DROUGHT ASSESSMENT AND FORECASTING USING INDICES AND ARTIFICIAL NEURAL NETWORKS FOR THE UPPER TANA RIVER BASIN, KENYA

RAPHAEL MULI WAMBUA

A Thesis submitted to the Graduate School in fulfillment for the requirements of the Doctor of Philosophy Degree in Agricultural Engineering of Egerton University

EGERTON UNIVERSITY

SEPTEMBER, 2016

DECLARATION AND RECOMMENDATION

DECLARATION

I declare that this thesis is my original work, and that it has not been wholly or in part presented for an award of any degree in any University known to me.

Signature: _____ Date: _____

Name: **Raphael Muli Wambua**

Reg. No: **BD11/0459/13**

RECOMMENDATION

This thesis is the candidate's original work and has been prepared with our guidance and assistance. It is presented for examination with our approval as official University Supervisors.

Signature: _____Date: _____

Name: **Prof. Dr.-Ing. Benedict M. Mutua**

Department of Agricultural Engineering

Egerton University

Signature: _____Date: _____

Name: **Dr. (Eng.) James M. Raude**

Department of Soil, Water and Environmental Engineering

Jomo Kenyatta University of Agriculture and Technology

COPYRIGHT

ACKNOWLEDGEMENT

I thank God for the Mercy, Grace, Guidance, Protection, Faithfulness, Strength, Good health and all the Blessings that has led success in education and other endeavours.

Thanks to Egerton University staff who assisted during my training. I express my appreciation to the academic staff in the Faculty of Engineering and Technology, Egerton University who have provided a very conducive environment and opportunity for my mentorship during my theoretical and research training period. I appreciate the support from the Egerton University Division of Research and Extension for offering support in the publications from my research.

Distinguished gratitude goes to my supervisors; Prof. Dr.-Ing. Benedict M. Mutua and Dr. (Eng.) James M. Raude. They offered scholastic guidance, constant encouragement, valuable suggestions and great support throughout the development of this thesis. Their approach allowed me great freedom to pursue independent work of interest. In addition, they kept my ideas focused in the right direction. They always listened to me and advised on how to get solutions to come up with a worthwhile research thesis. It is a great opportunity and fortune to work with such great supervisors, for they have immensely multiplied my determination in research conceptualization and development. I am thankful to Dr. Kamau, D. N, Prof. Onyando, J. O and Dr. Kundu, P. M. all of the Faculty of Engineering and Technology for their encouragement and advice during the proposal, research and thesis development. I wish to appreciate the support from Dennis Theuri, V.O. Odongo, and C. W. Maina for the guidance they offered in the GIS application.

I thank the African Development Bank (AfDB) through the Ministry of Education, Science and Technology (MoEST) for providing scholarship and financial support for the research project. I recognize Water Resources and Management Authority (WRMA) and Ministry of Environment and Natural Resources, Kenya Meteorological Department, for providing data that was used in this study. I would also like to appreciate all the field assistants that helped me during the data collection.

DEDICATION

This research work is dedicated to my wife Theresa Monthe, who has been a source of encouragement and inspiration during challenging times of my academic programme. I appreciate her prayers, understanding and support during my study and research period. To my children; Anthony Wambua and Daniel Mutuku, for they have always been sources of determination and happiness in the family.

To my parents Mr. Benard Wambua and Mrs. Elizabeth Mutono Wambua who have supported me through out my life; when challenged in various levels of my education, they prayed for God's blessings on my endeavours. They cared and/or encouraged me to develop good character and positive attitude anchored on persistence which is paramount for any successful journey. In addition, they also sacrificed immensely in numerous ways to ensure that I pursued education to great altitudes of success.

To my brothers; Urbanus Kioko, Stephen Ndolo and all my sisters: for continous support during my education. I also dedicate this thesis to my numerous friends and church members who have supported me throughout the process. I appreciate all what they have done, especially their prayers and advice needed to overcome challenges encountered along the research process.

In a special way, this thesis is dedicated to all my trainers of good will; the primary and secondary school teachers, University lecturers and all those who have immensely contributed to my success in the academic exploration.

Finally this thesis is dedicated to humankind who depends on the limited water resources for life and socio-economic development, yet droughts, climate change and human-induced activities adversely affect the water resource systems.

ABSTRACT

Drought is acritical stochastic natural disaster that adversely affects water resources, ecosystems and people. Drought is a condition characterized by scarcity of precipitation and/or water quantity that negatively affects the global, regional and local land-scales. At both global and regional scales, drought frequency and severity have been increasing leading to direct and indirect decline in water resources. For instance, increase in drought severity and frequency in the upper Tana River basin, Kenya, water resources systems quantity and quality have been adversely affected. Timely detection and forecasting of drought is crucial in planning and management of water resources. The main objective of this research was to formulate the most appropriate models for assessment and forecasting of drought using Indices and Artificial Neural Networks (ANNs) for the basin. Hydro-meteorlogical data for the period 1970-2010 at sixteen hydrometric stations was used to test the performance of the indices in forecasting of the future drought at 1, 3, 6, 9, 12, 18 and 24-months lead times, by constructing ANN models with different time delays. Drought conditions at monthly temporal resolution were evaluated using selected drought indices. The occurrence of drought was investigated using non-parametric Man-kendall trend test. Spatial distribution of drought severity was determined using Kriging interpolation techinique. In addition, a standard Nonlinear-Integrated Drought Index (NDI), for drought forecasting in the basin was developed using hydro-meteoroogical data for the river basin. The performance of the drought forecasting models at the selected lead times were assessed using Mean Absolute Error (MAE), correlation coefficient (R), Nash-Sutcliffe Efficiency (NSE), Ratio of mean square error (RSR) and modified index of agreement (d_1). The results of spaial drought show that the south-eastern parts of the basin are more prone to drought risks than the north-western areas. The Mann-Kendall trend test indicates an increasing drought trend in the south-eastern and no trend in north-western areas of the basin at both 90 and 95% significant levels. Another output of this research was the development of Surface Water Supply Index (SWSI) function, NDI and characteristic curves defining the return period and the probability of different drought magnitudes based on Drought Indices (DIs). In addition, drought Severity-Duration-Frequency (SDF) curves were developed. The formulated NDI tool can be adopted for a synchronized assessment and forecasting of all the three operational drought types in the basin. The results can be used in assisting water resources managers for timely detection and forecasting of drought conditions in prioritized planning of drought preparedness and early warning systems.

TABLE OF CONTENTS

LIST OF SYMBOLS

Symbol	Description
A	Sample statistic
b	bias at the cell body
$f(x)$	Gamma function
$g(x)$	Gamma distribution
$H(x)$	Cumulative probability of observed precipitation
K'	Function representing average water demand
L	Evapo-transpiration when precipitation is zero
m	Number of hidden neurons
n	Number of input neurons
PE	potential evapo-transpiration
R_e	Recharge
rn	Rainfall component
RO	Runoff
rs	Storage reservoir volume
Wi	Weight attached to the input signal i
x_0	Original SPI index
Xi	The input signal
Zi	Palmer moisture anomaly index
α	The shape parameter
β	Scale parameter
$\Gamma(\alpha)$	Value taken by gamma function
RO	Runoff
K_i'	Average water demand
D	Average absolute water deficiency
K	Weighting factor
P_{ij}	Transitional probability from state S_i to state S_j
n_{ij}	Entries of a P matrix
P^{t+1}	Transition matrix at any given time
P^{t+n-1}	Transition matrix of previous time step
P^t	Transition matrix at initial time
t_{jj}	Recurrent time

ABBREVIATIONS AND ACRONYMS

Abrevation/Acronym	Description
ADI	Aggregated Drought Index
ALRMP	Arid Lands Resource Management Project
ANNs	Artificial Neural Networks
ASALs	Arid and Semi-arid lands
CMI	Crop Moisture Index
DDDI	Data Driven Drought Index
DI	Drought Index
DRSRS	Department of Resource Survey and Remote Sensing
EDI	Effective Drought Index
FAO	Food and Agriculture Organization of the United Nations
FEWS-NET	Famine Early Warning system-Network
FFANNs	Feed Forward Artificial Neural Networks
GD	Ground water Drought
GDP	Gross Domestic Product
GIS	Geographical Information Systems
GKF	Gaussian Kernel Function
GoK	Government of Kenya
GUI	Graphical User Interface of the MATLAB
GRNN	Generalized Regression Neural Network
ICPAC	IGAD Climate Prediction and Application Centre
IFAD	International Fund for Agricultural Development
IGAD	Intergovernmental Authority on Development
IPCC	International Panel for Climate Change
ITCZ	Inter-tropical convergence zone
KenGen	Kenya Electricity Generating Company
KFSSG	Kenya Food Security Steering Group, of Government of Kenya
KMD	Kenya Meteorological Department
KNBS	Kenya National Bureau of Statistics
LINK	Livestock Network and Knowledge System
MAE	Mean Absolute Error
MATLAB	MATrix LABoratory

