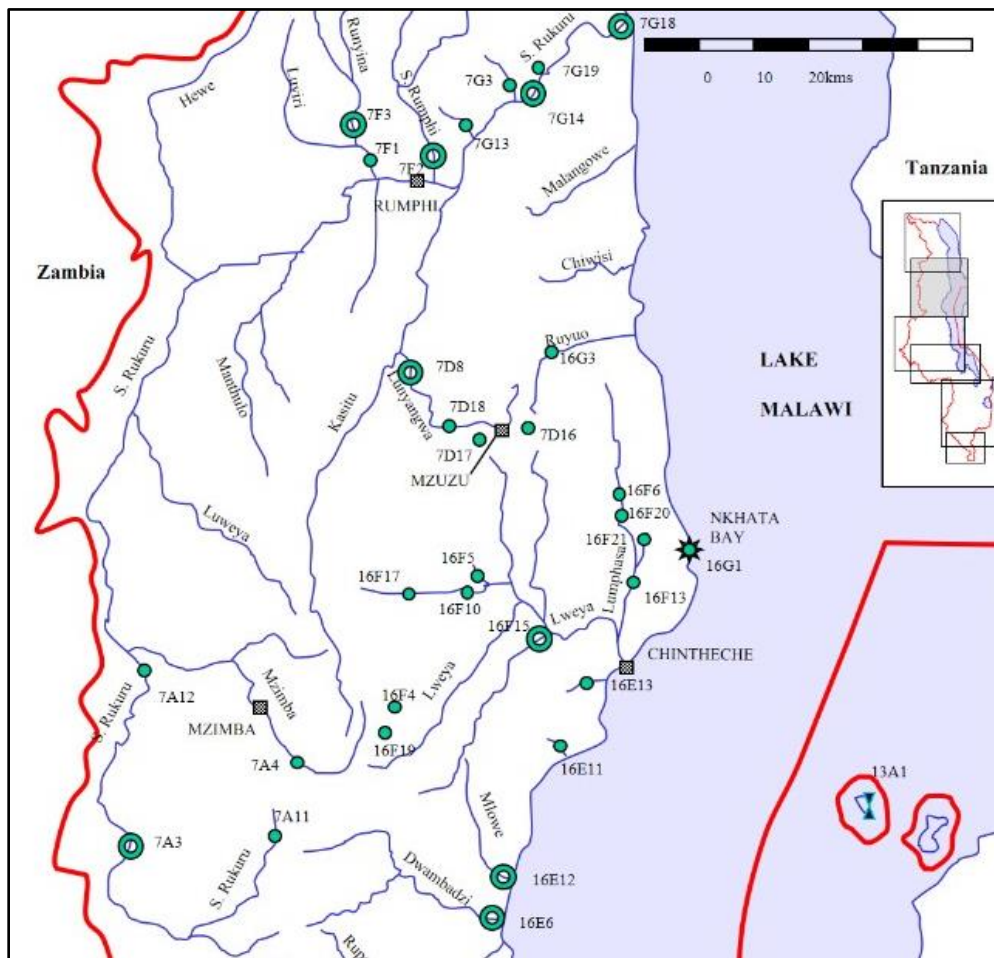


Figure 6: Hydro-metric network in the northern region of Malawi mainly in the South Rukuru River Basin (Source: Ministry of Irrigation and Water Development, 2014).

3.3 Research Design

For this research study, a quantitative research design was used. The whole research process dealt with data expressed in numerical values, graphs, and tables, then analyzed through mathematical and statistical analysis methods.

3.4 Data Collection

3.4.1 WEAP Model Input Data

To simulate the basin monthly hydrological processes with the WEAP hydrological model, the elevation data, climate data (monthly precipitation, temperature and wind speed) and land use land cover data are required. In addition, the observed streamflow data is also required which was used for model calibration and validation.

3.4.1.1 Remote Sensing Datasets

The remote sensing datasets that are in this study are commonly and widely used, and are freely available on a global scale. For the elevation data, it was derived from Hydro SHEDS digital elevation data based on high-resolution elevation data obtained from the Space Shuttle flight for NASA's Shuttle Radar Topography Mission (SRTM). The elevation data are available at two spatial resolutions: 500 meters (15 arc seconds) and 90 meters (3 arc seconds). For this study, the data at 90 meters (3 arc seconds) resolution was used.

For climatic data, Global Land Data Assimilation System (GLDAS) climate datasets of precipitation, temperature and wind speed on monthly basis at a spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$ was loaded and used in NetCDF format for a period 2000 to 2021 (22 years). The data was in NetCDF format for compatibility with the WEAP modelling software. This climatic data was accessed from the GIOVANNI website.

Land cover data from the European Space Agency (ESA) Climate Change Initiative (CCI) project was used. Surface Reflectance (SR) time series of Advanced Very High Resolution Radiometer (AVHRR) and PROBA-Vegetation satellite derives the land cover data from the period of 1992 to 2015 (24 years) which produces 22 Land Cover Classification System (LCCS) classes of land cover with a spatial resolution of 300 m (SEI 2022b). The elevation and land cover data was automatically downloaded in the WEAP modelling software.

3.4.1.2 Surface Observations

Stream flow data was obtained from Rumphi District Water Office for four hydrometric stations within the South Rukuru and North Rumphi river basin for a period of 22 years (2000 – 2021). The first 16 years (2000 – 2015) was used for calibration and the last 6 years (2016 – 2021) was used for validation.

Observed streamflow data contained gaps or missing values. Therefore, estimation of missing values was crucial to have a complete set of data in order to perform necessary statistical analysis. There are a number methods that have been adopted in infilling missing data ranging from simple interpolation to complex statistical methods. For this study, where streamflow data is missing, infilling of missing streamflow data was done using the spline interpolation methods in Microsoft Soft Excel.

3.5 Data Analysis

3.5.1 Observed Hydrological Data Availability Assessment

Data availability aims at assessing the available records which are based on the degree of completeness, which refers to the degree to which all data in a data set is available i.e. the wholeness or comprehensiveness of data (Emran 2015). According to Crochemore et al. (2020) data availability has to be characterized by the fraction (%) of the reference time period (2000 – 2021) covered by observed data.

For each time series, the overall availability was computed as the total fraction of 2000 – 2021 covered with data, as well as the longest availability was computed as the longest fraction of 2000 – 2021 covered with continuous data (i.e. without missing values). The longest availability was compared to the overall availability to assess the continuity of the time series. Table 2 below shows the data availability assessment criteria for the observed hydrological data which has been adapted from Crochemore et al., (2020). Scatter and bar plots were produced for visual inspection of the data availability.

Table 2: Data availability assessment criteria for the observed hydrological data adapted from Crochemore et al., (2020)

Assessment Criteria	Objective	Worst Case	Best Case
Overall availability for 2000 – 2021	Length of the observation series, as a fraction of the 2000 – 2021 period	No data from 2000 to 2021	All data from 2000 to 2021
Longest availability without gaps for 2000 – 2021	Length of the longest observation series without gaps, as a fraction of the 2000 – 2021 period	No continuous data from 2000 to 2021	Continuous data from 2000 to 2021
Continuity in available data	Length of the longest observation series without gaps, as a fraction of the overall availability, i.e. ratio between longest availability and overall availability	Only gaps in available data	No gaps in available data

3.5.2 The WEAP Model Calibration and Validation

Calibrating the model primarily involved obtaining a parameter set for each sub-catchment which gave the best possible fit between the simulated and observed flows (Jazim 2006). The WEAP hydrological model was calibrated to estimate land use and soil-related parameters using the manual (trial-and-error) method until a good fit is observed between the measured and simulated streamflow. The input parameter values were adjusted during the calibration to maximize output fit.

In general, the trial-and-error approach that was applied in this research involved a process wherein the model parameters were altered systematically and then the model was run several times until the solution matches observed values within an acceptable level of accuracy. This process was done for a period of sixteen years (2000 – 2015). The WEAP embedded soil moisture method involves seven soil and land use-related parameters (Ingol-Blanco & McKinney 2013; Abera Abdi & Ayenew 2021). Table 3 below shows the seven soil and land use-related parameters that are used to re-calibrate the hydrologic model.

Table 3: Soil and land use-related parameters that are used to re-calibrate the WEAP hydrologic model.

Parameters	Definition
K_c	Crop Coefficient
S_w	Root Zone Soil Water Capacity (mm)
D_w	Deep Soil Water Capacity (mm)
RRF	Runoff Resistance Factor
K_s	The Conductivity of Root Zone at Full Saturation (mm/month)
K_d	The Conductivity of Deep Zone at Full Saturation (mm/month)
F	Preferred Flow Direction
Z₁	Initial storage fraction at the beginning of simulation of upper soil layer
Z₂	initial storage fraction at the beginning of simulation of lower soil layer

A validation data set of six years (2016 – 2021) was used to assess the adequacy of the model. The adjusted calibration parameters of the hydrological model were now used in the validation data set and then the model was run. The hydrological model was regarded as a success when the resulted output at the validation stage was reasonably close to the observed measurements for the model.

Figure 7 below shows a schematic diagram of flow chart for input and output, and modeling processes using the WEAP hydrologic model.

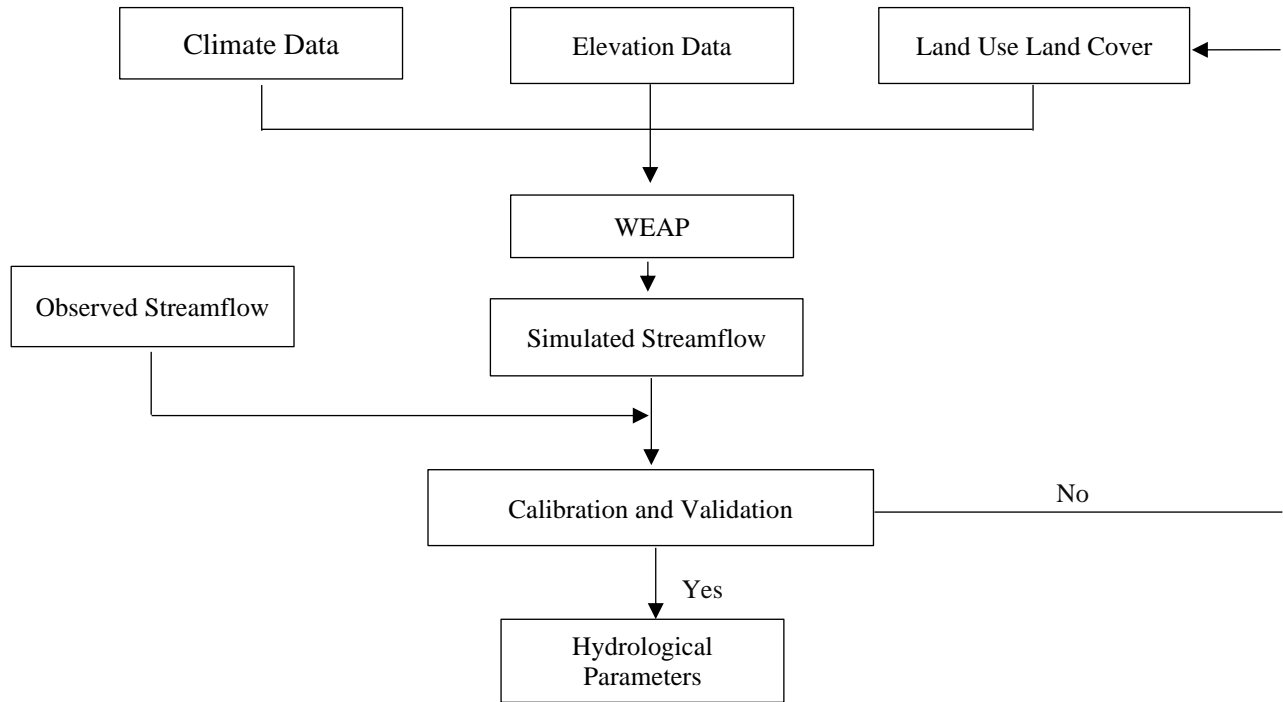


Figure 7: Schematic diagram for WEAP hydrologic model procedure

3.5.2 The WEAP Model Performance Evaluation Measures

The WEAP model performance was assessed using various techniques: joint plots of the monthly and mean monthly simulated and observed hydrographs and commonly used statistical methods. These methods include: The coefficient of determination (R^2) (Krause, Boyle & Bäse 2005); Nash-Sutcliffe coefficient of efficiency (NSE) (Nash & Sutcliffe 1970), index of agreement (IA) (Willmott 1981), Percent Bias (PBIAS) (Gupta, Sorooshian & Yapo 1999), Kling and Gupta Efficiency (KGE) (Gupta et al. 2009).

The coefficient of determination (R^2) represents how well a linear model fits the data. It varies in the range $0 \leq R^2 \leq 1$; a value of 1 means a perfect linear fit, and a value of 0 indicates that the linear model is not representative. According to Santhi et al. (2001); Van Liew et al. (2007), the values of R^2 that are greater than 0.5 are acceptable with higher values indicating less error variance. R^2 is computed as follows;

$$R^2 = \frac{(\sum [X_i - \bar{X}] [Y_i - \bar{Y}])^2}{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2} \quad (14)$$

where $X_i, Y_i; \bar{X}, \bar{Y}$; and n denote the i th measured monthly discharge data, the i th simulated monthly discharge data, the mean of the measured monthly discharge data, the mean of the simulated monthly discharge data, and the total number of observation data respectively.

The Nash-Sutcliffe efficiency (NSE) is a normalized statistic that determines the relative magnitude of the residual variance (“noise”) compared to the measured data variance (“information”) (Nash & Sutcliffe 1970). NSE indicates how well the plot of observed versus simulated data fits the 1:1 line. NSE is computed as shown in equation 15:

$$NSE = 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (15)$$

where $X_i, Y_i; \bar{X}, \bar{Y}$; and n denote the i th measured monthly discharge data, the i th simulated monthly discharge data, the mean of the measured monthly discharge data, the mean of the simulated monthly discharge data, and the total number of observation data respectively.

NSE values vary from minus infinity to one, with $NSE = 1$ being the optimal value. When it is lower than 0, the mean predicts the variable better than the model; when it is between 0 and 1, the model works with standard cases, and the closer to 1, the better the correspondence between observed and modelled data (Fernández-Alberti et al. 2021). According to Moriasi et al., (2007), values between $0.75 < NSE \leq 1$ rated as very good; $0.65 < NSE \leq 0.75$ as good; $0.5 < NSE \leq 0.65$ and as satisfactory; and $NSE \leq 0.5$ rated as unsatisfactory for monthly data.

Another model performance evaluation measure is the Index of Agreement (IA). The index of agreement (IA) was developed by Willmott (1981) as a standardized measure of the degree of model prediction error and varies between 0 and 1. A computed value of 1 indicates a perfect agreement between the measured and predicted values, and 0 indicates no agreement at all (Willmott 1981). IA is computed as shown in equation 16:

$$IA = 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (|Y_i - \bar{X}| + |X_i - \bar{X}|)^2} \quad (16)$$

where X_i , Y_i ; \bar{X} , \bar{Y} ; and n denote the i th measured monthly discharge data, the i th simulated monthly discharge data, the mean of the measured monthly discharge data, the mean of the simulated monthly discharge data, and the total number of observation data respectively.

PBIAS (percent bias) calculates the average tendency of simulated values to be larger or smaller than observed values (Gupta et al. 1999). PBIAS with a smaller magnitude shows better model performance. A positive PBIAS indicates overestimation and a negative PBIAS indicates an underestimation of flows (Condom et al. 2011). PBIAS can be represented as follows:

$$PBIAS = \frac{\sum_{i=1}^n (X_i - Y_i)}{\sum_{i=1}^n X_i} \times 100 \quad (17)$$

where X_i , Y_i ; \bar{X} , \bar{Y} ; and n denote the i th measured monthly discharge data, the i th simulated monthly discharge data, and the total number of observation data respectively.

Kling-Gupta Efficiency (KGE), is increasingly being used as an alternative or a complementary metric for the traditional Nash-Sutcliffe Efficiency (NSE). The Kling-Gupta Efficiency (KGE) addresses several shortcomings in NSE and is increasingly used for model calibration and evaluation (Gupta et al. 2009). Like NSE, $KGE = 1$ indicates perfect agreement between simulations and observations. It is preferred that to have higher efficiency values over lower values, because this indicates their model is closer to perfectly reproducing observations (i.e. $KGE = 1$). KGE is computed as follows;

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (18)$$

where r is the linear correlation between observations and simulations, α represents the measure of the flow variability error, and β is the bias.

These methods of model performance evaluation metrics have been widely used in WEAP evaluation studies and have successfully been implemented (Ingol-Blanco & McKinney 2013; Tena et al. 2019; Yaykiran et al. 2019; Abera Abdi & Ayenew 2021) for measured and simulated flows over the calibration and validation periods.

3.6 Methodology Matrix

Below is a table that is showing the data collected, its source, how the data was analyzed and the software used under each specific objective.

Table 4: Methodology Matrix

Specific Objective	Data collected	Data Source	Data Analysis	Software / Tools
a) To assess the availability of observed hydrological data within the basin.	<ul style="list-style-type: none"> • Observed Streamflow data 	<ul style="list-style-type: none"> • Rumphi District Water Office 	<ul style="list-style-type: none"> • Data Completeness Measures 	<ul style="list-style-type: none"> • MS Excel
b) To develop a remote sensing rainfall runoff model	<ul style="list-style-type: none"> • Elevation Data • Land Use Land Cover Data (LULC) • Climate Data (Precipitation, Temperature & Wind Speed) 	<ul style="list-style-type: none"> • Hydro SHEDS database • ESA – CCI¹ project • Global Land Data Assimilation System (GLDAS) 	<ul style="list-style-type: none"> • Descriptive Statistics 	<ul style="list-style-type: none"> • R Studio • WEAP Modelling Software
c) To conduct a performance analysis of the remote sensing rainfall runoff model	<ul style="list-style-type: none"> • Simulated streamflow data • Observed Streamflow data 	<ul style="list-style-type: none"> • WEAP • Rumphi District Water Office 	<ul style="list-style-type: none"> • Measures of goodness-of-fit of a model 	<ul style="list-style-type: none"> • R Studio • MS Excel

¹ European Space Agency (ESA) Climate Change Initiative (CCI) project

Figure 8 below is showing the flow chart of the methodology adopted in this study.

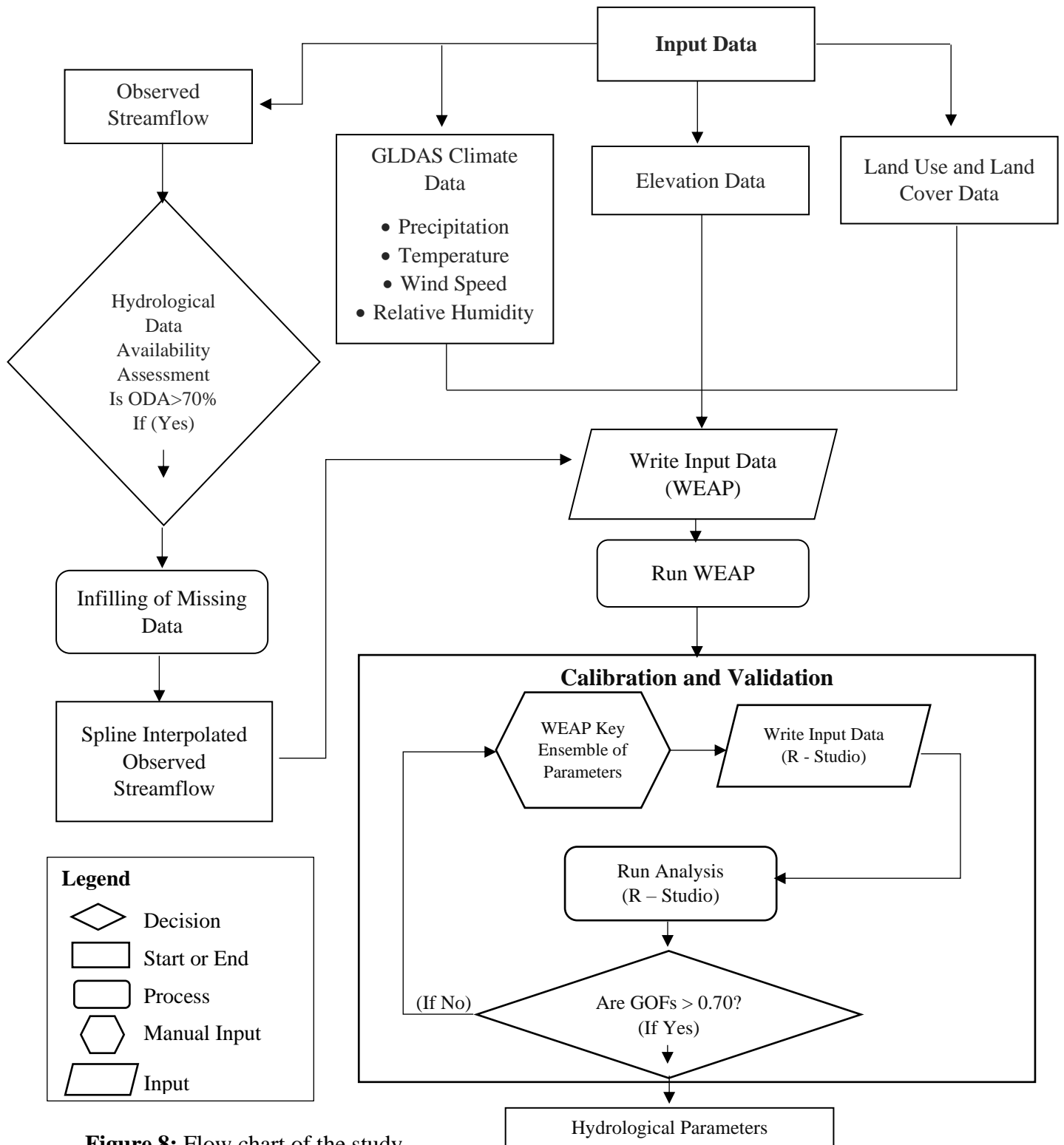


Figure 8: Flow chart of the study

CHAPTER 4: RESULTS

4.1 Observed Hydrological Data Availability Assessment

Table 5 presents the proportions of missing data and data availability assessment criteria for the collected observed daily streamflow data from four hydrometric stations; 7F1 Runyina at Chikwawa, 7G14 South Rukuru at Phwezi, 7G18 South Rukuru at Mlowe, 7H3 North Rumphu at Chiweta from 2000 to 2021.

Table 5: Proportions of missing data and hydrological data availability assessment

Station	Proportion of Missing Data (%)	Overall Availability (%)	Longest Availability Without Gaps (%)	Continuity in Available Data (%)
7F1 Runyina at Chikwawa	25	75	1	1
7G14 South Rukuru at Phwezi	31	69	9	13
7G18 South Rukuru at Mlowe	8	92	5	5
7H3 North Rumphu at Chiweta	18	82	2	3

4.1.1 Missing Data Proportions and Overall Data Availability

Each hydrometric station had gaps or missing data for the entire 2000 – 2021 period when the daily streamflow data was investigated. Figure 9 is showing the distribution of missing data in the time series stream flow data for the four hydrometric stations. To show the extent of missing data in the distribution plots, a single brown line shows a missing data value, thus where the brown line is thick, that shows a huge gap of data missing.

For instance, station 7F1 had an enormous missing data in the year 2004 from the month of May to August, with 199 missing data values which constitutes 84% of the missing data. In station 7G18, it can also be visualized that from July, 2009 to June, 2010 there were no gaps, not even a single brown line appears within the respective period. In addition, it can be clearly shown that for station 7G14 from 2017 to 2021 there is no daily streamflow data, consequently having a very thick brown line, which shows that there is a huge gap of missing data for the station.

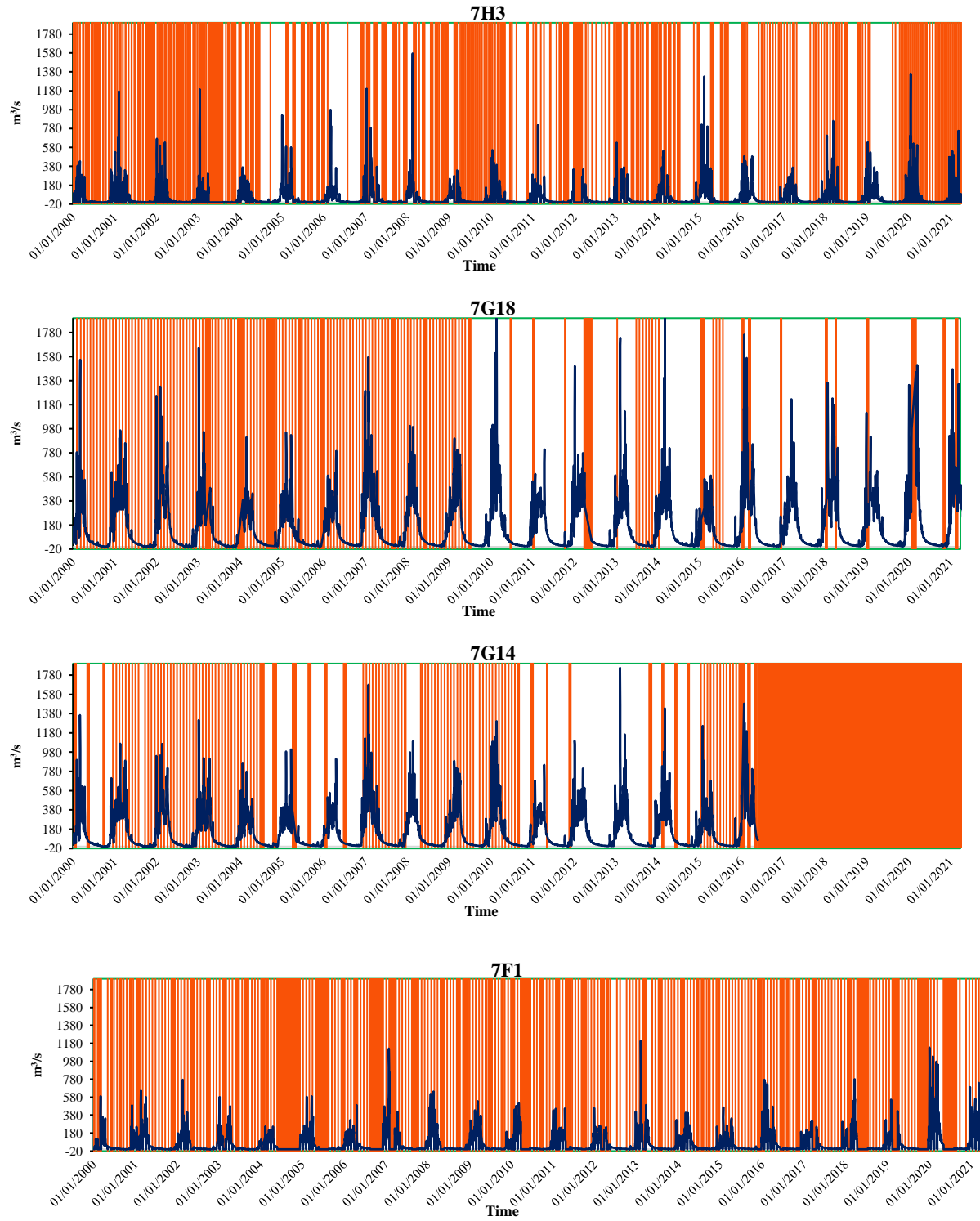


Figure 9: Distribution of missing data in the time series stream flow data for the four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period.

For each hydrometric station and stream flow time series data, The Overall Data Availability (ODA) was calculated as follows;

$$ODA = \frac{\text{Total data values recorded}}{\text{Total data values that could have been recorded from 2000–2021}} \times 100 \quad (19)$$

Where the ‘total data values recorded’ are the non-null values in the data set under measure, and the ‘total data values that could have been recorded’ are the reference data set that has no nulls (where all attributes are mandatory) for the whole study period; thus the total data values could have been recorded from 2000 – 2021 is 8,036. Below is Table 6, showing the total data values recorded for the entire period from 2000 – 2021.

Table 6: Total data values recorded for entire period from 2000 – 2021.

Station	Total Available Data from 2000 - 2021
7F1 Runyina at Chikwawa	6,019
7G14 South Rukuru at Phwezi	5,541
7G18 South Rukuru at Mlowe	7,419
7H3 North Rumphu at Chiweta	6,620

In Figure 9, it shows the results of the overall data availability that was assessed on the four stream flow data time series collected for the entire 2000 – 2021 period, the overall data availability for station 7F1 shows to be slightly higher (75%) compared to 7G14 (69%) which has the lowest overall data availability and station 7G18 has the best overall availability (92%) with only 8% missing data. Station 7H3 comes second best with overall availability of 82%.

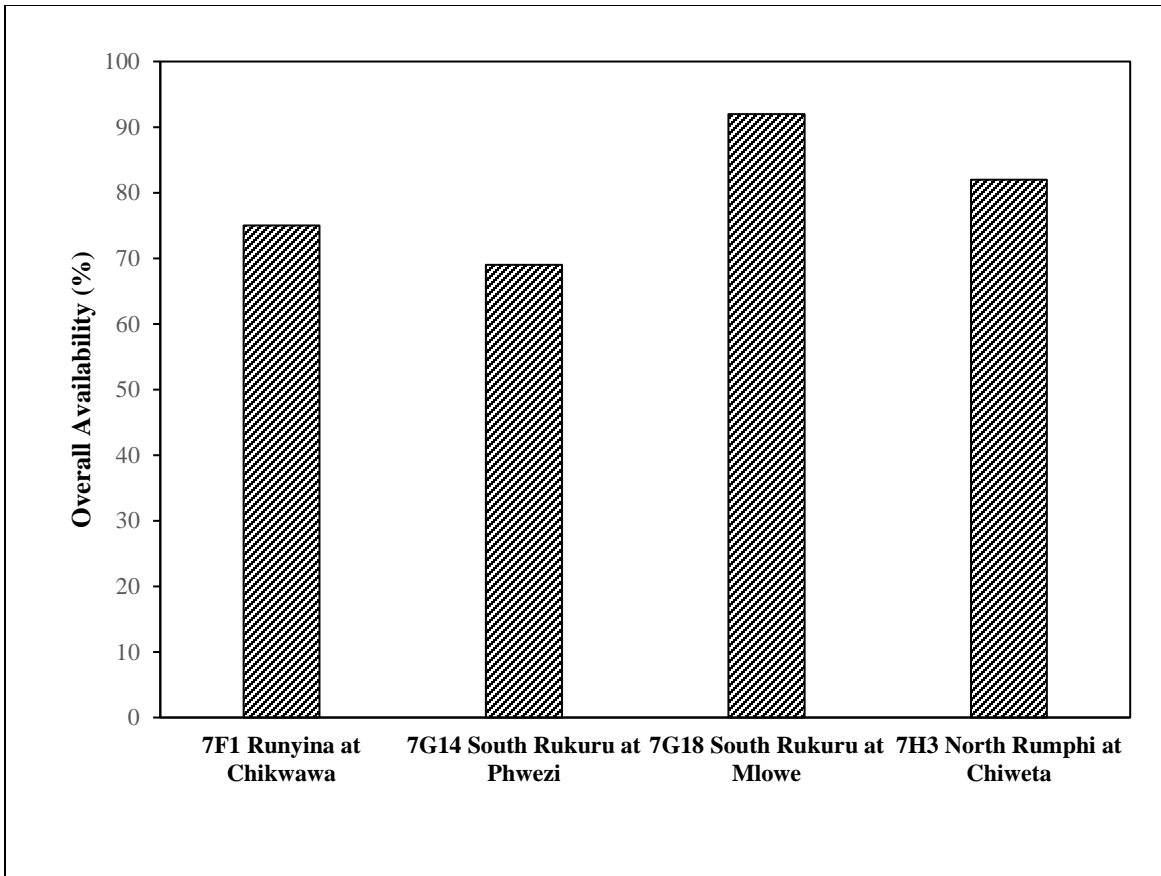


Figure 10: Overall data availability for four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period.

The overall data availability was also assessed yearly, and it was calculated as follows;

$$ODA = \frac{\text{Total data values recorded in a year}}{\text{Total data values that could have been recorded in a year}} \times 100 \quad (20)$$

where the total data values that could have been recorded in a year is 365 for a normal year and 366 for a leap year.

The bar plots show that all stations had lowest overall data availability in 2004 except station 7G14 which from 2017 to present there is no daily streamflow data that is being collected, showing 0%

overall data availability. Only two stations 7G14 and 7G18 achieved 100% overall data availability per year; the two stations had all daily streamflow data in years 2012 and 2017 respectively.