MDGs	Millennium Development Goals
MI	Model Intensive
MuI	Murger Index
MLP	Multilayer Perception
MSE	Mean square error
NDI	Nonlinear-Integrated Drought Index
NCPB	National Cereals and Produce Board
NF	Neuro-fuzzy Network
NIR	Near Infra Red Band
NSE	Nash-Sutcliffe Efficiency
PCA	Principal Component Analysis
PDSI	Palmer Drought Severity Index
PHDI	Palmer Hydrological Drought Index
RBF	Radial Basis Function
RCMRD	Regional Centre for Mapping of Resources for Development
RDI	Reclamation Drought Index
RMSE	Root Mean Square Error
RS	Remote Sensing
RSR	Ratio of root mean square error to standard deviation
SANN	Spatial Analysis Neural Network
SARIMA	Seasonal Autoregressive Integrated Moving Average model
SMDI	Soil Moisture Deficit Index
SPI	Standardized Precipitation Index
SSTAs	Sea Surface Temperature Anomalies
SVM	Support Vector Machine
SWD	Surface Water Drought
SWSI	Surface Water Supply Index
TARDA	Tana and Athi River Development Authority
UN	United Nations
UNDP	United Nations Development Programme
WFP	World Food Programme
WRE	Water Resources Engineering
WRMA	Water Resources Management Authority
NDWI	Normalized Difference Water Index

WSVI	Water Supply Vegetative Index
NDDI	Normalized Difference Drought Index
RH	Relative Humidity
MAE	Mean absolute error
SDI	Stream flow drought index

LIST OF TABLES

LIST OF FIGURES

DEFINITION OF TERMS

Term	Definition
Activation function	An equation used within the neurons for signal data processing
Aggregated drought index	A tool used as an indicator of three categories of drought such as hydrological, agricultural and meteorological droughts. It uses numerous hydro-meteorological variables asinput into the mode
Agricultural drought	It refers to the deficit of soil moisture due to meteorological effects with different timing and effects. It may depend on initial moisture content and water storage capacity of the soil. It adversely affects cultivated vegetation and crop production.
Artificial neural network	A model system for processing large and complex data which uses inputs to generate outputs in the process, it simulates the working principles of neuron of a human brain
Axon	Data output channel of the ANN through which the output signal or information is released
Cell body	A data processing unit within an ANN
Climate change	It is a statistically significant disparity in mean climatic conditions that occurs for an extensive time period in most cases a period of at least decade is considered
Climate variability	It is a phenomena of weather change comprising of shifts in the magnitude and frequency of erratic weather events, and slow and continuous increase in mean temperature
Crop moisture index	It is a tool used to assess the water deficit conditions on soils where crops are grown
Dendrite	Data input conduit is an ANN through which data of signal is transmitted to the cell body
Drought	A condition on land which is characterized by scarcity of water that fall below normal average, threshold or truncation levels
Drought index	A function of water-related variables(such as precipitation, streamflow, resrvoir volume, dam inflow, ground-water level...) for quantifying drought
Global warming	It refers to the critical changes in temperature and temperature dependent variables within the atmosphere and land, caused by high concentrations of hydrocarbon gases (carbon dioxide,

	methane and nitrogen) in the atmosphere
Ground water drought	It is a hydrological drought caused by significantly reduced recharge or reduced volumetric ground water storage and or yield
Hydrological drought	Effect of precipitation deficit on surface or sub-surface water resources such as rivers reservoirs, dams groundwater
Hydrological risk	It is a term is used to refer to a combination of the magnitude of the effects of climate variability impacts and the probability that the effects will occur
Long-term drought forecasting	Prediction of water deficit for a long period of time of up to 12 months lead time
Medium-term drought forecasting	Detection of water deficit for a medium range of time of 6 to 9 months lead time
Meteorological drought	A period of prolonged dry weather condition due to significantly below average precipitation
Non-linear-Integarated drought index	A tool that exhibit non-linear realtionship between input and output variables. It is used as an indicator of three categories of drought such as hydrological, agricultural and meteorological droughts
Palmer drought severity index	A water balance tool for assessing the total moisture status of a region using supply and demand over a two-layer soil model
River basin	An area of land surrounded by boundary line called drainage divide, where water flow on the surface, streams and rivers resulting from precipitation converge to a single point, called outlet.
Short-term drought forecasting	Detection of water deficit conditions for a short period of time of 1 to 3 months lead time
Socio-economic drought	Refers to all socio-economic effects of agricultural, hydrological and meteorological droughts on supply and demand of goods and services in the society
Soil Moisture Deficit Index	It is a tool used to assess the water scarcity conditions on different soil profiles where crops are grown
Standardized	A tool for evaluating drought conditions of an area in terms of

precipitation index	precipitation defict based on certain threshold levels
Surface water drought	This is a condition of water deficit below a certain threshold level for surface water sources within a basin
Surface water supply index	It is an indicator of surface water and or moisture conditions on a basin

CHAPTER ONE
INTRODUCTION

1.1 Background information

Drought is one of the critical stochastic natural disasters that adversely affect water resource systems, people and ecosystems (Zargar *et al*., 2011; Jahangir *et al*., 2013). Drought is defined as a hydro-meteorological event on land characterized by temporary and recurring water scarcity. According to Morid *et al*. (2007), the magnitude of the drought is indicated by the extent with which it falls below a defined threshold level over an extended period of time. Drought has been identified as the most complex natural hazards due to difficulty in its detection. Drought preparedness and mitigation depend upon timely information on its onset, and propagation in terms of temporal and spatial extent. Such information can be obtained if effective and continuous drought monitoring indices are used in drought evaluation. The study of spatial and temporal drought conditions has greatly been applied in planning and management of water resource systems such as water supplies, irrigation systems, and hydropower generation (Ceppi *et al*., 2014; Abad *et al*., 2013; Alaa, 2014; Okoro *et al*., 2014). These studies were undertaken in Lombarrdy region of norh Italy, Bashar river basin, Mashtul pilot area of Nile Delta and the river basins in Imo state of Nigeria respectively.

At a global scale, demand for water resources has continued to increase as a result of the population pressure and related socio-economic development needs. As a result, numerous sectors have been affected by water scarcity and thus, effective management of impacts of drought-induced water deficit is required. These drought impacts are more severe on Arid and Semi-Arid Lands (ASALs) than in the humid areas (UNDP, 2012). Therefore, management of drought has become an important issue in most of the countries in the world. The drought characteristics in terms of frequency, duration and severity have been assessed using Drought Indices (DI) in some parts of the world (Mishra and Sign, 2010; Barua, 2010; Belayneh and Adomowski, 2013). However in many regions of the world such as Kenya, drought forecasting is still inadequate and thus the need to develop forecasting techniques for Kenya.

Globally, drought has become more frequent and severe due to climate variability with some regions experiencing droughts at varying scales and times. Therefore, global impacts of drought on environmental, agricultural and socio-economic aspects need to be studied. Drought is classified into four distinct types namely; meteorological, agricultural, hydrological and socio-economic. According to IPCC (2014), these droughts have either direct or indirect impacts on river basins and

human lives. Direct impacts include degradation of water resources in terms of quantity and quality, reduced crop productivity, increased livestock and wildlife mortality rates, increased soil erosion and land degradation, and increased plant diseases and insect attacks (UN, 2008; Scheffran *et al*., 2012). On the other hand, the indirect impacts comprise reduced income, unemployment, and migration of people and animals (Figure 1B; Appendix B). Worldwide, more than eleven million persons have died since 1900 as a result of drought related impacts. In addition, two billion persons have been critically affected by the impacts (FAO, 2013). The main challenges associated with drought are that it causes ill health through water scarcity, malnutrition and famine (UN, 2008).

African countries are among the most vulnerable to adverse impacts of climate variability and drought. This is because the Gross Domestic Product (GDP) of these countries depends heavily on rain-fed agriculture. The impacts in Africa are compounded by numerous factors such as poverty, high population density, and human diseases. According to Mwangi *et al*. (2013) in East Africa, it has been projected that water availability will decline due to drought. In addition, there is a likelihood of increased drought frequency due to decline in precipitation especially during the dry months. Extreme weather events will cause more frequent and intense droughts in some areas and flooding in others. This may be accompanied by hurricanes and tropical storms especially on coastal areas (IPCC, 2001).

In Kenya, very notable droughts of 2009 and 2011 adversely affected the agricultural sector where crop yields were drastically reduced. During this period, the country's wheat yield dropped by 45% compared to the 2010 growing season (FAO, 2013). Similarly, between 2002 and 2010, Australia suffered a multi-year drought. The total wheat yield in Australia at the time dropped by 46% compared to the annual average level. In 2010, Russia suffered a long and severe drought which significantly affected the environment, human health and economy. In the US, the southern states experienced a severe drought in 2011while in 2012 more than 6.3 million persons were negatively affected by drought in China. During such drought episodes, people experienced challenges in food access and drinking water (FAO, 2013).

Drought and climate variability have significantly impacted on Kenyan river basins. Human activities have also led to human-induced climate variability in most of the Kenyan river basins. This has aggravated pressure on water resources in these river basins. Climate change is expected to impact on different sectors of Kenya's economy. For instance, the water sector which is the driver of other sectors of Kenya's economy will be adversely affected. Kenya is a water scarce

country with approximate renewable freshwater of 643 m^3 per capita per year and is projected to drop by a third to 235 m^3 by the year 2025 (GoK, 2009). These values are far below the recommended level of 1000 m^3 per capita per year (WHO, 2010).

The agricultural sector in Kenya is highly vulnerable to climate variability (FAO, 2011). This sector contributes more than 51% of the gross domestic product (GDP) (Mwangi *et al*., 2013) and has been critically affected by frequent droughts. Over the past 50 years, Kenya has experienced at least one main drought per decade (FAOSTAT, 2000). In addition, there has been a notable increase of drought in terms of frequency, duration and intensity. Any damage caused by drought on agriculture and water resources leads to famine, humanitarian crisis, rationing of water supply and decline in hydropower generation, the main source of power in the country that constitutes approximately 50.7% of the total power (World Bank, 2013). Effective drought forecast allow water resource decision makers to develop drought preparedness plans. Such plans are critical for advance formulation of programmes to mitigate drought-related environmental, social and economic impacts. Therefore, accurate drought assessment and forecasting with an adequate lead time is paramount in formulation of mitigation measures for river basins (Sharda *et al*., 2012).

With the current trend in climate change, the average temperatures in the highlands of Kenya are predicted to increase by approximately 1.4^0C whereas the annual rainfall will increase by 20% by 2050. This will lead to an increase in runoff in most parts of Kenya. However, some areas such as the northern and the western regions may have little increase of about 5% of annual rainfall. In addition, potential evapo-transpiration is expected to increase by 10% due to an increase in temperature by 2050 (WWF, 2006). These changes are going to affect drought and its impacts at varying magnitudes and time-scales within basins.

Most of the electrical power in Kenya is generated within the upper Tana River basin mainly from five hydro-power plants. These include Masinga, Kamburu, Kindaruma, Gitaru and Kiambere which have a generating capacity of 40, 44, 225 and 156 MW respectively (Oludhe, 2012). These plants are part of the Seven-Forks dams designed for hydro-power generation. Masinga dam was originally designed as the main water reservoir purposely to regulate water flow to the other plants. The quantity of hydro-power generation depends wholly on the availability of water in the reservoirs at reasonable design levels to drive the turbines. However, in recent times, the availability of water within the Masinga reservoir has greatly been adversely affected by

occurrence of droughts. The occurrence of drought is associated with low water levels in the reservoirs as a result of reduced stream flows feeding the major reservoir.

Drought forecasting has received a new approach especially with the development of the Artificial Neural Networks (ANNs). An ANN is a computing system made up of a number of simple and highly interconnected information processing elements. Such a system has performance characteristics that resemble biological neural networks of human brain. ANN has numerous merits when used for data processing. The system processes information based on their dynamic state response to external input (Morid, 2007). ANNs have the capacity to model relationships that are quite dynamic and can capture many kinds of relationships including non-linear functions which are difficult or impossible to determine using other approaches (Mustafa *et al*., 2012). The ANNs have recently been used in water resources engineering (WRE). WRE comprises the study of hydraulics, hydrology, environment and geological related variables. Such variables are dynamic and exhibit non-linear and stochastic characteristics. These properties make WRE variables complex and difficult to determine due to spatial and temporal variations. Thus due to their advantages, ANNs provide effective analytical techniques in modelling and forecasting non-linear and dynamic time series variables in WRE such as drought (Mustafa *et al*., 2012).

Agricultural drought can be monitored by assessing soil moisture content levels. However, direct soil moisture data measurement is not available at regional and basin scales at fine resolutions. To estimate soil moisture content, process based models are used where an integration of both the random variables of climate, physical properties of land are considered. One of the advanategs of these models is that they can be used to provide information at different spatial and temporal resolutions. Some of these process based models include the FAO developed AquaCrop model (Casa *et al*., 2013) and Soil and Water Assessment Tool (SWAT) (Fiseha *et al*., 2013). The Aquacrop model has been used in modelling crop response to soil water availability. For instance, AquaCrop model was applied in evaluating wheat grain yield and crop biomass in China for several irrigation systems (Du *et al*., 2011). The model was applied by Iqbal *et al*. (2014) to assess crop grain yield and biomass response to soil water content, actual evapotranspiration under deficit irrigation conditions.

Most basins in Kenya have limited or lack adequate quantifiable information on drought characteristics such as magnitude, frequency, duration and severity. In addition, there is very limited information on appropriate drought assessment and forecasting methods. Drought

assessment for different river basins is essential in understanding the trends in drought frequency, severity, magnitude and associated impacts. The information of such assessment can be adopted in informed decision making by governments and its support programmes on the affected communities. Drought models can be used to estimate and forecast drought conditions on a spatial and temporal domain. To prepare for effective mitigation of drought risks in Kenya, asessment and forecasting of drought conditions is vital. Thus in this research, assessment and formulation of drought forecasting models for upper Tana River basin through the application of Drought Indices and Artificial Neural Networks (ANNs) was accomplished.

1.2 Statement of the problem

There has been a problem of erratic drought occurence that has negatively affected water resource systems and consequently socio-economic development in the upper Tana River basin. The problem of drought in the basin is aggravated by the fact that its frequency and severity has been increasing over the years. Being a stochastic hydrological phenomenon, it is difficult to assess and forecast drought. Despite cascaded adverse impacts of the drought occurrence on decline of quantity and quality of water resources (WRMA, 2009; IFAD, 2012), hydropower generation in the Seven fork cascade dams (Word Bank, 2006; WSI, 2011), limited research has been conducted to assess and forecast its characteristics. Drought occurrence in the basin is attributed to the combined effects of both climate change and land use/cover change that lead to increased evapo-transpiration (ET). According to Nohara *et al*. (2006), increase in ET is caused by increase in temperature that negatively affect river basin water resource systems such as stream flows, reservoir levels and soil moisture levels at varying magnitudes. There is need to formulate sustainable drought mitigation and copying mechanisms for the basin. However, an appropriate tool(s)/methods for assessing and forecasting drought conditions (severity, duration and frequency) in the basin are limited. To address this challenge, this study therefore formulated effective models and Non-linear Integrated Drought Index (NDI) for assessment and forecasting of drought in the upper Tana River basin.

1.3 Objectives

1.3.1 Main objective

The main objective of this study was to formulate the most appropriate models for assessment and forecasting of drought using Indices and Artificial Neural Networks (ANNs) in the upper Tana River basin for guiding decision making in water resources planning and management.

1.3.2 Specific objectives

The specific objectives of this study were to:

i) Assess temporal and spatial drought using selected Drought Indices (DIs) based on hydro-meteorological data from 1970 to 2010 for the upper Tana River basin
ii) Evaluate the performance of the selected DIs and ANNs in forecasting of short, medium and long-term drought conditions in the upper Tana River basin
iii) Formulate a Nonlinear-Integrated Drought Index (NDI) using principle component analysis based on the basin hydro-meteorological data for 1970-2010
iv) Assess the performance of the formulated NDI and ANNs in forecasting and projecting short, medium and long-term drought conditions in the river basin

1.4 Research questions

i) How can the the spatial and temporal drought condition based on Drought Indices (DIs) assessment using hydro-metrological data of 1970 to 2010 for upper Tana River basin?
ii) How do the DIs and ANNs perform in forecasting drought indices for short, medium and long-term drought characteristics of upper Tana River Basin?
iii) How effective can a Nonlinear-Integrated Drought Index (NDI) be formulated using hydro-metrological data for upper Tana River basin?
iv) How do the formulated NDI and ANNs perform in forecasting NDI values for short medium and long-term drought characteristics for the basin?

1.5 Justification

Improved socio-economic development which is associated with sustainable water resources availability, food security and hydro- power generation and supply in the upper Tana River basin is one of the Kenya's priorities. To reduce the levels of poverty, hunger, improve sustainable access to safe drinking water, sustainable conservation of resources, as per the Kenya's Vision 2030 (GoK, 2007), understanding of drought characteristics is paramount (GoK, 2012). Although drought has affected water resources in the basin, quantification of its impacts and characterization of drought is limited. There is need to end poverty, hunger, promote sustainable water management and access to sustainable energy as stated in the UN sustaibale development goals 1, 2 6 and 7 respectively (UN, 2016). To achieve these goals, enhanced drought preparedness through planning of mitigation measures to reduce adverse drought impacts on water resources, food security, hydro-power generation and livelihoods is paramount. This requires data on spatial and temporal characteristics of drought. Such data is useful in identification of drought on-set, its propagation as

well as detection of drought risk areas in the basin. However, such data for the upper Tana River basin is scanty and not readily available.

1.6 Scope

Although the Tana River basin covers 126,026 km^2, the current research focused only on the upper Tana River basin with an area of 17,640 km^2. Assessment and forecasting of drought was based on data-driven drought indices that use hydro-meteorological data from 1970 to 2010 for eight gauged stream flow stations and eight meteorological stations. The data analysis is based on monthly and 90-m temporal and spatial resoulutions respectively.

CHAPTER TWO
LITERATURE REVIEW

2.1 Occurrence of droughts

Drought is a condition on land characterised by scarcity of water that falls below a defined threshold level. The term drought has been defined differently in numerous applications (UNDP, 2012). However, it is a challenge to quantitatively define the term. Droughts may be expressed in terms of indices that depend on precipitation deficit, soil-water deficit, low stream flow, low reservoir levels and low groundwater level. Drought may be defined differently depending on the sector involved. For example, a hydrological-drought occurs whenever river or groundwater levels are relatively low. In addition, water-resources drought occurs when basins experience low stream flow, reduced water reservoir volume and groundwater levels. The water resources drought is influenced by climatic and hydrological parameters within a river basin and drought management practices. The hydrological drought, mainly deals with low stream flows. This drought adversely affects various aspects of human interest such as food security, water supply and hydropower generation (Karamouz *et al*., 2009; Belayneh and Adamowski, 2013).

A sequence of drought occurrence in a river basin may lead to desertification of vulnerable areas such as arid, semi-arid and sub-humid areas. Within these fragile ecosystems, water resources, soil structure and soil fertility are critically degraded by drought occurrence (El-Jabi *et al*., 2013). The occurrence of any drought in terms of magnitude, frequency, duration and severity has not been clearly understood for numerous river basins in the world, and this calls for intensified research in drought and such related fields.