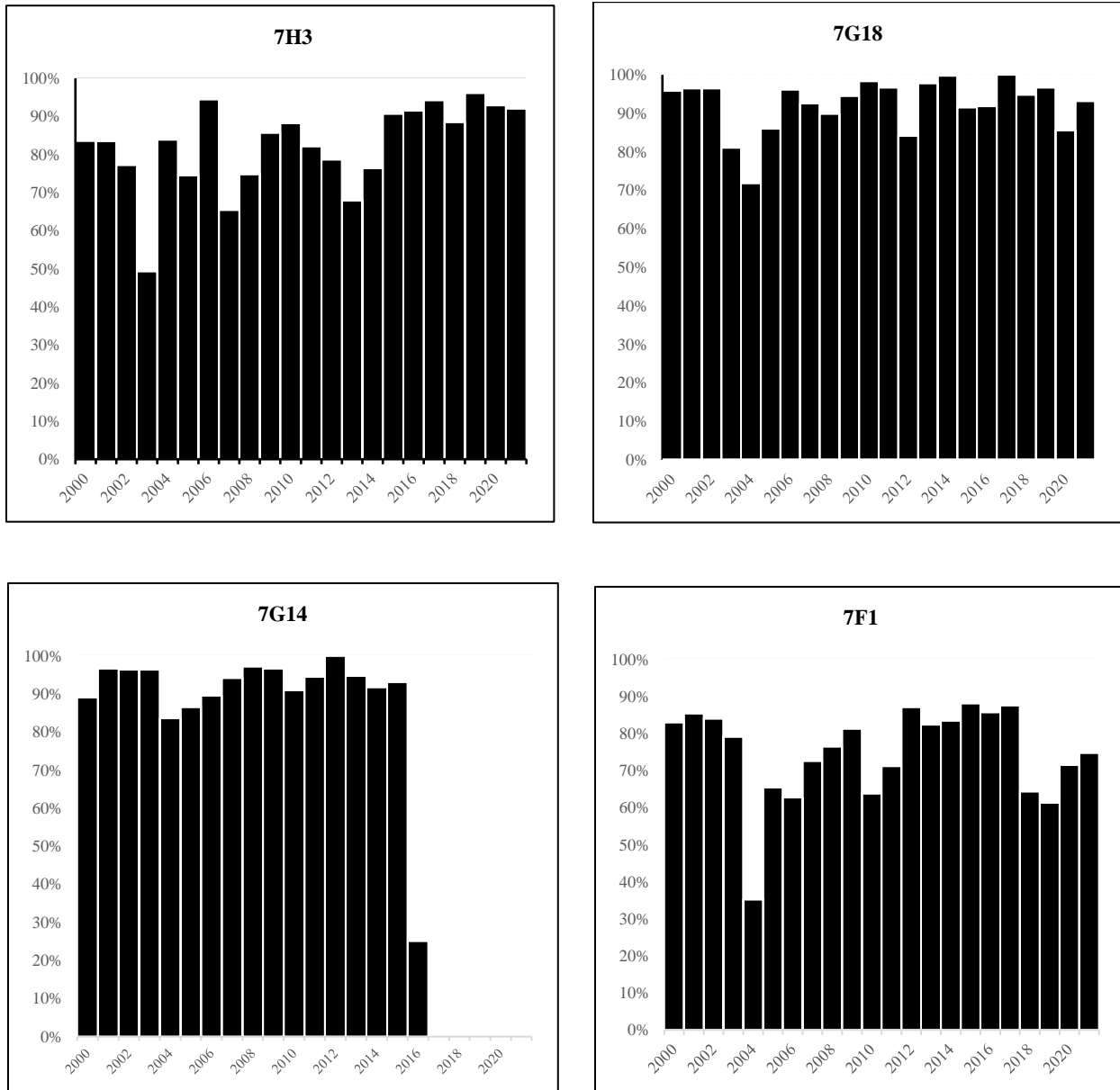


Figure 11: Results of the overall data availability per year from 2000 to 2021 for hydrometric stations 7H3, 7G18, 7G14 and 7F1 respectively

4.1.2 Longest Data Availability without Gaps

The time series of streamflow data for each station was inspected of having continuous data (i.e. without missing values) without gaps. The percentage of the longest data availability without gaps (LDA) was computed as;

$$LDA = \frac{\text{Longest Data Values Length without Gaps}}{\text{Total data values that could have been recorded from 2000–2021}} \times 100 \quad (21)$$

Figure 11 is showing the longest data availability without gaps for each hydrometric station for the entire 2000 – 2021 period.

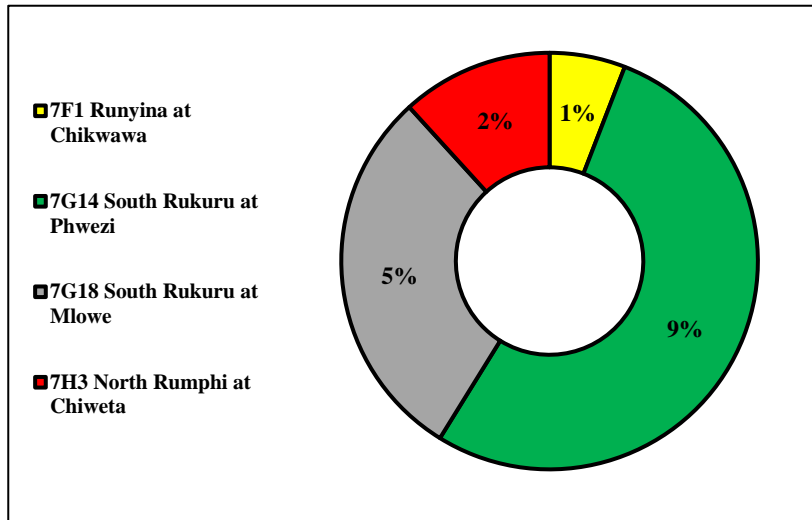


Figure 12: Longest data availability without gaps for the four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period.

From Figure 11, station 7G14 has 9% as the longest data availability without gaps having 688 values continuously without missing data which is the longest data length for the entire 2000 – 2021 period, this longest data length only occurred once for the entire 2000 – 2021. Nevertheless, station 7F1 has the lowest longest data availability only 1% having only 55 values continuously without missing data for the entire 2000 – 2021, still the station (7F1) has a continuous data length

of 27 values that occurred for 176 times (Table 6) for the 2000 – 2021 period at different time steps. Stations 7G18 and 7H3 have the longest data availability as 5% and 2% respectively, with longest continuous data lengths of 368 and 197 respectively.

Furthermore, the occurrence of the continuous data length was investigated of the entire time series streamflow data for each station, subsequently the total available data for each continuous data length and percentage of the total available data to the total available data for entire period was computed. Table 7 below shows the data length without gaps, occurrence of the continuous data length, the total available data and percentage of the total available data for each station.

The total available data for each continuous data length is computed as follows;

$$\text{Total Available Data} = \text{Continuous Data Length} \times \text{Number of Occurrences of the Data Length} \quad (22)$$

The percentage of the total available data has been calculated as follows;

$$\% \text{ Total Available Data} = \frac{\text{Total Available Data}}{\text{Total Available Data from 2000–2021}} \times 100 \quad (23)$$

From Table 7, the highest percentage of the total available data is occurring at the same continuous data length of 27 from stations 7G18, 7G14 and 7F1 is 45%, 47% and 79% respectively except 7H3 which has its highest percentage 14% at a continuous data length of 24. For further clarification, for station 7G18, it means that at continuous data length of 27, there is the maximum occurrence of the continuous data without gaps of 124, having a total available data of 3348. However, the longest data availability without gaps for 7G18 occurred at 368 and only occurred once which is only contributing 5% of the total available data.

Table 7: Continuous Data Length without Gaps (DL), Number of Occurrences (NOO), Total Available Data (TAD) and Percentage of the Available Data (% TAD) for each Continuous Data Length for stations 7H3, 7G18, 7G14 and 7F1 for the entire 2000 – 2021 period

7H3				7H3				7G18				7G14				7F1			
DL	NOO	TAD	% TAD	DL	NOO	TAD	% TAD	DL	NOO	TAD	% TAD	DL	NOO	TAD	% TAD	DL	NOO	TAD	% TAD
1	8	8	0%	34	2	68	1%	2	1	2	0%	8	1	8	0%	1	6	6	0%
2	5	10	0%	36	2	72	1%	3	3	9	0%	10	1	10	0%	2	3	6	0%
3	10	30	0%	37	2	74	1%	4	1	4	0%	14	2	28	1%	3	2	6	0%
4	4	16	0%	38	1	38	1%	6	1	6	0%	27	97	2619	47%	4	1	4	0%
5	10	50	1%	40	1	40	1%	10	1	10	0%	36	1	36	1%	5	3	15	0%
6	5	30	0%	42	1	42	1%	11	1	11	0%	41	1	41	1%	6	5	30	0%
7	7	49	1%	43	1	43	1%	12	1	12	0%	55	2	110	2%	7	3	21	0%
8	7	56	1%	45	2	90	1%	14	1	14	0%	72	1	72	1%	8	5	40	1%
9	6	54	1%	46	4	184	3%	18	1	18	0%	84	1	84	2%	9	4	36	1%
10	7	70	1%	47	1	47	1%	27	124	3348	45%	93	1	93	2%	10	2	20	0%
11	39	429	6%	49	1	49	1%	31	1	31	0%	99	1	99	2%	11	1	11	0%
12	38	456	7%	52	2	104	2%	33	1	33	0%	101	1	101	2%	12	3	36	1%
13	5	65	1%	54	1	54	1%	44	1	44	1%	103	2	206	4%	13	3	39	1%
14	5	70	1%	57	1	57	1%	46	1	46	1%	104	1	104	2%	14	3	42	1%
15	7	105	2%	58	1	58	1%	64	1	64	1%	110	1	110	2%	15	2	30	0%
16	11	176	3%	62	1	62	1%	73	1	73	1%	122	2	244	4%	17	3	51	1%
17	3	51	1%	63	1	63	1%	77	1	77	1%	125	1	125	2%	18	5	90	1%
18	12	216	3%	69	1	69	1%	79	1	79	1%	131	1	131	2%	19	1	19	0%
19	5	95	1%	88	1	88	1%	90	1	90	1%	143	1	143	3%	20	2	40	1%
20	2	40	1%	95	2	190	3%	165	1	165	2%	148	1	148	3%	22	2	44	1%
21	5	105	2%	100	1	100	2%	167	1	167	2%	151	1	151	3%	23	1	23	0%
22	6	132	2%	102	1	102	2%	170	1	170	2%	190	1	190	3%	24	2	48	1%
23	5	115	2%	116	1	116	2%	190	1	190	3%	688	1	688	12%	25	4	100	2%
24	39	936	14%	119	1	119	2%	223	1	223	3%					26	4	104	2%
25	2	50	1%	120	1	120	2%	241	1	241	3%					27	176	4752	79%
26	4	104	2%	123	1	123	2%	266	1	266	4%					29	1	29	0%
27	1	27	0%	138	1	138	2%	270	1	270	4%					48	1	48	1%
28	1	28	0%	174	1	174	3%	271	1	271	4%					55	6	330	5%
29	2	58	1%	197	1	197	3%	350	1	350	5%								
30	5	150	2%					371	1	371	5%								
31	2	62	1%					378	1	378	5%								
32	3	96	1%					386	1	386	5%								

4.1.3 Continuity in Available Data

In the previous sections of the results for observed hydrological data availability assessment, the overall availability and the longest data availability without gaps have been computed. The final assessment that was done on the stream flow time series collected was to check for the continuity in the available data, which is calculated as follows;

$$CAD = \frac{\text{Longest Data Availability}}{\text{Overall Availability}} \quad (24)$$

From Figure 12, the continuity of available data appears to be highest on station 7G14 which is 13%. Stations 7G18 and 7H3 have the continuity as 5% and 3% respectively, with station 7F1 having the lowest continuity in available data, only 1%

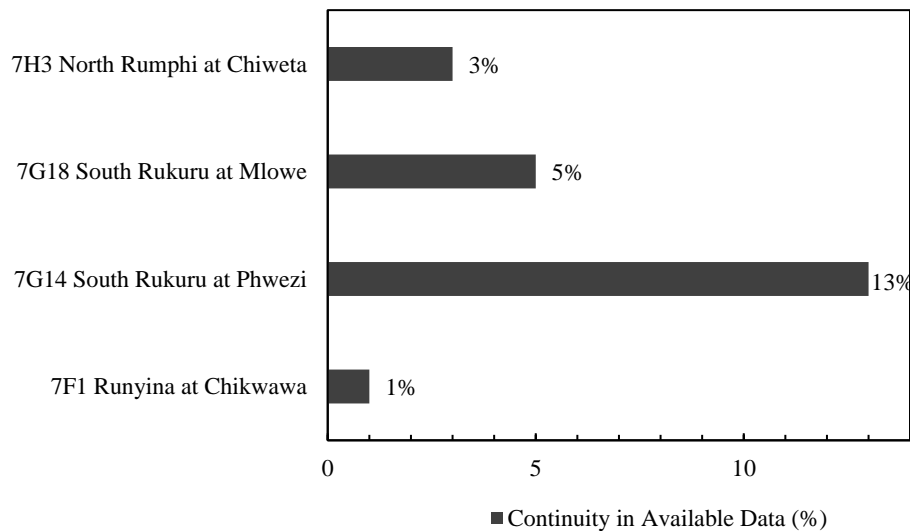


Figure 13: Bar plot of the continuity of available data for four hydrometric stations (7H3, 7G18, 7G14 and 7F1) for the entire 2000 – 2021 period.

The continuity in available data (CAD) was also assessed yearly, and it was calculated as follows;

$$CAD = \frac{\text{Longest Data Availability without Gaps}}{\text{Overall data availability in a year}} \quad (25)$$

The line plots were produced and they display that 7G14 and 7G18 are the only stations that achieved 100% continuity in available data per year. However, station 7G18 the 100% continuity in available data occurred twice (2017 and 2019) while 7G14 only occurred once in 2012. Station 7G18 also shows that it had low continuity from 2000 to 2009, but from 2010 to 2021 it has the highest continuity in comparison to other stations.

Station 7F1 has its highest continuity (21%) in available data in 2020. Station 7F1 shows to have the lowest continuity in available data for the entire 2000 – 2021 period except in some of the years where it surpasses station 7H3. Station 7H3 achieved the highest continuity in available data (56%) in 2019 but there is a decrease within 2020. Then in 2021, the station has the lowest continuity (4%). Worst case scenario for continuity in available data has occurred in station 7G14, where from 2017 to 2021 there are only gaps. Figure 13 below shows the line plots of the yearly continuity in available data for the stations 7H3, 7G18, 7G14 and 7F1 from 2000 – 2021 period.

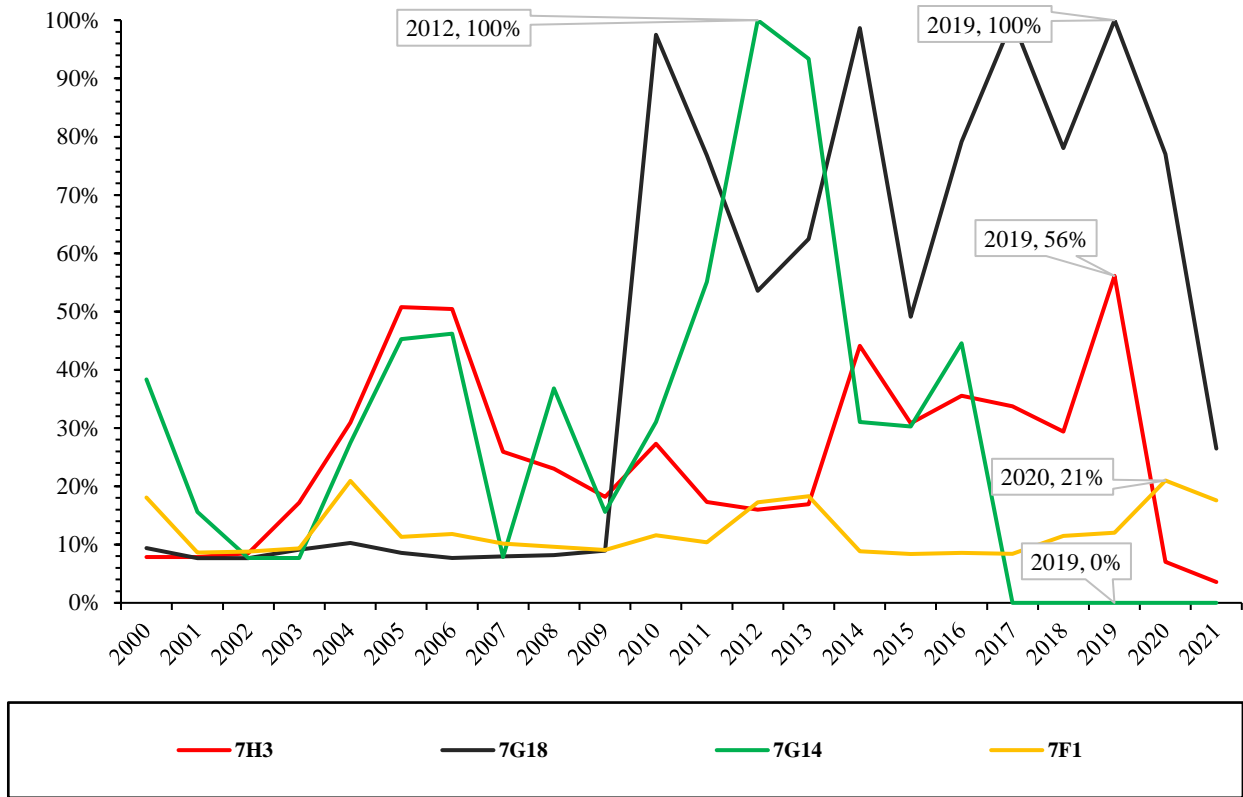


Figure 14: Results of the continuity of the available data per year from 2000 to 2021 for hydrometric stations 7H3, 7G18, 7G14 and 7F1 respectively.