2.1.1 Types of droughts

According to Zoljoodi and Didevarasl (2013), there are four main categories of droughts. These include the Hydrological, Meteorological, Agricultural and Socio-economic droughts. The first three types are called the operational droughts and can be integrated into a drought management algorithm. Their relation can then be applied in development of water resource strategy in a river basin (Karamouz *et al*., 2003). Propagation of hydrological and agricultural drought originates from meteorological droughts which develop from changing phenomena within the hydrological cycle as given in Figure 2.1. The main droughts may further be classified into other types of drought.

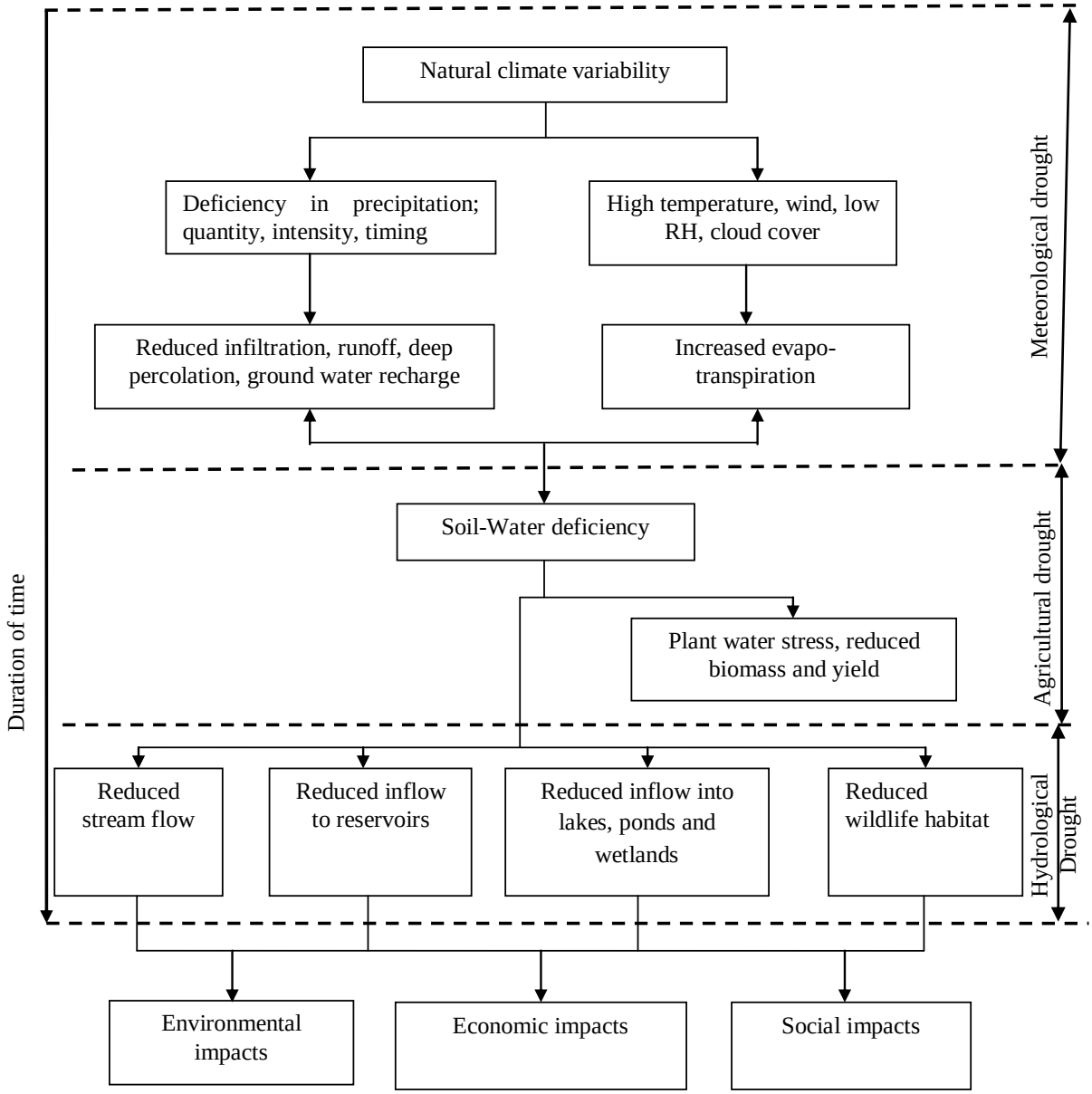


Figure 2. 1: Propagation of drought via hydrological cycle

The hydrological drought is associated with below average quantity of surface and sub-surface water resources resulting from deficit precipitation. Its characteristics which are defined by magnitude, severity, duration and frequency can be studied at a basin scale. Hydrological drought impacts on large areas and large human population and may be triggered by climate change and /or variability (Mondal and Mujumdar, 2015). Like other drought events, hydrological drought is considered to be a 'creeping hazard' because it develops slowly, it is not easily noticed, covers extensive areas and it lasts for a long period of time with adverse impact on water resources, ecological systems, and socio-economic development (Liu *et al*., 2015; Van-loon, 2015).

According to Van-loon and Laaha (2015) and as shown in Table 1A of Appendix A, hydrological drought has the most significant effects across different sectors compared to other types of droughts. Hydrological drought may be categorized into surface and ground water droughts (GD). The Surface Water Drought (SWD) is caused by direct reduction in precipitation that subsequently leads to low surface runoff. The SWD is also caused indirectly by reduced groundwater discharge to surface water resources. This may be attributed to reduced flow of groundwater into surface flow in influent rivers and springs. In some instances, increase in groundwater on specific areas within a basin for an effluent river contributes to the SWD. The common indicators of SWD are reduced river flows, low water levels in reservoirs and lakes. SWD results from a combined interaction of meteorological drought, water resources development infrastructure and operational management.

Groundwater Drought (GD) is a hydrological type of drought caused by significantly low quantity of water in aquifers that may be due to reduced recharge. The recharge normally takes place through permeation and inflow from sub-basins (Adindu *et al*., 2013). The GD is assessed by measuring the volumetric ground water storage. However, these data are not readily available in most river basins. Thus, aquifer level is considered to be a better indicator than the volumetric ground water storage. The GD is also determined from the evaluation of its secondary effects such as base flow into rivers. Ground water is a vital source of water supply especially in river basins where surface water exhibits a high temporal variability. In some cases, groundwater availability is used as an indicator of relative drought risk.

The meteorological drought which is the most commonly known drought is associated with long time intervals of significantly low or no precipitation and increased air temperature. The deficiency in rainfall leads into low infiltration, decreased runoff and ground water recharge. High air temperatures lead to changes in wind characteristics such as increased wind velocity, low Relative Humidity (RH) and increased evapo-transpiration.

The three operational types of droughts are interconnected. For instance, Agricultural drought links meteorological and/or hydrological drought to agricultural impact. Agricultural droughts impact negatively on farming systems whenever they occur. Their impacts are normally two-fold; environmental and economic impacts. The agricultural drought is a type associated with low agricultural production, increased food insecurity, decline in output from agro-processing industries and unemployment incidents in the agricultural sector. From the environmental perspective, agricultural drought is caused by insufficient precipitation, high temperature that causes elevated

rates of evapo-transpiration, increased salt concentration in the crop root zones and soils within irrigation systems (Mishra and Singh, 2010). The term environmental drought is sometimes used to address the adverse effects of extremely low flows on ecosystems, and may be analyzed in the emerging field of eco-hydrology.

The term socio-economic drought relates the supply and demand of economic goods with elements of meteorological, hydrological and agricultural drought. This drought occurs when the demand for an economic good exceeds the supply. It is caused by weather related deficit in water supply within a basin. Some of the noticeable impacts of socio-economic drought include increased unemployment, increased food prices, reduced income, reduced tax revenues and increased migration.

2.1.2 Drought modelling

Drought modelling is a technique of using simple and or complex mathematical, scientific and conceptual representations of drought charactersitics. The purpose of drought modelling is mainly to provide a concise understanding of its occurrence, characteristcics and forecast. The fundamental role of modelling the drought is to offer solutions to the challenges of increasing water scarcity due to population growth, expansion in agriculture, industrial and energy sectors. The scarcity of water in the world is compounded by the droughts that affect both surface and ground water resources. Drought modelling may be categorized into eight aspects such as drought forecasting, temporal assessment, spatial assessment, probability characteristics, management, impacts of climate change on drought, assimilation of land data systems into drought and impacts of drought on different sectors. All the aspects of drought are interconnected in drought modelling and are inter-linked as given in Figure 2.2 which was modified from Mishra and Singh (2011).