4.2 The Remote Sensing WEAP Model

4.2.1 WEAP Model: Calibration and Validation

As soon as the WEAP hydrological model was developed, the model was able to simulate the streamflow and other hydrological parameters with default soil and land use related parameters. Model calibration followed to estimate the parameters using the manual (trial-and-error) method until a good fit was observed between the measured and simulated streamflow. These estimated parameters that resulted from calibration were also used to validate the model.

Table 8 below is showing the values of these initial soil and land use – related parameters and values for the estimated soil and land use-related parameters.

Table 8: Values of these initial soil and land use – related parameters (default) and values for the estimated soil and land use-related parameters.

Parameters	Definition	Default	Range in Values
K_c	Crop Coefficient	1	0.45 – 0.5
S_w	Root Zone Soil Water Capacity (mm)	1,000	1,150
D_w	Deep Soil Water Capacity (mm)	1,000	100 - 135
RRF	Runoff Resistance Factor	2	2.25
K_s	Conductivity of Root Zone at Full Saturation (mm/month)	20	10
K_d	Conductivity of Deep Zone at Full Saturation (mm/month)	20	20
F	Preferred Flow Direction	0.15	0.15
Z₁	Initial storage fraction at the beginning of simulation of upper soil layer	30%	40%
Z₂	Initial storage fraction at the beginning of simulation of lower soil layer	30%	40%

For visual inspection, Figure 15 is showing monthly hydrographs of observed and simulated stream flows for hydrometric stations 7H3 and 7G18 for the entire 2000 – 2021 period. The model was run for 22 years of 2000 – 2021 period but the first year (2000) was used for stabilizing the model runs and the simulated streamflow from 2001 to 2021 were used for comparison purposes. The WEAP model was calibrated and validated using data of these two streamflow gauging stations (7H3 and 7G18). The calibration period extended from January 2001 to August 2015, which is 70% of the data period while the validation period is from September, 2015 to December 2021 which is 30% of the data period. The division of the periods is the same for both catchments.

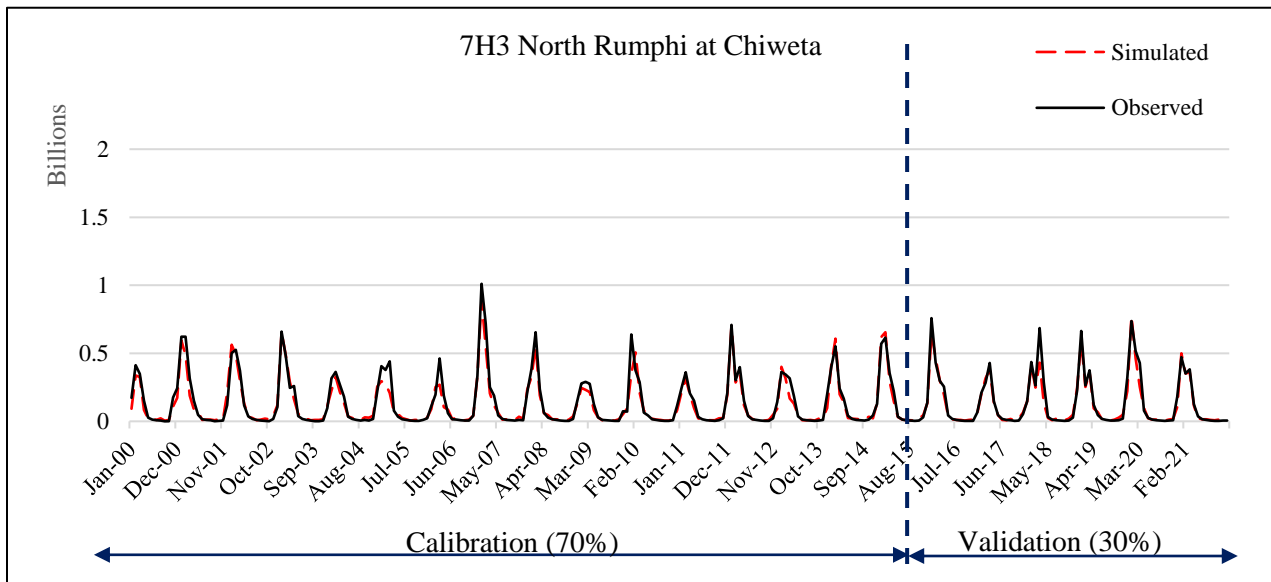
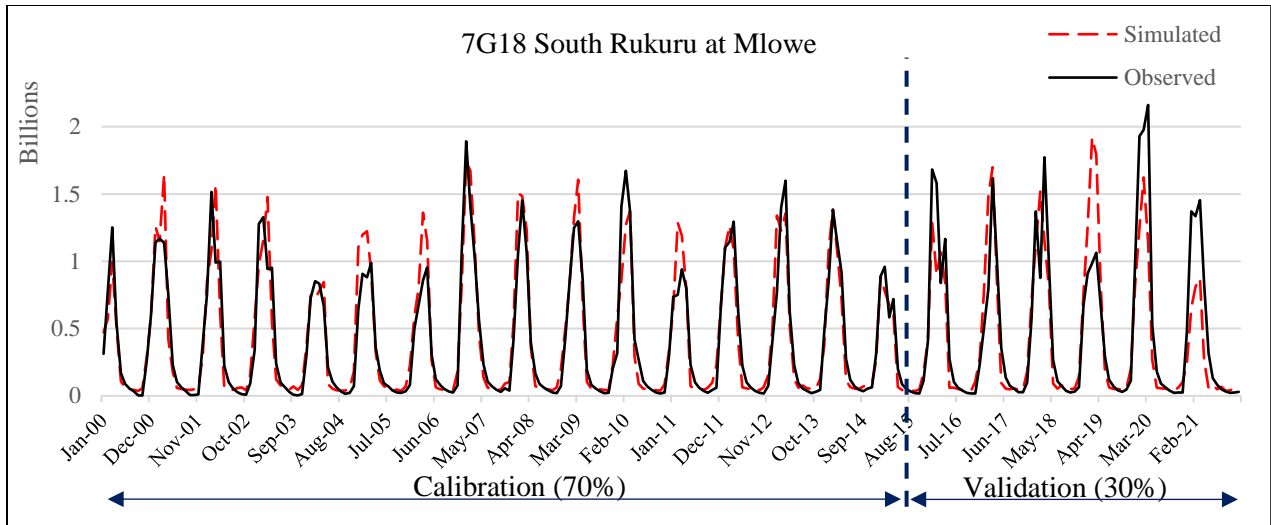


Figure 15: Monthly observed and simulated stream flows for hydrometric stations 7H3 and 7G18 for the entire 2000 – 2021 period.

4.2.2 WEAP Model Hydrological Processes

The WEAP model developed is capable of estimating nearly all components of the hydrological cycle. The main vital components for the hydrological cycle in a specific area are precipitation, surface runoff, and evapotranspiration. The monthly behavior of these terms for each catchment area, as simulated by the WEAP model using the available input data, is shown in Figure 16.

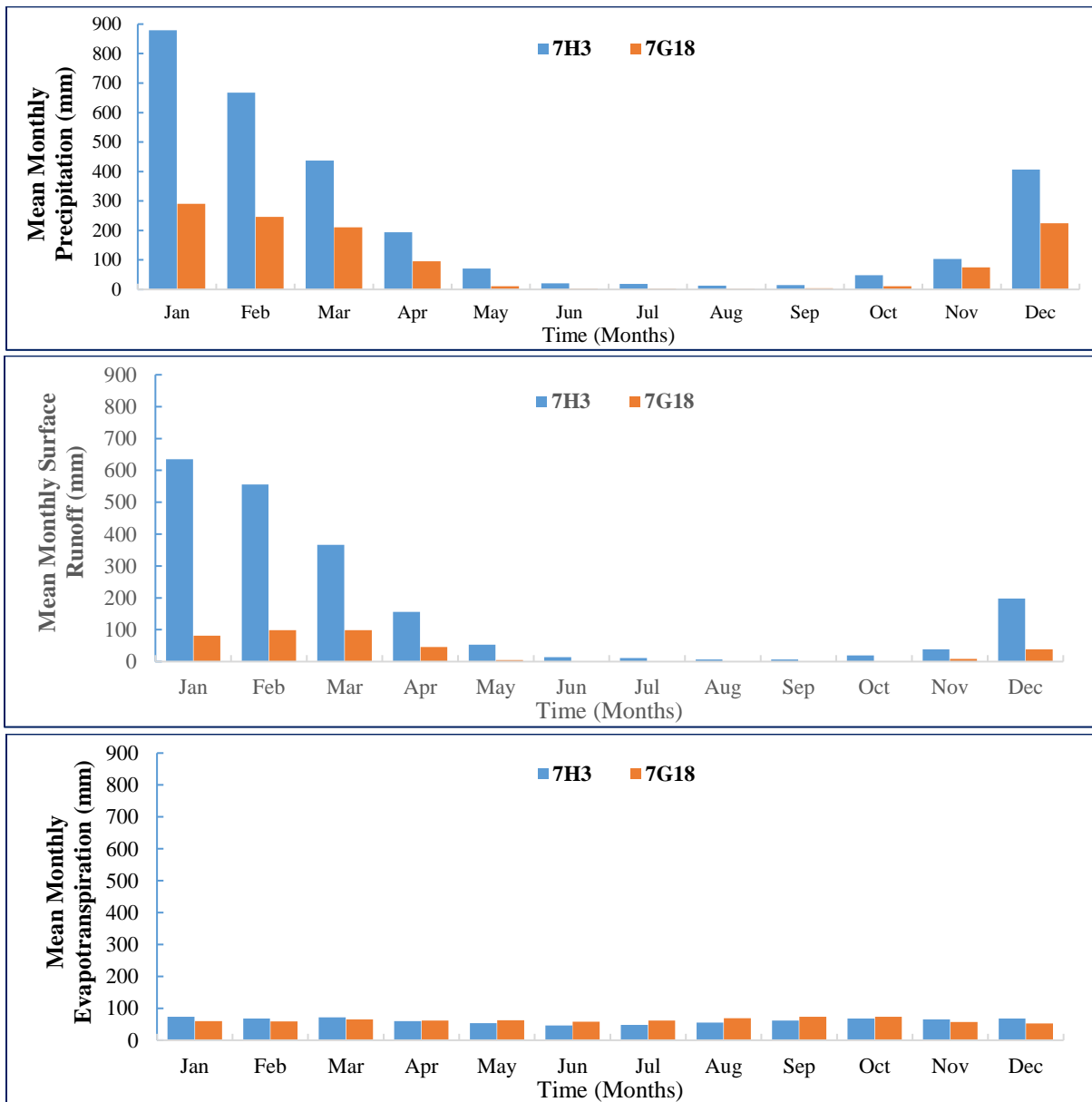


Figure 16: Results of the WEAP simulations: Calculated terms of the mean monthly water balance for each sub-basin (mm). Precipitation, Runoff and Evapotranspiration respectively.

The precipitation is around 2, 873 mm/ year, having higher values in December, January, February and March, while surface runoff and evapotranspiration reach 2, 060 and 740 mm/year, respectively. Similarly, with the South Rukuru sub-basin, the peak precipitation and surface runoff occurs each year from December to March. The precipitation, surface runoff and evapotranspiration reach around 1, 175; 381 and 756 mm/year, respectively.

Figure 17 is showing the water balance monthly average for hydrometric stations 7G18 and 7H3 respectively. For both sub-basins, the water availability is highest in the months of peak precipitation, with lowest water availability between June and September. In some months, the evapotranspiration exceeded the amount of precipitation in May, June, July, August, September and October. On the contrary, in months ranging from November to March, the evapotranspiration was less than the precipitation amounts.

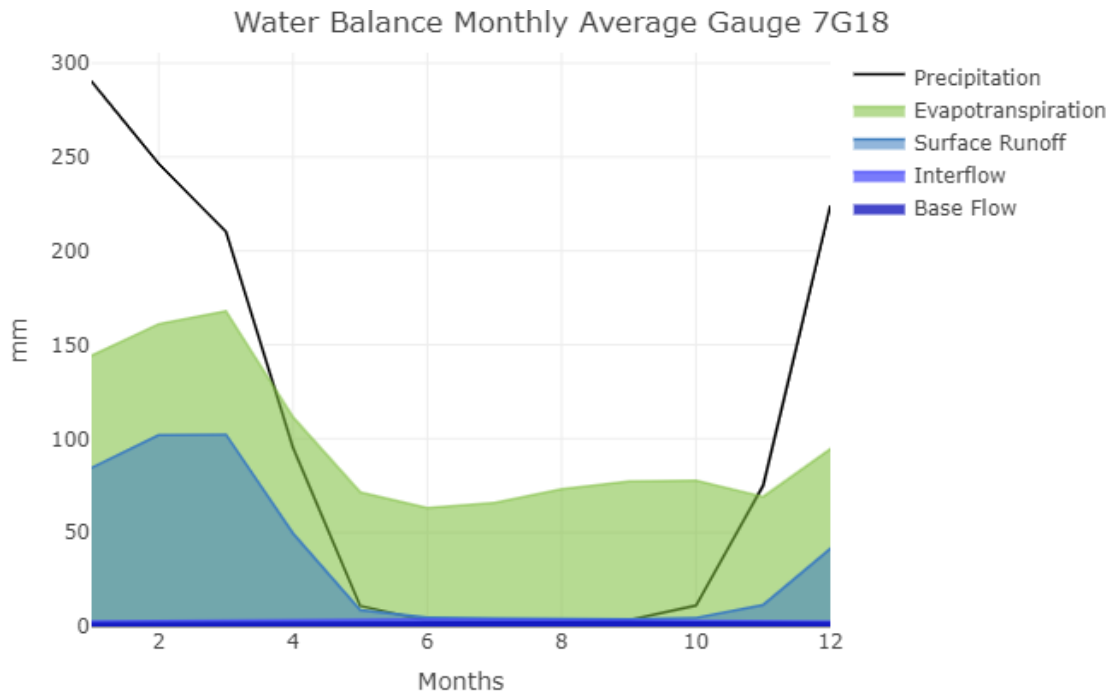
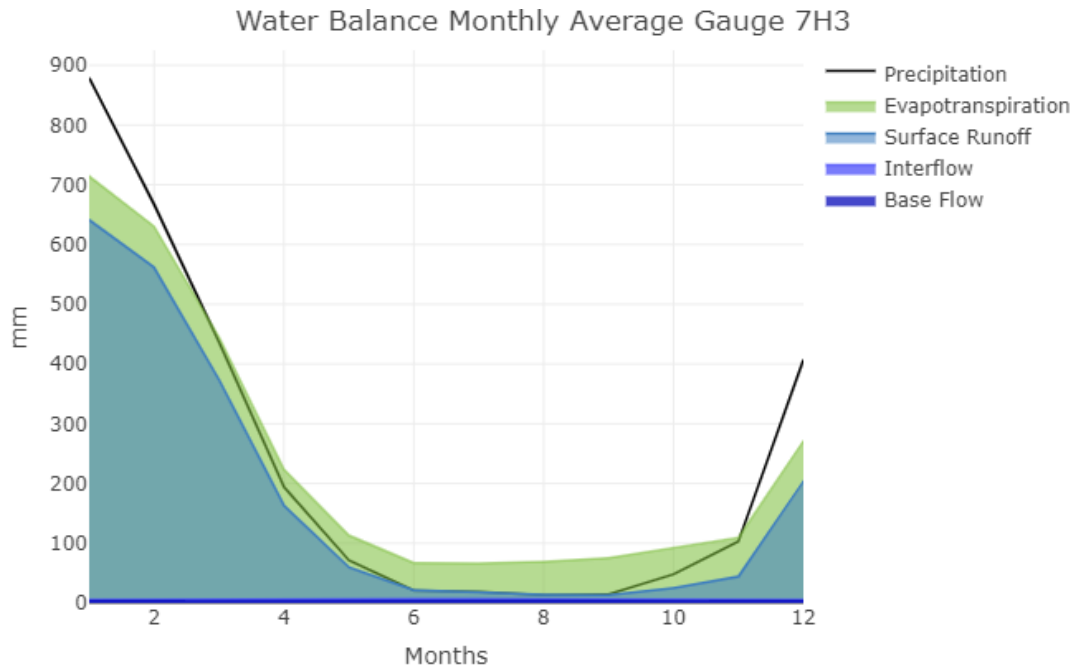


Figure 17: Results of water balance monthly average for hydrometric stations 7H3 and 7G18 respectively.

4.3 WEAP Model Performance Evaluation

As shown in Table 9, five well known statistical criteria were used to evaluate the ability of the model to simulate the measured stream flows; The coefficient of determination (R^2); Nash-Sutcliffe coefficient of efficiency (NSE); index of agreement (IA); Kling Gupta Efficiency (KGE) and Percent Bias (PBIAS). For comparison between the measured and simulated stream flows, goodness of fit metrics presented in Table 9, showed a stronger agreement for both stations with R^2 of 0.87, NSE of 0.86, IA of 0.97, KGE of 0.93 and PBIAS of -0.2 for station 7G18. Similarly with station 7H3, R^2 of 0.93, NSE of 0.92, IA of 0.98, KGE of 0.84 and PBIAS of -10 during the calibration period. For the validation period, there was a reasonably very good agreement between the measured and simulated streamflow during this period for the two stations with R^2 of 0.73, NSE of 0.72, IA of 0.92, KGE of 0.79 and PBIAS of -10.9 for station 7G18. Likewise for station 7H3, R^2 of 0.95, NSE of 0.94, IA of 0.98, KGE of 0.87 and PBIAS of -7.7 during the validation period.

Table 9: Goodness of Fit Metrics (GOF) for Hydrometric Stations 7G18 and 7H3 respectively.

Gauge	Type	Period	Goodness of Fit Metrics				
			NSE	IA	R^2	KGE	PBIAS
7G18	Overall	Jan 2001 to Dec 2021	0.81	0.95	0.81	0.89	-3.8
7G18	Calibration (70%)	Jan 2001 to Aug 2015	0.86	0.97	0.87	0.93	-0.2
7G18	Validation (30%)	Sept 2015 to Dec 2021	0.72	0.92	0.73	0.79	-10.9
7H3	Overall	Jan 2001 to Dec 2021	0.93	0.98	0.94	0.85	-9.3
7H3	Calibration (70%)	Jan 2001 to Aug 2015	0.92	0.98	0.93	0.84	-10
7H3	Validation (30%)	Sept 2015 to Dec 2021	0.94	0.98	0.95	0.87	-7.7

Overall, station 7H3 has the strongest agreement with R^2 of 0.94, NSE of 0.93 and IA of 0.98, whereas station 7G18 with R^2 of 0.81, NSE of 0.81 and IA of 0.95. On the other hand, station 7G18 achieved a good performance with KGE of 0.89 and PBIAS of -3.8 compared to station 7H3 which has KGE of 0.85 and PBIAS of -9.3. There was a reasonably very good agreement between the observed and simulated streamflow, however the goodness of fit metrics tend to be higher during the calibration period for station 7G18 while for station 7H3 the metrics are higher during the validation period. Visual comparison of observed and simulated monthly hydrographs (Figure 15) also shows this satisfactory agreement. Figure 16 shows the visual comparison of the observed and simulated streamflow using the coefficient of determination (R^2) during both the calibration and validation periods for both hydrometric stations 7G18 and 7H3.

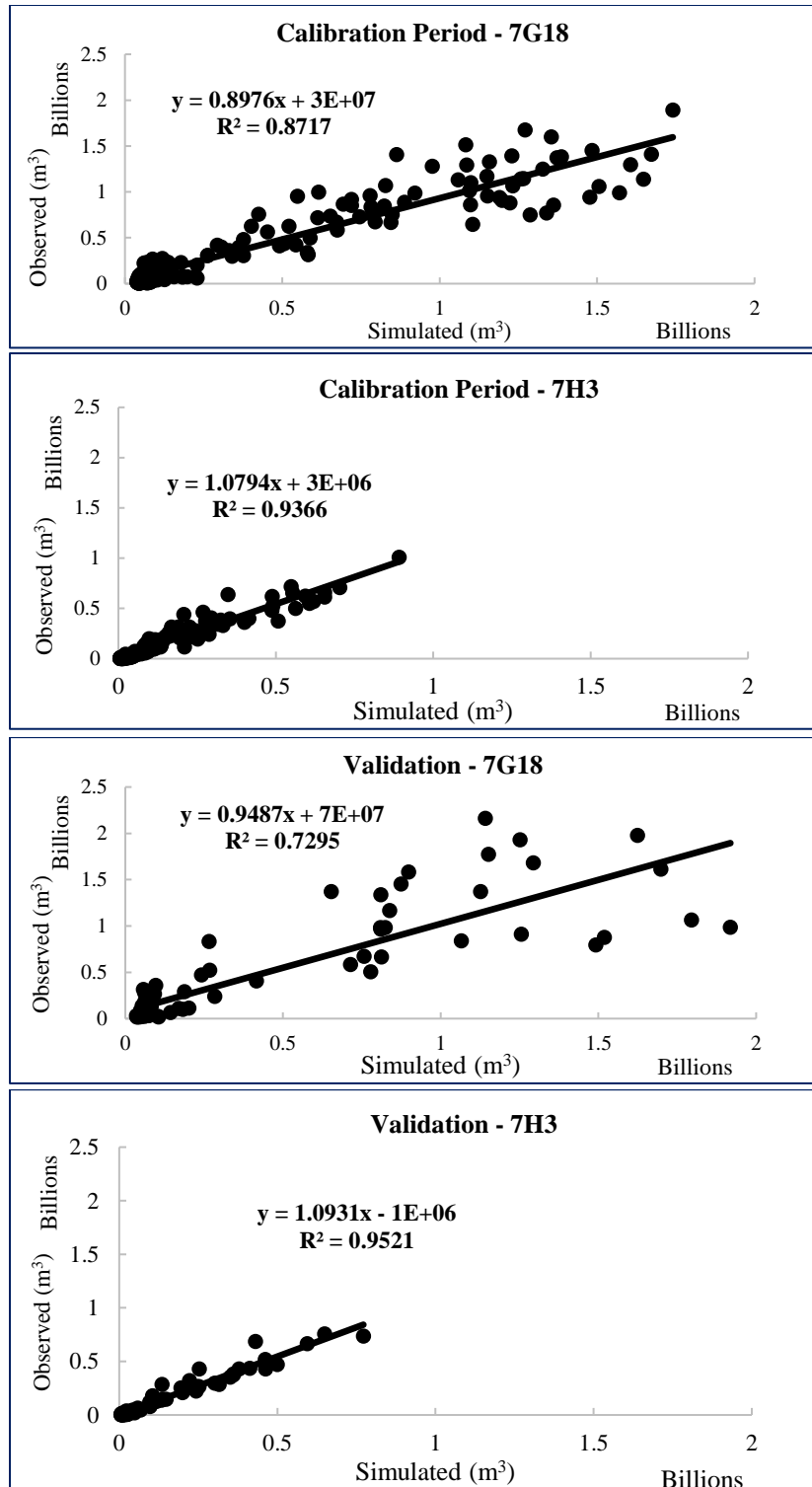


Figure 18: Visual comparison of the observed and simulated streamflow using the coefficient of determination (R^2) during both the calibration and validation periods for hydrometric station 7G18 and 7H3.

CHAPTER 5: DISCUSSION

5.1 Observed Hydrological Data Availability Assessment

Poor availability and accuracy of streamflow data makes it difficult to develop hydrological models. When a hydrological model is being developed, whether at model set up, calibration and validation, there is a need to undergo quality checks of available streamflow observations (Crochemore et al. 2020). For this research, quality checks were regarded as the hydrological data availability assessment which mainly focused on data completeness analysis i.e. to check if the data was complete enough to be used to develop a hydrological model.

Hence, the observed daily streamflow data from four hydrometric stations; 7F1 Runyina at Chikwawa, 7G14 South Rukuru at Phwezi, 7G18 South Rukuru at Mlowe, 7H3 North Rumphi at Chiweta from 2000 to 2021, undergone the minimum quality checks of data availability which were adapted from Crochemore et al. (2020); thus the overall availability, longest availability and continuity of available data. As indicated by Arriagada et al. (2021) incomplete streamflow series are more frequent and gaps in the data are generally longer in most of the developing countries including Malawi. From the results, it has been revealed that indeed there is incomplete time series streamflow data and having gaps throughout the entire period from 2000 – 2021 for all four hydrometric stations within WRA 7.

Up until now, there is no ascertained cut-off based on the literature concerning tolerable proportion of missing rate in streamflow data as convincing statistical inferences. According to Hamzah et al. (2020) the missing data rule of thumb is stated that missing rate of 5 percent or a lesser amount of it, is insignificant. 5% to 10% is minor, 10% to 25% is moderate, and 25% to 50% is high and above 50% is excessive. Herein, I set the threshold for the acceptable percentage of missing data

to consider a gauge station streamflow data usable at 30% of missing data, which is having the overall availability of 70%.

Although the streamflow time series data has a lot of missing data and gaps but with the threshold set, all stations under study had a good allowable overall data availability percentages except station 7G14, which has the overall availability as 69% which is just below the threshold. Station 7G14 has the lowest overall availability because the station had been destroyed during the road construction works within the area in the year 2016 and yet to be re-installed. The station has no streamflow data from 2017 to 2021, which shows how incomplete the data is for the station. These results agree with previous study by (Kumambala 2010), where station 7G18 had a missing proportion percentage of 1.23% from 1958 to 2000 and 7G14 having a percentage of 1.83%, which shows that stations 7G18 and 7G14 had high overall availability. It can be observed that from 2000 to 2021 there is a dense decline of data availability.

This means that the streamflow data from hydrometric stations 7G18, 7H3 and 7F1 can be used to develop the hydrological model. Infilling of missing streamflow data for the three hydrometric station (7F1, 7H3 and 7G18) was done using the spline interpolation methods in Microsoft Soft Excel because when developing a hydrological model, there is need of complete set of observed streamflow data (Kumambala 2010). Missing data introduce bias and uncertainty into hydrological models, making it difficult to accurately simulate the behavior of a particular hydrological system. These results give a clear indication that remotely acquired data can be used as an alternative or complementary data source to assess water availability within river basins of Malawi.

5.2 WEAP Model Hydrological Processes

The annual precipitation cycle for Malawi indicates that the wet season begins from November and ends in April (Tadeyo et al. 2020), which means that the rest of the year can be regarded as a dry season. It was observed, and as expected, that the precipitation for both sub-basins showed a similar cycle as that of Tadeyo et al. (2020). For both sub-basins, the lowest precipitation values occur between June and September. These results also agree with previous observations by Libanda et al. (2017) who stated that the onset and offset of the rainfall season in Malawi is triggered by the position variation of the Intertropical Convergence Zone (ITCZ). The ITCZ dominates most areas of southern Malawi from November before retreating to northern Malawi during March and April, bringing an end to the wet season.

It can be observed that the months at which the precipitation was at the peak in all the sub-basins, the surface runoff is also at peak, since surface runoff occurs when excess precipitation can no longer sufficiently rapidly infiltrate in the soil as the soil becomes saturated. In principle, the runoff component may be utilized to cover different water demands within the sub-basins. However, if there is no adequate infrastructure to accumulate the surface runoff in the rainy season, it will flow directly into the lake (Fernández-Alberti et al. 2021).

Over 60% of global precipitation is consumed by evapotranspiration and it is the largest withdrawal of water from the systems (Abera Abdi & Ayenew 2021). For this study, the evapotranspiration shows steadiness throughout the year, with highest values in January. The major contributing factors to evapotranspiration rate includes moisture availability, solar radiation, temperature and presence of vegetation (Brooks et al. 2013) at which during this period these factors are at the peak. It should be noticed that the evapotranspiration in North Rumphi sub-basin

is higher compared to that of South Rukuru sub-basin mainly because of the dense vegetation that North Rumphi sub-basin has. A dense vegetation leads to high moisture retention and furthermore, a greater amount of water gets intercepted and transpired which in turn increases the evapotranspiration values (Brooks et al. 2013).

It also clearly shows that even during the dry season when there is low water available in the sub-basins there is high evapotranspiration, which means there is continuous decrease in the total water storage of the system. In summary, the North Rumphi sub-basin presents the highest precipitation and surface runoff, with evapotranspiration being lower compared to South Rukuru sub-basin in the hot dry seasons for the entire period.

5.3 WEAP Model Performance Evaluation

The competence of the model performance was tested using goodness-of-fit statistics from the model simulated output and measured streamflow data. According to Moriasi et al. (2007), for calibration period, the model performance was achieved to simulate streamflow with NSE, R^2 , IA, PBIAS and KGE values that have shown a strong agreement between measured and simulated streamflow in the both catchments of the model. Likewise, for validation the period, there was a very good agreement between the measured and simulated flows in both catchments (Table 9). However, the PBIAS values for all the catchments and periods there were negative, which indicate that the model was slightly underestimating the streamflow values.

In general, the WEAP model was able to maintain a very good agreement in reproducing the overall streamflow characteristics. The model shows an outstanding representation of both catchments, with best fit to the North Rumphi Catchment. Thus, these results indicate a more than reasonable ability of the WEAP hydrological model to simulate the flows using remotely acquired

data. Considering the sensibility of the WEAP model, the parameters that have a stronger influence on the model outputs were almost all the soil and land use-related parameters except the preferred flow direction and deep conductivity (Table 8).

Similarly, previous studies have confirmed the capability of the WEAP hydrologic model in reproducing catchment hydrology processes in different parts of the world (Table 10). Among these, Abera Abdi & Ayenew (2021) reported that WEAP hydrologic model attained the R^2 , NSE and IA values of 0.82, 0.80, 0.95 between monthly measured and simulated streamflow in the Ketar River Basin, Ethiopia for the calibration period. During the validation period, the model showed a reasonably good agreement between the measured and simulated flows with R^2 of 0.91, NSE of 0.91 and IA of 0.98. The performance of the WEAP model simulation in The Chongwe River Catchment, Zambia showed a positive strong relationship using the R^2 of 0.97 and the NSE of 0.64 (Tena et al. 2019). Ingol-Blanco & McKinney (2013) developed the WEAP model to assess the hydrologic processes in Rio Conchos Basin, Mexico. Six gauging stations were used in the basin for model performance evaluation. Goodness of fit metrics values ranged as follows; NSE from 0.65 to 0.87; R^2 from 0.92 to 0.97 during calibration and for validation NSE from 0.60 to 0.88 and R^2 from 0.92 to 0.97 between measured and simulated flows respectively.

In the Indus River Basin, The WEAP model calibration and validation statistics of Nash–Sutcliffe efficiency and coefficient of determination values were 0.85, 0.86 and 0.89, 0.87, respectively (Asghar et al. 2019). From five gauging stations in the Western Algeria watersheds, ranging values of NSE 0.23 to 0.88, and R^2 0.74 to 1.0 were achieved between measured and simulated average monthly flows (Hamlat et al. 2013). Finally, to characterize and evaluate the water supply and downstream demands of the hydrological basins associated with the Conguillío National Park, located in the Andes Mountains in central-southern Chile, the WEAP hydrologic model was

developed by Fernández-Alberti et al. (2021) and the model has demonstrated the capability from four gauging stations with values ranging from NSE (0.80 – 0.86, 0.77 – 0.84); R^2 (0.86 – 0.94, 0.88 – 0.96) and BIAS (0.04 – 0.19, 0.08 – 0.23) at calibration and validation periods between measured and simulated flows data respectively.

Table 10: Comparison of WEAP hydrologic model performance evaluation with other similar studies elsewhere

Country of Origin	Goodness of Fit Metrics		Modelling Periods	Reference
	R^2	NSE		
Malawi	0.93	0.92	Calibration	Present Study
	0.95	0.94	Validation	
Ethiopia	0.82	0.80	Calibration	(Abera Abdi & Ayenew, 2021)
	0.91	0.91	Validation	
Zambia	0.97	0.64		(Tena et al., 2019)
Pakistan	0.96	0.85	Calibration	(Asghar et al. 2019)
	0.87	0.89	Validation	
Algeria	0.74 – 1.0	0.23 – 0.88		(Hamlat et al. 2013)
USA	0.92	0.91	Calibration	(Mehta et al. 2013)
	0.83	0.78	Validation	
Mexico	0.92 – 0.97	0.65 – 0.87	Calibration	Ingol-Blanco & McKinney (2013)
	0.92 – 0.97	0.60 – 0.88	Validation	
Chile	0.86 – 0.94	0.80 – 0.86	Calibration	(Fernández-Alberti et al. 2021)
	0.88 – 0.96	0.77 – 0.84	Validation	

CHAPTER 6: CONCLUSION AND RECOMMENDATIONS

The purpose of this research study was to assess the water availability using remote sensing in South Rukuru and North Rumphu river basin. This final chapter presents a review of the work presented in this thesis together with recommendations.