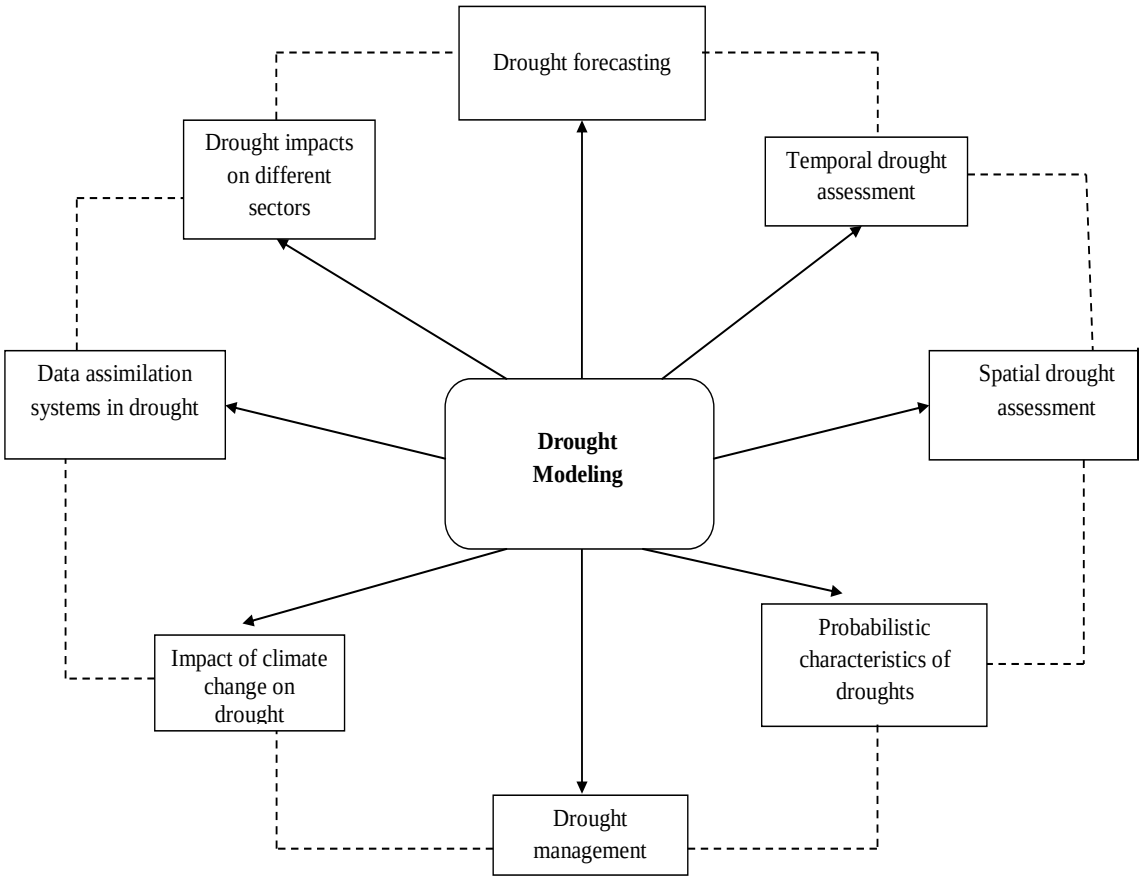


Figure 2. 2: Network of phases of drought modeling
(Modified from Mishra and Singh 2011)

2.1.3 Determination of drought threshold level

The threshold level of any drought is based on the theory of crossing technique where the properties of runs above and below a truncation level are determined. The truncation level may be considered as the long-term mean or median value computed from hydro-meteorological data as shown in Figure 12B, Appendix B. The truncation level for a drought may be smaller than the lowest available value of a particular data set. The truncation level is used to specify some statistics of a drought variable. This is achieved by portioning the time series of the variable deficit and surplus segment. The truncation level may be constant or vary with time. When using a constant threshold level in most cases, the absence of trend should be first checked. Given a time series data such as stream flow, reservoir and ground water levels, a threshold level may be determined using a modified function according to Peters *et al*. (2003) given as:

$$\int_0^M \left(Q_{T(c)} - Q_{(t)}\right) dt \;=\; c \int_0^M \left(Q_b - Q_{(t)}\right) dt \qquad (2.1)$$

$$Q_{(t)} = \begin{cases} Q \, if \; Q \geq 0 \\ 0 \, if \; Q < 0 \end{cases} \quad (2.2)$$

Where;

$Q_{(t)}$ = stream flow (m^3/s)

$Q_{T(c)}$ = threshold value

Q_b = minimum value (average supply of average demand) (m^3/s)

M = period of time in series

c = drought criterion factor

The drought criterion factor c is the ratio of deficit value below the threshold to the deficit below the average value. The drought criterion parameter c determines the height of the threshold level. If the value of c is one, the threshold level is equal to the average of Q. If the value of c is zero, then the threshold level is equal to the minimum of Q. The definition of the threshold also ensures that the total drought deficit decreases with decreasing amplitude of $Q_{(t)}$. The last line of the function is included by considering that deficiency is influenced by supply and demand (Awass and Foerch, 2006a). The $Q_{T(c)}$ is an arbitrary value that depends upon the objective of the study of drought based on water scarcity as a relative concept that can occur at any level of supply or demand. The truncation level is normally used to objectively demarcate the on-set and end of a drought event.

2.1.4 Selection of drought threshold level

The truncation level for any drought assessment may be chosen based on mean, median and percentage of exceedance of the available data set and purpose of the drought study. For instance, the truncation levels for stream flow may be taken as a percentile such as Q_{50}, Q_{70}, Q_{90} (Awass, 2009). The first step in selecting a threshold level Q_0 is to define its value, below which a stream flow or precipitation data is considered as drought. Then the method of 'Crossing Theory' also called 'Run-sum Analysis' is applied to investigate drought characteristics.

When the plotted data for drought falls below the threshold value, a drought event starts, and when it rises above the threshold, the drought event ends. The beginning and ending of drought are defined using the start and end times. For river basins with mainly perennial rivers, relatively low fixed thresholds in the range of Q_{70} to Q_{95} may be considered appropriate (Meigh *et al*., 2002; Sung *et al*., 2013). However, for basins dominated by intermittent and ephemeral rivers, which have large proportion of zero flow, Q_{70} or mean flow is recommended. If the trend of the data changes with time, then a variable threshold can be applied for detection of deviations and drought

assessment. For instance, monthly flow duration curves could be developed and a variable threshold value of Q_{75} used for the given river basin (Kjeldsen *et al*., 2000).

2.2 Climate change and variability

Climate change is explicitly defined as statistically significant and long-term (at least a decade) continuous disparity form mean climatic conditions (IPCC, 2001). Climate variability refers to the deviation in mean state of climate variables on temporal and spatial scales. It is a deviation of climate statistics over a defined period of time such as a month, season or a year compared to long-term staitistics of the same calendar period (WMO, 2015). This change comprises shifts in the magnitude and frequency of erratic weather events and slow and continuous increase in mean temperature. Climate variability dynamics was not well understood in early times on all scales. There has been a great improvement in understanding dynamics of climate variability on a large-scale in recent times. However, more research needs to be done on its cascaded effects to hydrology, quantity and quality of water and sustainable management of water resources at the river basin scale (Peng *et al*., 2013).

Climate variability has the ability to drive climate and ecosystems across certain thresholds and create a new condition. There may be a shift in mean climatic conditions and extreme events such as more frequent floods and droughts, severe soil erosion and prolonged periods of low stream flow. The climatic change and variability plays a key role in the modification of spatial-temporal patterns of hydrology. For instance, according to Ma *et al*. (2009), climate variability has led to alteration of the hydrological processes, reduced glaciers and water supply downstream of Himalaya's catchment. In addition, it influences the human socio-economic activity through land cover and land use within a watershed (Hundecha and Bardossy, 2004). In addition, climate has asignificant adverse effects on all aspects of economy. For instance the 1998-2000 droughts in Kenya were estimated to have economic decline of 2.8 billion United States Dollars. This was due to the losss of crops, livestock damage to fisheries, reduced hydropower generation, reduced industrial production and decline in water supply (Gok, 2012)

Due to climate variability, the natural ecosystems require sustainable management for resilience. The process of resilience involves recovery of ecosystems from significant multi-hazard threats with minimum damage to environment, social and economic wellbeing. However, due to systems going beyond thresholds, managing resources for resilience may not remain a viable strategy. Instead, it might be more viable to manage systems in their new stable states. It is therefore

important to conserve, adapt and/or mitigate climatic variability impacts and reduce their risks using integrated approaches to hydrological risks (UNEP, 2015).

The term hydrological risk is used to refer to a combination of the magnitude of the effects of climate variability impacts and the probability that the effects will occur. The effect of climate variability on hydrological extreme events has a great challenge in water resources and basin management. There is the need to formulate strategies to efficiently manage the challenges at river basin level (WRMA, 2010).

The threat of climate change on basins degradation in the past decade has been of great concern as far as conservation of biodiversity is concerned (Ayyad, 2003; Chen and Rao, 2008). Based on the period in years of occurrence, climate variability may be categorized as short-term and long-term (IPCC, 1995). Climate change has been attributed to global warming. For instance, global surface air temperature has been noted to change by 0.74 ± 0.18^{o}C from 1906 to 2005 (IPCC, 2007). The climatic variables which impact immensely on water resources include the precipitation and temperature. One of the resulting adverse effects on water resource systems is the climatic-induced water scarcity. The impact of climate variability on the hydrological response of a basin may be investigated using accurate field data and modelling techniques.

Some of the hydrologic variables affected by the climatic variability include surface runoff, sub-surface runoff, stream flow, sediment yield and soil erosion (Kamga, 2001; Tong and Cheng, 2002). The climatic variables which impact immensely on water resources include the precipitation and temperature. One of the resulting adverse effects on water resource systems is the climatic-induced water scarcity that needs detailed investigation. The impact of climate variability on the hydrological response of a basin may be investigated using accurate field data and modelling technique.

Prediction of hydrological response associated with climate variability requires the use of simple or complex distributed hydrological models. However, the use of complex models may require large datasets for both the pre-change and post-change basin conditions. Alternative approaches that combine spatial modelling and qualitative techniques may therefore be applied (Bassey and Akinkunmi, 2013). A number of approaches have been used to study changes within different basins. Hydro-meteorological measurements and the application of Remote Sensing (RS) and

Geographical Information Systems (GIS) techniques have been widely used to detect effects of climate change on basins.

Some of the efforts made to monitor climatic change are measurements of stream flow, water balance components, levels of lakes, tree ring growth rates and drought risks. The regional trends in climate change may be reflected in the projections for countries such as Kenya. It is expected that the mean annual temperatures in Kenya will increase from 1 to 2.8°C by the 2060s, and 1.3 to 4.5°C by the 2090s (IPCC, 2007). On the other hand, the mean annual rainfall will increase by up to 48%. It is anticipated that there will be a general increase in precipitation due to climate change in most arid and semi-arid lands (ASALs) of Kenya (IPCC, 2007).

2.2.1 Impact of climate change on water resources

Climate change and land-cover change within a basin may greatly impact on the quantity and quality of water resources. Reservoir water availability is influenced by drivers of climate variability especially the ones that alter runoff within a basin. To manage water resources under climate variability, the capacity of water reservoirs should be increased for offsetting the impacts of climate variation and maintaining existing water yields.

Climate change influences the patterns of hydrology, droughts and water resources systems (Ma *et al*., 2010). Although hydrology and water resources are interrelated, the two terms differ in their definition. Hydrology is the study of the water on, above and below the earth in terms of its occurrence, circulation, distribution, chemical and physical properties, and the reaction of the water with the environment, including the interaction with ecosystem within a basin. On the other hand, water resources refer to the quantity and quality of water based on a defined temporal and/or spatial resolution.

Due to climate change and land-cover change, hydrological processes in basins have significantly been altered especially by extreme climate events (Jones, 2010). Extreme climate events such as droughts are series of occurrences that happen with greater intensity or frequency than common events. The climate variability has led to increased climatic uncertainty with variation in the weather pattern, mainly between the seasons and years. The impact of climate change on water resources and agriculture is considered as both regionally distinct and spatially heterogeneous. One of the challenges in monitoring trends in climate change in river basins is the paucity of meteorological data both in terms of period of record and the distribution of stations. For instance,

in many countries, precipitation, temperature, wind and other hydro-meteorological data recording began only after the Second World War. This makes it difficult to statistically determine the extent of variability of an event in terms of magnitude and frequency (Awass, 2009). However, in some areas there are enough data to study annual trends.

2.2.2 Effect of water balance components on drought

Different components of the hydrological water balance equation are influenced by the drought. The water balance equation components for a river basin are summarized in the expression:

$$P = R - ET \pm \Delta S - G \tag{2.3}$$

Where;

P = Precipitation (mm)

ET = Evapo-transpiration (mm)

ΔS = Change in soil water storage (mm)

G = The ground water recharge (mm)

From Equation (2.3), it implies that precipitation is a vital contributor to the runoff and thus it greatly influences stream flow. During a drought event, the quantity of precipation is significantly reduced below the normal average while temperature in a river basin may increase above long-term mean value. As a result, there is an increase in evapo-transpiration, and thus alteration of the magnitude of water balance equation components. When evapo-transpiration (ET) on the soil surface is high, then the quantity of water stored in the top soil horizon or the root zone is low. This condition contributes to agricultural drought. The ET, ΔS and G components of the water balance are greatly influenced by physiological charatceristics of the river basin which may cause accelerated or delay in hydrological response.

2.2.3 Effect of global warming on droughts

Global warming is due to firstly, the climate change which slows down the global circulation of ocean currents due to moderated differences in temperature between tropical and temperate sea water bodies. Secondly, due to the ice melting in the Polar regions implying cold water entering the oceans and drifting into the tropics affect global warming. The ice melting flowing leads into cooling of tropical oceans whose effect is picking significantly of low moisture by the prevailing winds (IPCC, 2011). The wind carries with it the little moisture picked along its course.

Global warming influences the rate and timing of evapo-transpiration. Due to global warming, some regions in the world are likely to get wetter while those that are already under dry conditions are likely to get drier. Thus, global warming is likely to increase drought occurrence and expansion of the dry areas (Dai, 2011). For instance, the regions in southern Africa, the Sahel region of Africa, southern Asia, south west of United States of America have generally been getting drier over the years (FAO, 2013).

In addition, the quantity of water resources is expected to decline by up to 30% in the areas affected by climate change. These notable changes will occur partly because of an expanding atmospheric circulation pattern. This pattern is called Hadley cell in which warm air in the tropics rises, loses moisture, and descends in the sub-tropics as dry air. During this process, jet streams shift to the higher latitudes, and storm patterns shift along with them leading to expansion of Arid and Semi-arid Lands (ASALs) (Dai, 2011).

2.2.4 IPCC and IGAD approach

The inter-governmental panel for climate change (IPCC) projects changes in climate system using a hierarchy of climate and earth system models. The models simulate changes in climtate based on a set of scenarios of anthropogenic forcing. One of the new sets of the scenarios such as Representative Concentration Pathways (RCP) is used for climate monitoring. This scenario projects atmospheric carbon concentration to be higher in the year 2100 than at present.

The IPCC monitors global surface temperature data sets for the purpose of projecting global warming using integrated climatic models such as simple, atmospheric chemistry and global carbon cycle models (IPCC, 2013; IPCC, 2014). The Intergovenmental Authority on Development (IGAD) also applies a number of climatic models to predict ten-day, monthly and other periodical climate data. It uses climate data bank comprising decadal precipitation and temperature for respective regions. These data sets are then used to develop regional precipitation and temperature risk maps (IGAD, 2007). However, the IPCC and IGAD provide information based on global scales and thus does not give a direct and effective indicator at basin scale scenarios. Most of the information at a basin scale is obtained through downscaling of the global data.

2.3 Major causes of drought in Kenya

Climate variability and global warming that affect atmospheric circulation play a fundamental role in influencing drought occurrences in Kenya. When the Indian Ocean surface water temperature is

abnormally low, it leads to the cooling of South-East and North-East trade winds. These two air masses move across the land and converge near the equator within a region called Inter-Tropical Convergence Zone (ITCZ). When the winds are cool, they do not pick up enough moisture from the ocean water surface and thus lead to erratic rainfall patterns in eastern parts of Kenya (GoK, 2009). On the western parts of Kenya, the Atlantic and Congo prevailing winds bring the same drought conditions in case they are abnormally too cold to pick sufficient moisture.

Other factors leading to increase in drought frequency and severity are human-induced such as poor land use practices, deforestation and destruction of catchment areas. Generally the forest cover has decreased by 72% between 2000 and 2007 (GoK, 2009) and further to 6.1% of the total land mass in 2011 (World Bank, 2014). Among the five main water towers in Kenya, the Mau Complex lost the highest with 70% of the forest cover destroyed within the stated period. The major causes of deforestation have been the need to expand agricultural land, uncontrolled exploitation of forest resources, overgrazing and establishment of new settlements on forest land as accelerated by increasing population growth.

2.3.1 Impact of drought in Kenya

Kenya has experienced approximately 30 major droughts during the last 100 years according to UNDP (2012). Over 70% of the natural disasters in Kenya are associated with droughts and extreme weather conditions. The severity and frequency of droughts in the country have been increasing over the years (Arnell, 2004). Some of the recognizable droughts include the 1952-1955, 1973-1974, 1983-1984, 1992-1993, 1999-2000 and 2009-2011 (UNDP, 2012). The occurrence of droughts in Kenya has impacted negatively on people's livelihoods in the sense that, huge resources which would have otherwise been used for other socio-economic activities are normally diverted to cater for food shortages and water scarcity. The 2008-2009 droughts for instance, caused adverse effect on 3.8 million pastoralists and agro-pastoralists in Kenya in terms of food insecurity. Approximately 1.5 million children in primary schools needed to be fed due to famine (KFSSG, 2009). Many people within the arid and semi-arid lands (ASALs) suffered malnutrition due to lack of food and water.

The drought of 2011 was also severe in terms of food and water scarcity in Kenya where over 3.2 million persons needed urgent care (KFSSG, 2011). In addition, over half a million persons were forced to migrate from their settlements in search of forage, food and water. The cattle were grazed through wildlife protected areas bringing competition and conflict between people, livestock and

wildlife on pasture and water resources. Additionally, the tourism sector which is a major foreign exchange earner for Kenya was greatly affected. This is because drought affects the environmental resources that are crucial attraction for tourists including the wildlife, biodiversity, water quantity and quality.

Drought has impacted negatively on numerous river basins in Kenya. For instance, the upper Tana River basin has been negatively affected by notable droughts such as the La Niña of 1999 to 2000, and 2008 to 2009 (Oludhe, 2012). These led to severe water scarcity in the region and a significant reduction in hydro-power generation that was characterized with rationing of power. To address the problem, there is need to have information for planning ahead. Therefore, this led to a detailed assessment evaluation and characterization of drought occurrences within the basin and finally the development of a prediction system that could be used for early warning on drought occurrence in the upper Tana River basin.

2.3.2 Drought monitoring in Kenya

To minimize the impacts of drought in Kenya, an effective and timely monitoring system is necessary. Such monitoring activities are meant to provide critical information in the development of early warning systems. Since drought has become a recurrent phenomenon in Kenya, different organizations are involved in coming up with methods to address drought-induced challenges.