6.1 Conclusion

The availability of complete quality streamflow data of sufficiently long duration is a requirement when developing a hydrological model. It is required that before using a hydrological data, it has to undergo quality checking, to check if it is fit for use. For this study, hydrological data availability assessment was performed which was adapted from Crochemore et al. (2020). The results show that each hydrometric station had gaps or missing data. The overall data availability for station 7F1 showed to be slightly higher compared to 7G14 which had the lowest overall data availability and station 7G18 has the best overall availability with station 7H3 coming second best. Station 7G14 had the longest data availability without gaps while station 7F1 had the lowest longest data availability. Lastly, station 7G14 appeared to have the highest continuity of available data with 7F1 with the lowest continuity in data. Although there were gaps or missing data, with the threshold set, all stations under study had a good allowable overall data availability.

After the quality checking of the hydrological data, the rainfall runoff model was developed in WEAP modelling software using Global Land Data Assimilation System (GLDAS) datasets of precipitation, air temperature and wind speed. The hydrological model is able, not only to estimate water availability trends over longer periods for South Rukuru and North Rumphu river basin but also monitor nearly all components of the hydrological cycle. Assessing water availability in a temporal context, the WEAP model hydrological response showed that months at which the

precipitation was at the peak in all the sub-basins, the surface runoff is also at peak. The lion's share of available water returns to the atmosphere via evapotranspiration in both sub-basins. However the annual stream discharge is mainly composed of interflow and base flow which contributes a major part of flow during the dry season and during the wet season the main contributor is surface runoff. Therefore, the sustainability of water availability in South Rukuru and North Rumphu river basin throughout the year is largely dependent on surface runoff in wet seasons and base flow in dry seasons. Anything that alters the hydrologic behavior of the sub-basins may impact the amount and the sustainability of water availability.

The WEAP model performance was evaluated using various commonly used goodness of fit metrics and joint plots of simulated and observed streamflow. The observed streamflow data was also used during the calibration and validation of the remote sensing WEAP model. The WEAP model was calibrated using a trial-and-error method over sixteen years (2000 – 2015) and validated for an independent six-year period (2016 – 2021). Model calibration was performed to estimate the soil and land use – related parameters until a good fit was observed between the measured and simulated streamflow. Model validation followed to check the adequacy of the remote sensing WEAP model. The model was accepted only when the goodness of fit metrics showed a strong agreement between the measured and simulated streamflow. The WEAP model was able to maintain a very good agreement in reproducing the overall streamflow characteristics within the sub-basins.

6.2 Recommendations

According to (GoM 2014) overall assessment of observed hydrological data here in Malawi was conducted up to February, 2010. After this period, with regards to the knowledge of the author, only a few assessments of such kind have been done. Only a single basin was used for the research study, an obvious avenue for future research on observed hydrological data availability assessments is being proposed to be conducted on other river basins within Malawi. In addition, the study only focused on four hydrometric stations, it being recommended that assessments for the other hydrometric stations within the Water Resources Area 7 should be conducted.

The WEAP model developed in this thesis has shown to be a reliable hydrological tool for water resources management and decision making for the South Rukuru and North Rumphu river system especially in applications such as water allocation planning. This could help all the interested parties in the system in taking necessary steps to balance water supplies with demands, particularly to manage the natural variability of water availability, and to avoid frequent or unexpected water shortfalls.

Investigations in this thesis are based in the Northern Region of Malawi and only a single basin. It is therefore necessary that future work has to be done on the usage of GLDAS datasets to assess water availability in other river basins in Northern Region, Central and Southern of Malawi. In addition, the model requires the use of up to date GLDAS climate data inputs, it is recommended that before using the model, the GLDAS climate data should be up to date.

In summary, based on the remote sensing rainfall-runoff WEAP model developed, the Global Land Data Assimilation System (GLDAS) datasets can be used to assess water availability in a river basin when the availability of observed hydrological data is insufficient, inconsistent or fragment.

CHAPTER 7: REFERENCES

- Abera Abdi, D. & Ayenew, T., 2021, 'Evaluation of the WEAP model in simulating subbasin hydrology in the Central Rift Valley basin, Ethiopia', *Ecological Processes*, 10(1), 41.
- Anees, M.T., Abdullah, K., Nawawi, M.N.M., Ab Rahman, N.N.N., Piah, Abd.R.Mt., Zakaria, N.A., Syakir, M.I. & Mohd. Omar, A.K., 2016, 'Numerical modeling techniques for flood analysis', *Journal of African Earth Sciences*, 124, 478–486.
- Armanios, D.E. & Fisher, J.B., 2014, 'Measuring water availability with limited ground data: assessing the feasibility of an entirely remote-sensing-based hydrologic budget of the Rufiji Basin, Tanzania, using TRMM, GRACE, MODIS, SRB, and AIRS: Measuring Water Availability With Limited Ground Data', *Hydrological Processes*, 28(3), 853–867.
- Arriagada, P., Karelovic, B. & Link, O., 2021, 'Automatic gap-filling of daily streamflow time series in data-scarce regions using a machine learning algorithm', *Journal of Hydrology*, 598, 126454.
- Asghar, A., Iqbal, J., Amin, A. & Ribbe, L., 2019, 'Integrated hydrological modeling for assessment of water demand and supply under socio-economic and IPCC climate change scenarios using WEAP in Central Indus Basin', *Journal of Water Supply: Research and Technology-Aqua*, 68(2), 136–148.
- Bhave, A.G., Bulcock, L., Dessai, S., Conway, D., Jewitt, G., Dougill, A.J., Kolusu, S.R. & Mkwambisi, D., 2020, 'Lake Malawi's threshold behaviour: A stakeholder-informed model to simulate sensitivity to climate change', *Journal of Hydrology*, 584, 124671.

Bi, H., Ma, J., Zheng, W. & Zeng, J., 2016, 'Comparison of soil moisture in GLDAS model simulations and in situ observations over the Tibetan Plateau: EVALUATE GLDAS SOIL MOISTURE OVER TP', *Journal of Geophysical Research: Atmospheres*, 121(6), 2658–2678.

Brooks, K.N., Ffolliott, P.F. & Magner, J.A., 2013, *Hydrology and the management of watersheds*, 4th ed, Wiley-Blackwell, Ames, Iowa.

Condom, T., Escobar, M., Purkey, D., Pouget, J.C., Suarez, W., Ramos, C., Apaestegui, J., Zapata, M., Gomez, J. & Vergara, W., 2011, *Modelling the hydrologic role of glaciers within a Water Evaluation and Planning System (WEAP): a case study in the Rio Santa watershed (Peru)*, Water Resources Management/Modelling approaches.

Costa, J.E., Cheng, R.T., Haeni, F.P., Melcher, N., Spicer, K.R., Hayes, E., Plant, W., Hayes, K., Teague, C. & Barrick, D., 2006, 'Use of radars to monitor stream discharge by noncontact methods: NONCONTACT STREAMFLOW', *Water Resources Research*, 42(7).

Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A. & Arheimer, B., 2020, 'Lessons learnt from checking the quality of openly accessible river flow data worldwide', *Hydrological Sciences Journal*, 65(5), 699–711.

Dobriyal, P., Badola, R., Tuboi, C. & Hussain, S.A., 2017, 'A review of methods for monitoring streamflow for sustainable water resource management', *Applied Water Science*, 7(6), 2617–2628.

Donigian, A.S., 2002, 'Watershed Model Calibration And Validation: The Hspf Experience', *Proceedings of the Water Environment Federation*, 2002(8), 44–73.

EC, 2015, *Guidance document on the application of water balances for supporting the implementation of the WFD: final : version 6.1-18/05/2015.*, Publications Office, Luxembourg.

Emran, N.A., 2015, 'Data Completeness Measures', in A. Abraham, A.K. Muda & Y.-H. Choo (eds.), *Pattern Analysis, Intelligent Security and the Internet of Things*, Advances in Intelligent Systems and Computing., vol. 355, pp. 117–130, Springer International Publishing, Cham.

FAO and UN Water, 2021, *Progress on the level of water stress*, FAO and UN Water.

FAO & IHE Delft, 2020, *Water Accounting In The Awash River Basin: Wapor Water Accounting Reports.*, Food & Agriculture Org, S.L.

Fernández-Alberti, S., Abarca-del-Río, R., Bornhardt, C. & Ávila, A., 2021, 'Development and Validation of a Model to Evaluate the Water Resources of a Natural Protected Area as a Provider of Ecosystem Services in a Mountain Basin in Southern Chile', *Frontiers in Earth Science*, 8, 539905.

García, L., Rodríguez, J.D., Wijnen, M. & Pakulski, I., 2016, *Earth Observation for Water Resources Management: Current Use and Future Opportunities for the Water Sector*, Washington, DC: World Bank.

GoM, 2005, *National Spatial Data Center*, vol. 1, Malawi Government, Ministry Lands Housing Physical Planning and Surveys, Lilongwe, Malawi.

GoM, 2010, *Malawi state of environment and outlook report: environment for sustainable economic growth*, Malawi Government, Ministry of Natural Resources, Energy and Environment, Lilongwe, Malawi.

GoM, 2011, *National Water Development Program; Situation And Needs Assessment Report*, Malawi, Lilongwe.

GoM, 2012a, *Malawi Sector Performance Report; Irrigation, Water and Sanitation*, Lilongwe, Malawi.

GoM, 2012b, *Malawi Water Sector Investment Plan Volume II*, Malawi Government, Ministry of Agriculture, Irrigation and Water Development, Lilongwe, Malawi.

GoM, 2014, *National Water Resources Master Plan*, vol. I, Malawi Government, Ministry of Agriculture, Irrigation and Water Development, Lilongwe, Malawi.

GoM, 2020, *Malawi Statistical YearBook*, Malawi Government, National Statistical Office, Malawi, Zomba.

Greve, P., Kahil, T., Mochizuki, J., Schinko, T., Satoh, Y., Burek, P., Fischer, G., Tramberend, S., Burtscher, R., Langan, S. & Wada, Y., 2018, 'Global assessment of water challenges under uncertainty in water scarcity projections', *Nature Sustainability*, 1(9), 486–494.

Gupta, H.V., Kling, H., Yilmaz, K.K. & Martinez, G.F., 2009, 'Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling', *Journal of Hydrology*, 377(1–2), 80–91.

Gupta, H.V., Sorooshian, S. & Yapo, P.O., 1999, 'Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration', *Journal of Hydrologic Engineering*, 4(2), 135–143.

Guug, S.S., Abdul-Ganiyu, S. & Kasei, R.A., 2020, 'Application of SWAT hydrological model for assessing water availability at the Sherigu catchment of Ghana and Southern Burkina Faso', *HydroResearch*, 3, 124–133.

Hamlat, A., Errih, M. & Guidoum, A., 2013, 'Simulation of water resources management scenarios in western Algeria watersheds using WEAP model', *Arabian Journal of Geosciences*, 6(7), 2225–2236.

Hamzah, F.B., Mohd Hamzah, F., Mohd Razali, S.F., Jaafar, O. & Abdul Jamil, N., 2020, 'Imputation methods for recovering streamflow observation: A methodological review', F. Li (ed.), *Cogent Environmental Science*, 6(1), 1745133.

Houghton-Carr, H. & Fry, M., 2006, 'The decline of hydrological data collection for development of integrated water resource management tools in Southern Africa', 5.

Hulsman, P., Winsemius, H.C., Michailovsky, C.I., Savenije, H.H.G. & Hrachowitz, M., 2020, 'Using altimetry observations combined with GRACE to select parameter sets of a hydrological model in a data-scarce region', *Hydrology and Earth System Sciences*, 24(6), 3331–3359.

Ingol-Blanco, E. & McKinney, D.C., 2013, *Hydrologic Modeling for Assessing Climate Change Impacts on the Water Resources of the Rio Conchos Basin*, *World Environmental and Water Resources Congress 2009*, 1–10, American Society of Civil Engineers, Kansas City, Missouri, United States.

Jayantari, M.W., Wardoyo, W. & Maulana, M.A., 2019, 'Satellite data use in the WEAP Model as an evaluation of Water Availability in Unda River Basin', *IOP Conference Series: Earth and Environmental Science*, 389(1), 012027.

Jazim, A.A., 2006, 'A Monthly Six-parameter Water Balance Model and Its Application at Arid and Semiarid Low Yielding Catchments', *Journal of King Saud University - Engineering Sciences*, 19(1), 65–81.

Ji, L., Senay, G.B. & Verdin, J.P., 2015, 'Evaluation of the Global Land Data Assimilation System (GLDAS) Air Temperature Data Products', *Journal of Hydrometeorology*, 16(6), 2463–2480.

Karimi, P. & Bastiaanssen, W.G.M., 2014, *Spatial evapotranspiration, rainfall and land use data in water accounting – Part 1: Review of the accuracy of the remote sensing data*, Water Resources Management/Remote Sensing and GIS.

Khasmakhi, H.P., Vazifedoust, M., Marofi, S. & Tizro, A.T., 2020, 'Simulation of river discharge in ungauged catchments by forcing GLDAS products to a hydrological model (a case study: Polroud basin, Iran)', *Water Supply*, 20(1), 277–286.

Krause, P., Boyle, D.P. & Bäse, F., 2005, 'Comparison of different efficiency criteria for hydrological model assessment', *Advances in Geosciences*, 5, 89–97.

Kumambala, P.G., 2010, *Sustainability of Water Resources Development for Malawi with Particular Emphasis on North and Central Malawi* – PhD thesis, University of Glasgow, United Kingdom .

Kumar, D.N. & Reshmidevi, T.V., 2013, 'Remote Sensing Applications in Water Resources', 93, 27.

Libanda, B., Zheng, M. & Banda, N., 2017, 'Variability of extreme wet events over Malawi', *Geographica Pannonica*, 21(4), 212–223.

Lorenz, D.L. & Ziegeweid, J.R., 2016, *Scientific Investigations Report*.

Madsen, H., 2000, 'Automatic calibration of a conceptual rainfall–runoff model using multiple objectives', *Journal of Hydrology*, 235(3–4), 276–288.

Madsen, H., 2003, 'Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives', *Advances in Water Resources*, 26(2), 205–216.

Madsen, H., Wilson, G. & Ammentorp, H.C., 2002, 'Comparison of different automated strategies for calibration of rainfall-runoff models', *Journal of Hydrology*, 261(1–4), 48–59.

Moreira, A.A., Ruhoff, A.L., Roberti, D.R., Souza, V. de A., Rocha, H.R. da & Paiva, R.C.D. de, 2019, 'Assessment of terrestrial water balance using remote sensing data in South America', *Journal of Hydrology*, 575, 131–147.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner R. L., Harmel R. D., & Veith, T. L., 2007, 'Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations', *Transactions of the ASABE*, 50(3), 885–900.

Muthuwatta, L.P., Ahmad, M.-D., Bos, M.G. & Rientjes, T.H.M., 2010, 'Assessment of Water Availability and Consumption in the Karkheh River Basin, Iran—Using Remote Sensing and Geo-statistics', *Water Resources Management*, 24(3), 459–484.

Nash, J.E. & Sutcliffe, J.V., 1970, 'River flow forecasting through conceptual models part I — A discussion of principles', *Journal of Hydrology*, 10(3), 282–290.

Oliveira, P.T.S., Nearing, M.A., Moran, M.S., Goodrich, D.C., Wendland, E. & Gupta, H.V., 2014, 'Trends in water balance components across the Brazilian Cerrado', *Water Resources Research*, 50(9), 7100–7114.

Salehie, O., Ismail, T. bin, Shahid, S., Hamed, M.M., Chinnasamy, P. & Wang, X., 2022, 'Assessment of Water Resources Availability in Amu Darya River Basin Using GRACE Data', *Water*, 14(4), 533.

Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R. & Hauck, L.M., 2001, 'VALIDATION OF THE SWAT MODEL ON A LARGE RWER BASIN WITH POINT AND NONPOINT SOURCES', *Journal of the American Water Resources Association*, 37(5), 1169–1188.

SEI, 2005, *WEAP (Water Evaluation and Planning System) Tutorial*, Stockholm Environment Institute, Stockholm, Sweden.

SEI, 2022a, *WEAP: Soil Moisture Method: Two-bucket method*.

SEI, 2022b, *Automatic Catchment Delineation*.

Sheffield, J., Wood, E.F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A. & Verbist, K., 2018, 'Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions', *Water Resources Research*, 54(12), 9724–9758.

Solomatine, D.P. & Wagener, T., 2011, 'Hydrological Modeling', *Treatise on Water Science*, pp. 435–457, Elsevier.

Speed, R., Yuanyuan, L., Quesne, T.L., Pegram, G. & Zhiwei, Z., 2013, *Basin water allocation planning: principles, procedures, and approaches for basin allocation planning*, Asian Development Bank, GIWP, UNESCO, and WWF-UK, Metro Manila, Philippines.

Stisen, S., Jensen, K.H., Sandholt, I. & Grimes, D.I.F., 2008, 'A remote sensing driven distributed hydrological model of the Senegal River basin', *Journal of Hydrology*, 354(1–4), 131–148.