The organizations involved in data collection for early drought warning systems in Kenya include; the Kenya Meteorological Department (KMD), Ministry of Agriculture, livestock and fisheries, Department of Resource Survey and Remote Sensing (DRSRS), Kenya National Bureau of Statistics (KNBS), Inter-governmental Authority for Development (IGAD) Climate Prediction and Application Centre (ICPAC), World Food Programme (WFP) Kenya Office, Regional Centre for Mapping of Resources for Development (RCMRD), Livestock Network and Knowledge System (LINK), Famine Early Warning system-Network (FEWS-NET) and the Arid Lands Resource Management Project (ALRMP) (WFP, 2011; UNDP, 2012). To prepare adequately for mitigation of the drought impacts, a thorough assessment, evaluation and forecasting of drought conditions is very critical. However, the main challenges or gaps with drought preparedness and mitigation in Kenya include the fact that:

i) Drought assessment has been based on past and present drought indices (DIs) developed for specific regions in other countries and their suitability in Kenya has not been sufficiently tested

ii) Efficiency of the performance of DIs in drought forecasting for different lead times is not well explored in most basins in Kenya
iii) Spatial and temporal drought assessment and forecasting of drought using hydro-meteorological variables is limited for most basins in Kenya

Due to these challenges, there was need to test the applicability of the Drought Indices (Dis) in Kenya using data forms available in concerned river basins. To effectively define the droughts in Kenya, a combination of hydro-meteorological data should be used as input variables in drought assessment and forecasting models.

2.4 Drought forecasting

The terms short, medium and long-term forecasting have been used in drought studies as indicators of lead time in months of future drought. In most of the drought forecasting research, 1 to 3 months lead time is considered as short-term forecast. The medium to long-term drought forecast is lumped into one category of 4 to 12 months lead time (Mishra *et al*., 2006; Cassiamani *et al*., 2007; Belayneh, 2012). Forecasting of short-term drought conditions is useful for monitoring the effect of drought on agricultural systems. Under the short-term drought forecasting soil moisture and crop water stress may be defined especially during growing seasons. On the other hand, forecasting medium and long-term droughts helps to understand the overall effect of drought on water resources at basin and regional scales. The medium to long-term forecasting is critical in water resources management. It may be used for drought risk management as emerging early warning systems in Kenya. The three categories of drought forecasting can be used to formulate long-term plans for sustainable management of water resources and agricultural systems. Droughts are likely to persist in river basins such as upper Tana, with varied projections. Some models have projected an intensification of drought events in some areas, although other models indicate a reduction in drought severity (GoK, 2012).

2.5 Drought mitigation

Numerous methods are used to alleviate, protect or reduce the impacts of drought on people's livelihoods. The methods are directed to alter the drought effect on water resources, agriculture, livestock and other basin resources. Some of the commonly used methods at the basin scale for drought mitigation include:

i) Drought monitoring: This involves gathering of information about drought in terms of water and rainfall levels. Such information is used in formulation of early warning systems and design of other mitigation strategies

ii) Rainwater harvesting: It involves collection, conveyance and storage of water during wet conditions for use during the drought. The water may be stored in dams reservoirs or in storage tanks

iii) Desalination: This is a process of treating and removing dissolved salts in sea water or irrigation water within a river basin

iv) Recycling: It is the treatment and purification of wastewater through recycling, recovery and reuse in community water supply and irrigation systems

v) River engineering: It is the design and construction of river training channels, water canals or to divert and re-direct water to drought prone areas for water supply or irrigation

vi) Cloud seeding: A rainfall inducing substance is sprayed in the atmosphere to act as nuclei on which clouds form and generate precipitation

vii) Development of Crop varieties: This is the development of new plant varieties that can tolerate drought with significant crop yields

viii) Development of farming technologies that cope with drought: It is the development and use of farming technologies that have high water use efficiency and crop production and animal husbandry such as drip and greenhouse systems

ix) Reduction in outdoor water use and wastage: This involves the regulation and control of water wastage via the use of sprinklers, hoses, water containers, pipes and other usage and maintenance activities

2.6 Drought assessment methods

Drought indices or models are used for assessment of occurrence and severity of droughts. The Drought Indices (DIs) were developed for specific regions using specific structures and forms of data input. There is limited information in the application of drought indices that combines both temporal and spatial drought evaluation at river basin scales. Drought has been assessed in terms of temporal and spatial domain using evapotranspiration mapping as illustrated by Eden (2012). There are two broad categories of drought indices; satellite based and the data driven drought indices (Belayneh and Adamowski, 2013).

2.7 Satellite based drought indices

The satellite Remote Sensing (RS) may be defined as the science and art of obtaining information of points, objects, areas or phenomena through analysis of data acquired by a sensor, which is not in direct physical contact with the target of investigation (Sayanjali and Nabdel, 2013). The RS provides an aerial view of land, water resources and vegetation cover. This technique gives a

spatial and temporal context of assessing drought and has the ability to monitor vegetation dynamics over large surface areas. Currently, there is a considerable interest in collecting remote sensing data at multiple time scales. Such data is used to conduct a near real time information management (Mulla, 2013). Examples of satellite drought indices are the Vegetation Condition Index (VCI), Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Water Supply Vegetative Index (WSVI) and Normalized Difference Drought Index (NDDI).

2.7.1 Vegetative condition index

The Vegetative Condition Index (VCI) is computed from an advanced accurate and high resolution radiometer radiance data. This data is usually adjusted to match land conditions, climate, and ecology and weather conditions. The index is used for drought detection and trend tracking. It can be used to determine the time of on-set of drought, intensity, duration and associated impacts on vegetation (Mishra and Singh, 2010). The main challenge with the use of VCI is that it is used during dry seasons and the areas under consideration should have significant vegetation cover. The VCI for a month *j* is computed from the relation:

$$VCI_j = \left(\frac{NDVI_j - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}} \right) \times 100 \quad (2.4)$$

Where;

VCIj = Vegetation condition index for month j

$NDVI_j$ = the NDVI for a month j

$NDVI_{min}$ = minimum NDVI for the period under consideration

$NDVI_{max}$ = maximum NDVI for the period

2.7.2 Normalized difference vegetative index

The Normalized Difference Vegetative Index (*NDVI*) is a satellite data driven index that is used to monitor ground vegetation which could be linked to drought conditions. The index can filter out green vegetation using Landsat Multispectral Scanner (MSS) digital data (Musaningabe, 2012). It is normally expressed as a function of the near-infrared and red bands given as:

$$NDVI = \frac{NIR - R}{NIR + R} \quad (2.5)$$

Where;

NIR = near-infra red band

R = the red band

The *NDVI* is the most commonly used satellite based index. One advantage of the *NDVI* is that it has distinct values ranging from -1 to 1 with zero taken as an approximate value denoting absence of vegetation. The negative *NDVI* values indicate a non-vegetative surface while values closer to 1 represent dense vegetation.

2.7.3 Normalized difference water index

The Normalized Difference Water Index (NDWI) is determined based on leaf water content and vegetative type. Its value ranges from -1 to +1. The higher the NDWI value, the higher the vegetative water content and the higher the proportion of vegetative cover. The values of NDWI are computed by processing the satellite data in which green and near infra red bands are used as per the relation:

$$NDWI = \frac{G - NIR}{G + NIR} \tag{2.6}$$

Where;

NDWI = normalized difference water index

G = green band

NIR = near infra red band

The NDWI is very sensitive to soil moisture content, vegetation cover and leaf moisture content (Tychon *et al*., 2007). Although NDWI is used for drought detection, it is sometimes affected by land cover and pests and diseases on vegetation. However, it has an advantage of detecting drought more effectively as compared to the NDVI (Gu *et al.*, 2007).

2.7.4 Water supply vegetative index

The Water Supply Vegetative Index (WSVI) is a drought indicator based on relationship between the NDVI and the land surface temperature. The higher the values of WSVI, the higher the moisture levels, canopy temperature and the lower the NDVI. On the other hand, lower values of WSVI give an indication of extreme drought. The WSVI values range from -4 for extreme drought to +4 for highly moist conditions (Luke *et al*., 2001). The values of WSVI are obtained by analyzing the effect of vegetation on the reflection of red, near infra red and thermal bands. This index is more effective in drought detection under the conditions when the NDVI is greater than 0.3. Combining the WSVI and the NDVI in drought detection provides a more sensitive approach and better results (Jain *et al*., 2010).

2.7.5 Normalized difference drought index

The Normalized Difference Drought Index (NDDI) is used for detection of drought by combining the outputs of NDWI and NDVI derived from satellite data. The values of the two indices decrease with decrease in slope gradient of the cumulative precipitation. However, NDDI values decrease more abruptly during dry period than the NDVI. Thus, NDDI is more sensitive to water content and it is a better index for drought detection than NDVI. The NDDI has been noted to detect drought conditions on grassland than NDVI (Gu *et al*., 2007). The NDDI values can be computed from the following relation:

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI} \tag{2.7}$$

Where;

NDDI = Normalized difference drought index

NDVI = normalized difference vegetative index

NDWI = normalized difference water index

While the satellite based drought indices can be used to detect the on-set, intensity and duration of drought, their limitations and advantages need to be examined. These indices are limited to the areas with significant vegetation cover. Thus, in case the vegetation is infested by the pests and diseases, such indices can give misleading results. In addition, the indices are difficult to harmonise their drought characterization in terms of magnitude as they result from analyzing different bands of the satellite imageries.

2.8 Data driven drought indices

The Data Driven Drought Indices (DDDI) use a single or a combination of hydro-meteorological variables as input parameters to assess drought intensity, duration, severity and magnitude. Some of the data driven indices as reported by Belayneh and Adamowski (2013) include; the Standardized Precipitation Index (SPI), Palmer Drought Severity Index (PDSI), Surface Water Supply Index (SWSI), Aggregated Drought Index (ADI), Effective Drought Index (EDI), Reclamation Drought Index (RDI), Crop Moisture Index (CMI) and Murger Index (MuI). These indices use different input data such as rainfall, temperature, catchment soil moisture content, snow water content, stream flow, storage reservoir volume, and potential evapo-transpiration (Zoljoodi and Didevarasl, 2013). However, the suitability of the indices and their testing for Kenyan conditions has not been adequate. Therefore, Kenya does not have generic indices for drought forecasting. Due to scanty

and lack of drought assessment indices that can be used for defining critical drought conditions in Kenya, this research assessed selected suitable indices.