Stisen, S. & Sandholt, I., 2010, 'Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling', *Hydrological Processes*, 24(7), 879–891.

Tadeyo, E., Chen, D., Ayugi, B. & Yao, C., 2020, 'Characterization of Spatio-Temporal Trends and Periodicity of Precipitation over Malawi during 1979–2015', *Atmosphere*, 11(9), 891.

Tena, T.M., Mwaanga, P. & Nguvulu, A., 2019, 'Hydrological Modelling and Water Resources Assessment of Chongwe River Catchment using WEAP Model', *Water*, 11(4), 839.

Van Liew, M.W., Veith, T.L., Bosch, D.D. & Arnold, J.G., 2007, 'Suitability of SWAT for the Conservation Effects Assessment Project: Comparison on USDA Agricultural Research Service Watersheds', *Journal of Hydrologic Engineering*, 12(2), 173–189.

Wada, Y. & Bierkens, M.F.P., 2014, 'Sustainability of global water use: past reconstruction and future projections', *Environmental Research Letters*, 9(10), 104003.

Wada, Y., Flörke, M., Hanasaki, N., Eisner, S., Fischer, G., Tramberend, S., Satoh, Y., Vliet, M.T.H. van, Yillia, P., Ringler, C., Burek, P. & Wiberg, D., 2016, 'Modeling global water use for the 21st century: the Water Futures and Solutions (WFaS) initiative and its approaches', *Geoscientific Model Development*, 9(1), 175–222.

Willmott, C.J., 1981, 'On The Validation Of Models', *Physical Geography*, 2(2), 184–194.

WMO, 2010, *Manual on stream gauging*, vol. 1, World Meteorological Organization, Geneva, Switzerland.

WWAP (ed.), 2014, *United Nations World Water Development Report 2014. Water and Energy*, UNESCO, Paris.

WWAP (ed.), 2016, *United Nations World Water Development Report 2016. Water and Jobs*, UNESCO, Paris.

WWAP (ed.), 2018, *United Nations World Water Development Report 2018. Nature-Based Solutions for Water*, UNESCO, Paris.

WWAP (ed.), 2020, *United Nations World Water Development Report 2020. Water and Climate Change*, UNESCO, Paris.

Yates, D., Purkey, D., Sieber, J., Huber-Lee, A., Galbraith, H., West, J., Herrod-Julius, S., Young, C., Joyce, B. & Rayej, M., 2009, 'Climate Driven Water Resources Model of the Sacramento Basin, California', *Journal of Water Resources Planning and Management*, 135(5), 303–313.

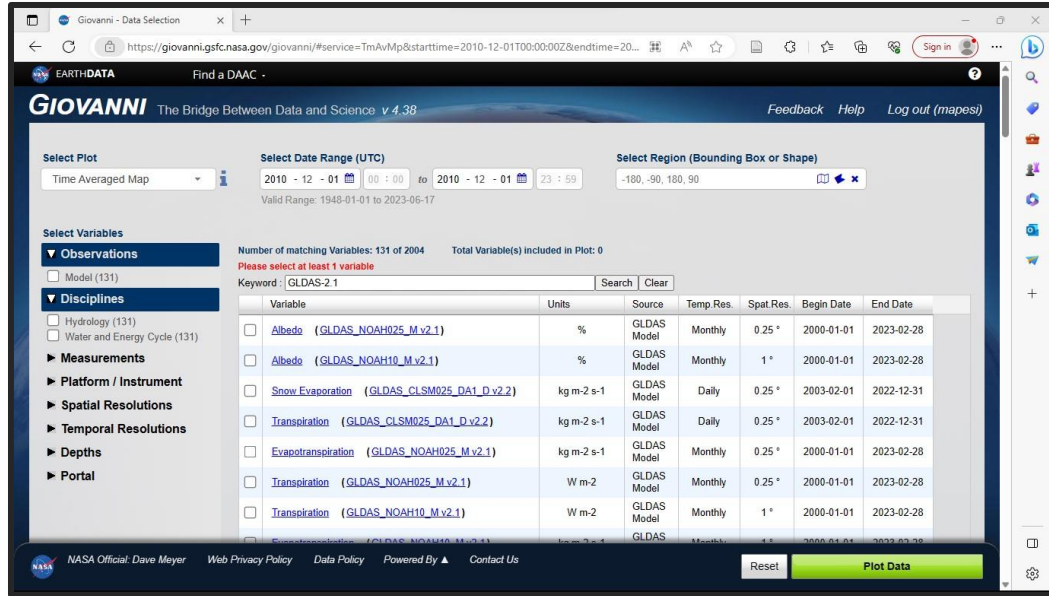
Yaykiran, S., Cuceloglu, G. & Ekdal, A., 2019, 'Estimation of Water Budget Components of the Sakarya River Basin by Using the WEAP-PGM Model', *Water*, 11(2), 271.

Zhang, H., Zhang, L., Li, J., An, R. & Deng, Y., 2018, 'Climate and Hydrological Change Characteristics and Applicability of GLDAS Data in the Yarlung Zangbo River Basin, China', *Water*, 10(3), 254.

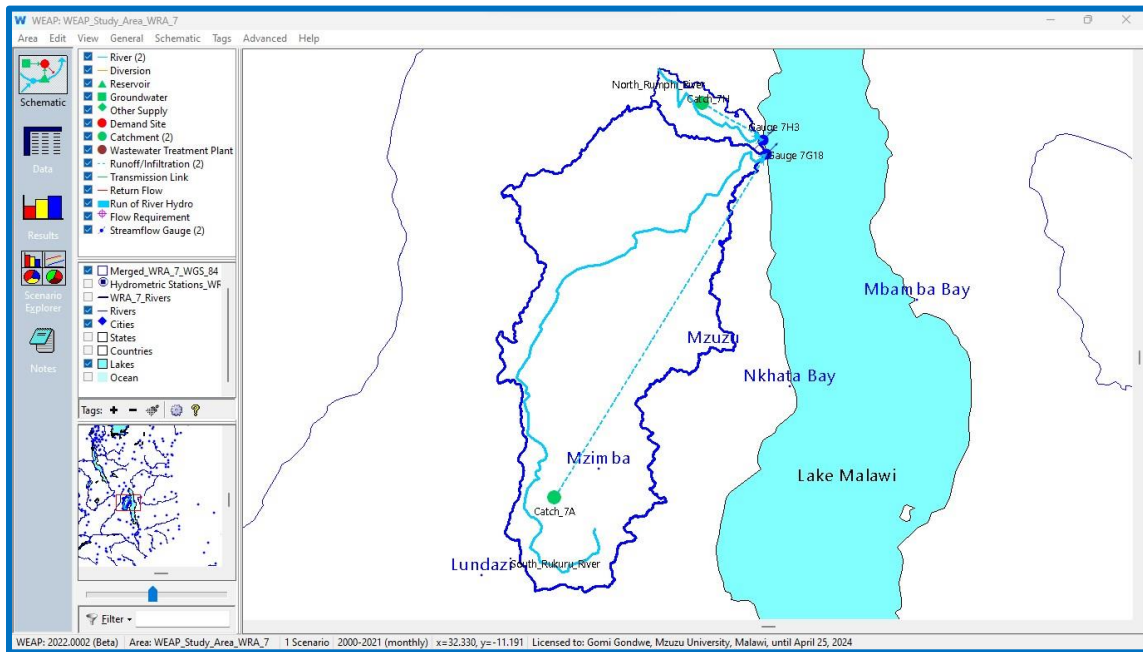
Appendix 1: Relevant Tables and Schematic Diagrams

A) Variables on the GIOVANNI Website, where GLDAS-2.1 climate data for Precipitation, Wind Speed, Specific Humidity and Temperature were Obtained;

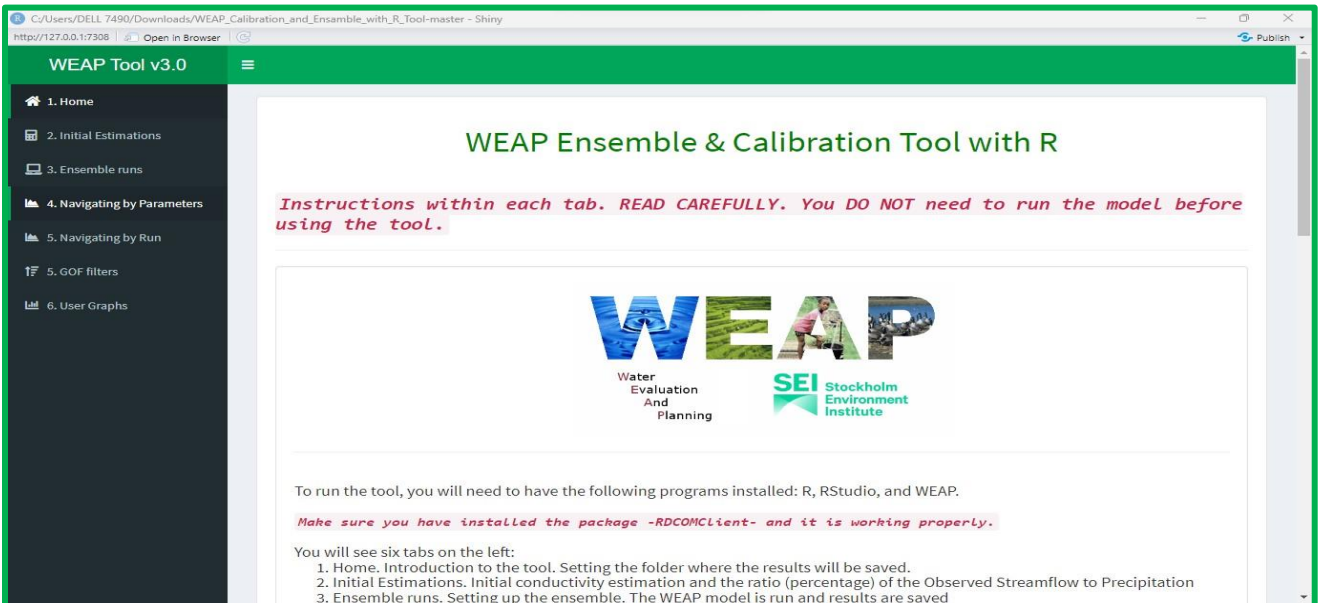
Variable	Units	Source	Temporal Resolution	Spatial Resolution	Data Begin Date	Data End Date
Specific Humidity (GLDAS_NOAH025_Mv2.1)	kg kg-1	GLDAS Model	Monthly	0.25°	2000 – 01 - 01	Present
Near Surface Air Temperature (GLDAS_NOAH025_3H_v2.1)	K	GLDAS Model	3 – hourly	0.25°	2000 – 01 - 01	Present
Total Precipitation Rate (GLDAS_NOAH025_Mv2.1)	kg m-2 s-1	GLDAS Model	Monthly	0.25°	2000 – 01 - 01	Present
Near Surface Wind Speed (GLDAS_NOAH025_Mv2.1)	m s-1	GLDAS Model	Monthly	0.25°	2000 – 01 – 01	Present



B) Schematic Diagram of the GIOVANNI Website, where GLDAS climate data for Precipitation, Wind Speed, Relative Humidity and Temperature were Obtained Accessed; [Giovanni - Data Selection \(nasa.gov\)](https://giovanni.gsfc.nasa.gov)



C) Schematic Diagram of the Remote Sensing Rainfall-Runoff (Soil Moisture Method) WEAP Model



D) Schematic Diagram of Shiny App in R Studio of the WEAP Ensemble & Calibration Tool with R. Developed by Angelica Moncada (SEI-LAC Water Group member) Used in Calibrating and Validating the WEAP Model.

Accessed; [GitHub - ammoncadaa/WEAP Calibration and Ensamble with R Tool](https://github.com/ammoncadaa/WEAP_Calibration_and_Ensamble_with_R_Tool)

Appendix 2: Stream Flow Data

Time	Gauge 7G18		Gauge 7H3	
	Interpolated Observed Flow (m ³)	Simulated Flow (m ³)	Interpolated Observed Flow (m ³)	Simulated Flow (m ³)
Jan-00	311,388,106	471,856,000	173,158,560	91,715,100
Feb-00	776,154,355	571,146,000	412,449,408	336,391,000
Mar-00	1,252,213,603	994,840,000	349,906,176	325,772,000
Apr-00	545,958,144	508,730,000	151,320,960	83,842,200
May-00	170,841,744	107,482,000	25,685,856	19,806,800
Jun-00	88,400,160	63,021,100	13,089,600	19,684,900
Jul-00	52,962,682	51,259,700	7,740,576	9,941,770
Aug-00	34,647,782	42,821,800	6,133,536	21,452,800
Sep-00	4,484,160	36,662,600	181,440	4,649,230
Oct-00	2,715,898	54,777,400	1,526,688	9,698,620
Nov-00	264,676,896	292,691,000	176,074,560	108,768,000
Dec-00	601,054,387	594,668,000	243,895,104	172,499,000
Jan-01	1,140,912,640	1,258,259,968	621,951,232	593,564,992
Feb-01	1,167,474,432	1,149,260,032	621,323,136	488,104,992
Mar-01	1,136,956,672	1,647,180,032	315,917,280	189,567,008
Apr-01	755,650,944	424,807,008	150,258,240	87,463,800
May-01	228,681,792	178,106,000	49,068,288	71,968,400
Jun-01	104,312,448	58,754,800	14,929,920	10,931,000
Jul-01	62,998,648	51,003,700	10,017,216	12,317,400
Aug-01	37,363,680	46,079,500	7,097,760	12,012,100
Sep-01	4,999,968	43,410,500	1,244,160	7,138,390
Oct-01	6,069,254	48,197,900	2,785,536	14,766,400
Nov-01	11,321,856	57,788,200	4,976,640	12,331,100
Dec-01	398,588,768	306,048,992	117,313,920	209,587,008
Jan-02	754,339,200	849,080,000	502,119,648	562,268,992
Feb-02	1,514,707,072	1,082,610,048	526,369,536	490,257,984
Mar-02	989,242,944	1,571,069,952	373,235,040	298,596,992
Apr-02	997,650,432	615,315,968	118,532,160	133,084,000
May-02	221,771,520	60,407,900	35,783,424	33,419,000
Jun-02	101,901,888	58,049,300	16,770,240	23,343,900
Jul-02	54,617,932	44,777,100	9,561,888	8,159,730
Aug-02	28,369,612	57,218,900	6,240,672	12,957,400
Sep-02	13,797,216	63,181,800	4,250,880	20,240,000
Oct-02	9,122,630	40,861,900	1,848,096	15,028,100
Nov-02	90,958,464	154,440,992	16,407,360	40,682,200

Dec-02	339,896,992	578,881,984	104,618,304	118,358,000
Jan-03	1,278,427,136	975,742,016	659,957,760	654,972,992
Feb-03	1,327,078,784	1,157,560,064	482,896,512	487,189,984
Mar-03	943,439,616	1,476,310,016	244,618,272	287,975,008
Apr-03	952,985,088	547,347,968	259,044,480	160,014,000
May-03	242,113,968	123,061,000	38,113,632	54,906,700
Jun-03	98,400,096	77,244,600	16,329,600	24,379,000
Jul-03	65,789,540	69,088,600	10,954,656	19,735,100
Aug-03	29,349,908	45,179,200	5,437,152	8,602,290
Sep-03	7,558,272	70,225,504	1,969,920	11,291,300
Oct-03	3,013,200	44,799,200	1,928,448	11,447,800
Nov-03	12,029,472	79,068,704	5,184,000	15,157,500
Dec-03	304,274,272	376,377,984	96,663,456	112,712,000
Jan-04	727,740,032	744,881,984	314,953,056	225,778,000
Feb-04	851,768,896	719,238,976	363,194,496	326,900,992
Mar-04	833,777,856	781,619,008	270,679,104	217,872,000
Apr-04	667,629,184	845,094,976	182,373,120	135,735,008
May-04	213,155,104	85,891,696	33,881,760	35,233,800
Jun-04	106,282,368	52,790,700	17,003,520	24,946,200
Jul-04	63,770,024	42,355,700	9,454,752	8,721,770
Aug-04	33,030,028	39,965,200	5,544,288	9,178,170
Sep-04	14,484,096	38,514,900	10,160,640	28,611,600
Oct-04	23,028,884	54,669,000	6,401,376	26,445,100
Nov-04	72,301,248	182,352,992	16,381,440	40,695,200
Dec-04	645,622,976	1,104,870,016	218,798,496	255,964,000
Jan-05	907,519,616	1,197,410,048	406,500,768	294,886,016
Feb-05	881,135,552	1,222,950,016	376,113,024	276,409,984
Mar-05	987,068,096	920,865,984	440,248,608	207,776,000
Apr-05	361,428,480	329,788,000	78,719,040	69,025,000
May-05	188,280,800	114,694,000	34,524,576	52,085,300
Jun-05	94,874,976	73,451,600	15,448,320	23,461,700
Jul-05	70,401,744	51,530,700	10,874,304	14,071,700
Aug-05	40,039,400	41,664,200	6,053,184	5,779,680
Sep-05	25,487,136	46,721,700	4,354,560	9,412,370
Oct-05	23,594,026	37,046,600	4,231,872	9,658,810
Nov-05	33,283,872	70,808,496	9,927,360	17,410,100
Dec-05	79,125,296	223,620,000	21,373,632	26,125,100
Jan-06	497,003,904	587,430,016	112,305,312	90,831,800
Feb-06	672,530,368	794,334,976	196,293,888	252,208,992
Mar-06	856,809,472	1,360,979,968	461,461,536	268,991,008
Apr-06	955,491,520	1,151,079,936	188,801,280	104,185,000

May-06	304,713,536	262,324,992	58,603,392	85,464,000
Jun-06	124,861,824	61,773,200	16,148,160	16,226,500
Jul-06	80,330,576	48,374,100	12,347,424	23,459,100
Aug-06	49,978,944	46,467,000	8,088,768	10,777,800
Sep-06	31,334,688	41,720,800	5,054,400	9,053,780
Oct-06	25,693,892	40,764,300	5,356,800	13,586,400
Nov-06	76,808,736	182,732,000	40,538,880	41,787,500
Dec-06	1,070,827,008	827,822,976	331,719,840	332,364,000
Jan-07	1,890,936,960	1,740,220,032	1,010,587,136	891,808,000
Feb-07	1,410,304,128	1,672,390,016	718,816,896	548,224,000
Mar-07	1,013,862,784	1,093,880,064	250,457,184	205,210,000
Apr-07	562,274,816	452,460,000	189,501,120	116,420,000
May-07	233,216,320	136,156,992	44,247,168	62,120,700
Jun-07	106,155,360	59,406,500	16,899,840	12,596,600
Jul-07	70,452,632	53,413,800	11,409,984	13,831,600
Aug-07	45,519,408	51,650,200	7,606,656	11,438,600
Sep-07	29,263,680	45,044,400	5,365,440	10,267,100
Oct-07	54,883,096	91,602,800	10,204,704	32,704,300
Nov-07	40,020,480	102,824,000	7,024,320	19,268,200
Dec-07	422,903,296	543,115,008	218,048,544	235,144,992
Jan-08	1,057,485,888	1,505,090,048	384,778,944	324,708,992
Feb-08	1,451,737,728	1,483,849,984	654,296,832	553,505,984
Mar-08	1,067,160,256	1,232,060,032	235,404,576	191,590,000
Apr-08	394,893,792	361,344,992	63,167,040	60,611,000
May-08	168,682,960	73,265,400	27,587,520	46,037,600
Jun-08	90,204,192	52,527,900	14,074,560	13,945,400
Jul-08	59,859,560	52,201,900	9,508,320	15,841,400
Aug-08	41,520,556	49,381,200	6,588,864	10,697,500
Sep-08	23,789,376	41,193,700	3,862,080	5,436,610
Oct-08	20,808,490	63,155,800	4,124,736	12,437,900
Nov-08	73,288,800	197,372,000	17,340,480	35,003,600
Dec-08	436,188,160	505,864,000	148,249,440	129,133,000
Jan-09	846,240,512	823,486,976	278,366,112	246,486,000
Feb-09	1,249,816,832	1,327,260,032	289,263,744	230,756,000
Mar-09	1,295,689,344	1,606,210,048	275,848,416	213,758,000
Apr-09	864,872,640	693,502,976	125,660,160	81,547,104
May-09	192,539,456	104,667,000	28,792,800	28,474,200
Jun-09	81,108,864	51,214,800	11,430,720	7,864,790
Jul-09	54,671,500	57,713,300	7,472,736	11,282,300
Aug-09	34,235,308	47,908,500	5,142,528	8,226,950
Sep-09	20,590,848	43,487,300	3,110,400	4,780,760

Oct-09	21,836,996	41,293,400	3,642,624	14,296,500
Nov-09	203,622,336	229,138,000	74,312,640	51,994,800
Dec-09	316,948,480	582,374,016	70,522,272	94,400,200
Jan-10	1,405,364,480	863,019,008	639,414,400	348,369,984
Feb-10	1,673,445,376	1,270,850,048	374,492,160	507,353,984
Mar-10	1,371,488,128	1,371,879,936	287,686,944	209,279,008
Apr-10	420,386,112	292,684,992	65,370,240	80,589,104
May-10	267,735,536	87,878,000	44,381,088	37,664,900
Jun-10	110,154,816	53,745,300	15,033,600	19,301,100
Jul-10	69,839,280	51,276,600	9,695,808	13,585,200
Aug-10	44,959,624	51,391,100	6,401,376	9,794,540
Sep-10	25,839,648	40,649,200	3,991,680	5,961,440
Oct-10	16,812,316	37,971,300	2,865,888	4,859,530
Nov-10	25,748,928	86,379,000	8,424,000	15,658,100
Dec-10	299,474,592	340,374,016	114,930,144	81,241,696
Jan-11	736,016,256	652,353,024	246,278,880	214,316,992
Feb-11	750,443,072	1,286,610,048	360,509,184	314,463,008
Mar-11	939,411,328	1,189,869,952	204,415,488	197,568,000
Apr-11	799,357,248	786,382,976	151,372,800	98,379,504
May-11	231,831,584	71,839,000	31,203,360	22,656,400
Jun-11	99,846,432	49,648,800	14,152,320	10,921,800
Jul-11	59,781,888	56,980,400	8,758,368	9,716,150
Aug-11	37,618,128	43,115,600	5,838,912	7,508,650
Sep-11	23,610,528	60,639,900	4,095,360	7,105,410
Oct-11	44,718,568	99,141,400	7,820,928	21,738,300
Nov-11	60,603,552	228,714,000	20,969,280	24,915,900
Dec-11	670,382,080	672,020,992	208,567,008	194,316,000
Jan-12	1,100,894,720	1,098,610,048	707,954,688	702,891,008
Feb-12	1,145,111,424	1,265,449,984	300,948,480	284,415,008
Mar-12	1,294,615,296	1,085,840,000	398,063,808	354,092,000
Apr-12	626,618,624	401,723,008	151,424,640	117,873,000
May-12	225,317,728	62,485,800	38,542,176	42,943,800
Jun-12	101,378,304	56,798,600	15,733,440	16,850,400
Jul-12	60,025,624	49,365,100	9,454,752	11,116,400
Aug-12	37,251,188	58,343,100	6,026,400	7,856,620
Sep-12	22,120,128	40,396,400	3,706,560	5,505,070
Oct-12	18,638,986	66,112,700	3,883,680	10,932,100
Nov-12	74,118,240	155,771,008	21,980,160	47,602,400
Dec-12	411,102,272	490,252,000	142,303,392	91,996,000
Jan-13	770,090,880	1,338,909,952	362,923,200	399,875,008
Feb-13	1,393,635,840	1,228,400,000	344,929,536	316,150,016

Mar-13	1,598,506,624	1,354,360,064	316,720,800	168,639,008
Apr-13	625,970,560	520,979,008	186,753,600	129,927,000
May-13	227,923,808	124,045,000	35,354,880	44,865,500
Jun-13	94,620,960	60,351,900	13,271,040	7,309,010
Jul-13	54,095,644	76,751,296	8,035,200	8,543,200
Aug-13	36,771,752	56,641,200	6,267,456	10,039,100
Sep-13	19,784,736	53,885,400	3,369,600	5,289,180
Oct-13	28,549,066	50,590,700	5,999,616	18,862,200
Nov-13	42,882,048	125,363,000	9,460,800	27,567,200
Dec-13	479,840,704	376,187,008	199,139,040	97,108,704
Jan-14	860,077,120	1,097,650,048	399,644,064	413,974,016
Feb-14	1,382,442,112	1,385,810,048	550,271,232	606,748,032
Mar-14	1,131,658,880	1,058,659,968	240,761,376	197,538,000
Apr-14	916,290,176	719,094,976	165,473,280	139,291,008
May-14	274,868,128	118,153,000	39,827,808	23,671,000
Jun-14	120,108,096	64,224,000	17,573,760	20,777,800
Jul-14	70,715,120	52,426,100	10,901,088	16,547,400
Aug-14	45,168,536	50,296,900	7,258,464	15,083,900
Sep-14	33,506,784	69,406,496	6,246,720	12,873,000
Oct-14	53,230,520	73,946,600	12,749,184	34,315,500
Nov-14	63,713,952	91,592,600	44,245,440	21,818,700
Dec-14	310,008,736	262,986,000	127,357,920	135,450,000
Jan-15	887,110,208	888,038,976	570,740,224	619,566,976
Feb-15	959,515,200	778,412,032	614,767,104	654,985,024
Mar-15	584,234,048	673,918,976	355,289,760	294,027,008
Apr-15	718,787,520	611,801,024	222,419,520	149,619,008
May-15	199,891,664	64,762,700	34,444,224	32,621,300
Jun-15	86,038,848	44,632,800	13,711,680	12,093,000
Jul-15	52,006,492	40,404,700	8,276,256	8,123,440
Aug-15	32,009,558	38,274,500	5,303,232	7,863,360
Sep-15	19,349,280	36,034,300	3,395,520	6,157,380
Oct-15	16,579,296	45,186,500	6,321,024	23,867,500
Nov-15	111,087,936	169,596,000	28,045,440	43,889,700
Dec-15	406,610,592	414,995,008	136,437,696	133,442,000
Jan-16	1,681,598,592	1,293,129,984	757,103,360	649,390,016
Feb-16	1,582,014,080	897,716,992	430,254,720	462,360,000
Mar-16	838,076,736	1,065,009,984	297,382,752	302,011,008
Apr-16	1,165,300,992	837,942,976	253,834,560	194,938,000
May-16	276,876,928	60,934,500	44,113,248	39,801,400
Jun-16	103,247,136	61,529,400	16,485,120	14,274,700
Jul-16	60,047,048	51,820,000	9,508,320	9,577,520

Aug-16	37,947,572	42,081,400	6,294,240	9,561,930
Sep-16	21,796,128	39,679,400	3,706,560	5,293,960
Oct-16	16,978,378	39,251,800	3,348,000	10,021,200
Nov-16	17,768,160	105,315,000	3,473,280	11,830,700
Dec-16	238,176,720	283,593,984	65,379,744	57,848,200
Jan-17	506,126,528	778,025,984	209,504,448	200,014,000
Feb-17	796,419,968	1,491,949,952	283,844,736	316,004,992
Mar-17	1,613,157,504	1,697,430,016	429,401,088	378,398,016
Apr-17	983,651,072	823,372,992	146,136,960	137,340,000
May-17	358,241,344	96,186,896	46,523,808	46,043,200
Jun-17	137,137,536	53,460,600	18,040,320	10,792,900
Jul-17	75,043,408	49,625,200	10,151,136	7,867,120
Aug-17	50,324,456	56,519,800	7,606,656	16,467,000
Sep-17	26,702,784	45,019,900	3,991,680	7,078,240
Oct-17	26,459,914	46,724,500	6,240,672	21,827,300
Nov-17	99,239,904	181,976,992	51,528,960	65,346,900
Dec-17	583,559,104	713,638,976	146,320,992	147,632,992
Jan-18	1,369,830,144	1,125,830,016	435,963,168	412,584,000
Feb-18	875,747,968	1,518,610,048	264,757,248	248,039,008
Mar-18	1,772,632,064	1,150,770,048	684,866,880	430,127,008
Apr-18	982,355,072	808,777,984	285,897,600	135,320,000
May-18	267,355,216	92,510,704	32,033,664	24,205,000
Jun-18	107,736,480	54,494,000	12,441,600	7,574,540
Jul-18	71,652,560	78,527,000	9,026,208	20,068,500
Aug-18	37,781,512	46,355,500	4,794,336	5,952,760
Sep-18	24,748,416	48,368,100	3,602,880	8,056,370
Oct-18	31,473,878	55,106,400	4,606,848	19,752,800
Nov-18	65,875,680	143,072,992	26,516,160	46,994,400
Dec-18	672,157,888	757,289,984	223,164,288	242,904,000
Jan-19	909,590,016	1,255,309,952	664,939,584	594,292,992
Feb-19	984,940,992	1,918,310,016	264,418,560	251,554,000
Mar-19	1,063,131,968	1,794,429,952	374,520,672	360,895,008
Apr-19	665,467,456	812,198,976	121,124,160	95,762,896
May-19	292,526,816	187,170,000	46,148,832	66,860,000
Jun-19	120,188,448	63,009,200	16,485,120	12,569,100
Jul-19	69,263,424	51,808,000	10,044,000	8,992,520
Aug-19	41,458,952	51,282,000	6,240,672	6,742,640
Sep-19	28,644,192	48,617,700	4,924,800	11,669,300
Oct-19	45,425,664	53,315,000	7,633,440	27,905,800
Nov-19	111,943,296	201,251,008	18,895,680	48,257,500
Dec-19	964,183,808	810,304,000	321,649,056	222,398,000

Jan-20	1,929,050,624	1,251,180,032	737,042,112	771,972,992
Feb-20	1,976,716,288	1,623,609,984	516,595,968	461,830,016
Mar-20	2,160,946,432	1,141,650,048	429,481,440	253,208,000
Apr-20	524,169,792	267,054,000	79,574,400	97,535,000
May-20	182,768,656	60,800,500	24,132,384	23,077,600
Jun-20	90,331,200	56,224,600	12,286,080	11,728,200
Jul-20	61,999,604	53,301,800	8,249,472	12,919,100
Aug-20	38,571,640	48,432,400	5,276,448	7,339,360
Sep-20	23,346,144	41,410,700	3,291,840	5,728,510
Oct-20	24,354,692	61,283,200	4,446,144	13,796,600
Nov-20	25,308,288	103,217,000	7,231,680	16,236,800
Dec-20	472,745,632	240,844,992	178,997,472	105,497,000
Jan-21	1,370,644,480	652,361,024	471,639,456	499,839,008
Feb-21	1,335,700,736	810,212,992	351,340,416	350,791,008
Mar-21	1,453,701,632	874,214,976	381,404,160	360,328,992
Apr-21	833,861,952	265,715,008	124,519,680	116,741,000
May-21	315,475,328	56,827,800	36,801,216	22,251,300
Jun-21	132,925,536	82,336,896	14,644,800	16,096,600
Jul-21	81,300,152	49,550,800	9,267,264	16,156,200
Aug-21	48,682,600	67,993,696	5,812,128	9,328,890
Sep-21	31,008,096	35,489,300	3,758,400	8,742,120
Oct-21	22,353,926	41,980,800	3,267,648	13,346,000
Nov-21	25,614,144	53,342,400	5,832,000	13,447,000
Dec-21	30,665,002	74,296,600	6,187,104	19,124,700

Appendix 3: Research Ethics and Regulatory Approval and Permit by MZUNIREC



MZUZU UNIVERSITY

DIRECTORATE OF RESEARCH

Mzuzu University
Private Bag 201
Luwinga
Mzuzu 2
MALAWI
TEL: 01 320 722
FAX: 01 320 648

MZUZU UNIVERSITY RESEARCH ETHICS COMMITTEE (MZUNIREC)

Ref No: MZUNIREC/DOR/22/94

24/11/2022.

Gomi Gondwe,
Mzuzu University,
P/Bag 201,
Luwinga,
Mzuzu 2.

gondwegomi@gmail.com

Dear Mr. Gondwe,
**RESEARCH ETHICS AND REGULATORY APPROVAL AND PERMIT FOR
PROTOCOL REF NO: MZUNIREC/DOR/22/94: POTENTIAL OF USING REMOTE SENSING TO
ASSESS WATER AVAILABILITY IN SOUTH RUKURU AND NORTH RUMPHI RIVER BASIN**

Having satisfied all the relevant ethical and regulatory requirements, I am pleased to inform you that the above referred research protocol has officially been approved. You are now permitted to proceed with its implementation. Should there be any amendments to the approved protocol in the course of implementing it, you shall be required to seek approval of such amendments before implementation of the same.

This approval is valid for one year from the date of issuance of this approval. If the study goes beyond one year, an annual approval for continuation shall be required to be sought from the Mzuzu University Research Ethics Committee (MZUNIREC) in a format that is available at the Secretariat. Once the study is finalised, you are required to furnish the Committee with a final report of the study. The Committee reserves the right to carry out compliance inspection of this approved protocol at any time as may be deemed by it. As such, you are expected to properly maintain all study documents including consent forms.

Wishing you a successful implementation of your study.

Committee Address:

Secretariat, Mzuzu University Research Ethics Committee, P/Bag 201, Luwinga, Mzuzu 2; E-mail address: mzunirec@mzuni.ac.mw