Some of the most critical elements of drought which are used for the design of water storage systems to cope with drought impacts include; longest duration, largest severity, highest intensity and spatial and temporal variation of droughts (Sharma and Panu, 1997; Manikandan and Tamilmani, 2013). Drought duration, severity and intensity are fundamental characteristics of any drought event. Drought duration refers to any continuous period of the sequence with deficit, while intensity is the magnitude below a truncation level. Severity is the cumulative deficit below a truncation level during drought period and may mathematically be defined as the product of the drought intensity and duration.

For better understanding of drought characteristics, assessment of the influencing variables is paramount. The statistical analysis for stream flow and precipitation as drought variables include parameters such as: the mean, coefficient of variation, log-1 serial correlation coefficient and the probability distribution function of the sequence under study. The extreme values of drought which include; duration, severity and intensity may be modelled with reference to a certain truncation level. The truncation level is usually taken as the long-term mean of the drought variable (Dracup *et al*., 1980a; Bonacci, 1993).

2.8.1 Standardized precipitation index

The Standard Precipitation Index (SPI) was developed by Mckee *et al*. (1993) to quantify the rainfall deficit and monitor drought conditions within Colorado, USA. For calculation of SPI, long-term historical precipitation record of at least 30 years is integrated into a probability distribution function which is then transformed into a normal distribution function. The SPI requires less input data compared to most other drought indices and this makes it flexible for wide applications (Mckee and Edwards, 1997; Bacanli *et al*., 2008).

The SPI has several advantages which make it more applicable in many river basins. First, it requires only the precipitation as the input data. This makes it ideal for river basins that do not have extensive hydrological data records. Secondly, its evaluation is relatively easy since it uses precipitation data set only. Thirdly, it is a standardized index and this makes it independent of geographical location as it is based on average precipitation values derived from the area of interest. In addition, the SPI exhibits statistical consistency, and has the ability to present both

short-term and long-term droughts over time scales of precipitation variation (Belayneh and Adamowski, 2012). However, the SPI has some disadvantages in its use as a drought assessment tool. First, it is not always easy to find a probability distribution function to fit and model the raw precipitation data. Secondly, most river basins do not have reliable time-series data to generate the best estimate of the distribution parameters. In addition, application of SPI in arid and semi-arid lands of time-series of less than three months may give inaccurate values.

To overcome the challenge of simulating and modelling the data for SPI outputs, application of different probability distribution functions may be employed. These include the Gamma, Pearson type III, Lognormal, Extreme Value and Exponential distribution functions (Cacciamani *et al*., 2007). However, the Gamma probability distribution function is preferred in hydrological studies. In hydrology, it has an advantage of fitting only positive and zero values since hydrological variables such as rainfall, and runoff are always positive or equal to zero as lower limit values (Markovic, 1965; Aksoy, 2000). The Gumbel and Weighbull distribution functions are used for study of extreme hydrological variables. The Gumbel distribution function is used for frequency analysis of floods, while the Weibull distribution function is used to analyse low flow values observed in rivers (Bulu an Aksoy, 1998).

Although the SPI can be used to present significant drought conditions within a river basin, identification of key dry periods requires an analysis of data for time scales greater than 6 months. This is because the high frequency of SPI values at shorter time scales conceal the critical dry periods. For time scales shorter than 6 months, there is insignificant autocorrelation while for time scales greater than 6 months, the autocorrelation increases significantly (Awass, 2009).

2.8.2 Palmer drought severity index

The Palmer Drought Severity Index (PDSI) was developed based on a criterion for determining the beginning and end of drought or wet period spell (Palmer, 1965; Wang, 2010). It is a simple monthly water balance model which requires rainfall, temperature and catchment soil moisture content as input parameters. This index applies a concept of supply and demand over a two-layer model. In this concept, the difference between the quantity of precipitation needed to maintain a natural water balance level and the actual precipitation is determined. The index does not consider stream flow, reservoir water balance, and other hydro-meteorological variables that influence the drought (Karl and Knight, 1985; Yan *et al*., 2013).

Several coefficients which are calculated to define local hydrological characteristics influenced by precipitation and temperature are calculated for use in PDSI. These coefficients depend on soil water capacity of the principal layers. The original PDSI has been modified to yield Palmer Hydrological drought Index (PHDI). The original PDSI does not take into account the human-induced impacts on water balance such as irrigation. However, the new version is a model mainly for evaluation and monitoring of water supply. The model has been applied on a number of catchments for detecting and planning of drought relief programmes (Loucks and Van Beek, 2005).

The PDSI has some limitations or disadvantages as a drought index. In some regions, the PDSI assumes that all the precipitation is rain. This may sometimes give misleading results in regions which experience winter season and also on high elevation areas. In addition, it under-estimates runoff since it assumes that overland flow occurs after all soil layers have been saturated. The other disadvantage is that the PSDI responds slowly to developing or ending of a drought event (Mishra and Sigh, 2010). Lastly, the original model is more suitable for agricultural drought than hydrological drought based on the applied time series.The original PDSI has some advantages and disadvantages (Narasimhan and Srinivasan, 2005). The major advantages of the original PDSI include:

i) The two indices provide decision makers with measurement of abnormality of recent weather condition for a basin or region
ii) It provides an opportunity to place current drought condition on a historical perspective
iii) It has the capacity to express historical drought conditions on spatial and temporal domain

The disadvantages of the PDSI are that:

i) The index uses two-layers in water balance computation and this is an over-simplification which leads to inaccurate values
ii) Potential evapo-transpiration (PE) in PDSI is computed based on Thornthwaite method which is a poor method of estimating the PE
iii) The original PDSI considered coarse resolution of land use and land cover parameters of 700-100,000 km^2 yet the land use changes within such a large area may be great

2.8.3 Surface water supply index

The Surface Water Supply Index (SWSI) was developed in Colorado USA, as an indicator of surface water or moisture levels (Shafer and Dezman, 1982). The index requires input variables which include; snow water content, stream flow, rainfall and storage reservoir volume (Castano, 2012). Normally the snow water content, rainfall and storage reservoir volume are used for

computing the SWSI values for winter season. However, during the summer season, stream flow substitutes snow water content. At a basin scale, the SWSI values are determined from monthly catchment average values of rainfall, reservoir volumes, snow water content and stream flows measured at stations within the catchment. One of the advantages of the SWSI is that it gives a representative measurement of surface water supplies across the river basin.

The SWSI is unique in specific basins or regions. It requires long term record data for its calibration and thus may be limited in basins that lack sufficient data. Another limitation of the SWSI is that any additional change in the water management within a basin calls for modification of its algorithm. The change may be due to an addition of new water reservoirs and flow diversions that based on their weights, require to be accommodated in the algorithm (Barua, 2010). Thus, it is difficult to have a homogeneous time series of the index for several basins.

2.8.4 Aggregated drought index

The Aggregated Drought Index (ADI) is used for determination of three categories of drought; hydrological, agricultural and meteorological droughts. In the use of each category, the specific drought is determined by selectively inserting input variables required into the model. This index uses rainfall, stream flow, potential evapo-transpiration, soil moisture content, snow water content and reservoir storage volume as input data (Keyantash and Dracup, 2004).

The Principal Component Analysis (PCA) is used as a numerical method for construction of ADI using appropriate input data sets. The PCA is used to transform spatially correlated data series from a basin into two sets of orthogonal and uncorrelated functions. The principle components are used to express the original *p*-variable data set in terms of uncorrelated component Z_j, where $1<j>p$. The *p*-model is used where the analysis explains temporal fluctuations of the input variables (such as precipitation, steram flow, reservoir levels, soil moisture content, temperature and evapotranspiration) of a basin. The calculation of the principle components involves the construction of a *pxp* symmetric correlation matrix C_x. The matrix gives the correlation between the original data where *p* is the number of variables. This matrix is expressed using the relation:

$$C_x = E \times \left\{ (x - u_x)(x - u_x)^T \right\} \tag{2.8}$$

Where;

C_x = correlation matrix

x = vector observation data

u_x = mean value of x

T = the transpose matrix

The correlation matrices developed undergo the PCA through the application of Eigenvectors. The Eigenvectors are unit vectors that establish the relationship between the principle components and standardized data. A unit vector may be derived from the relation:

$$Z = X \times E \tag{2.9}$$

Where;

Z = n x p matrix of principle components

X = n x p matrix standardized observation data

E = p x p matrix of eigenvectors

The first PC to represent ADI is determined and normalized by use of its standard deviation function defined by:

$$ADI_{i,k} = \frac{Z_{i,l,k}}{\sigma_k} \tag{2.10}$$

Where;

$ADI_{i,k}$ = ADI value for month k in year i

$Z_{i,l,k}$ = the first PC for month k in year i

σ_k = the sample standard deviation over all years for month k

To determine ADI thresholds, the empirical cumulative distribution of the ADI values given in Equation (2.10) are constructed. The ADI thresholds are then calculated using empirical cumulative distribution function and used to classify drought conditions based on the specified thresholds as summarized in Table 2A (Appendix A).

2.8.5 Deciles index

The deciles index was developed by Gibbs and Maher (1967) and has found a wide application in some regions such as Australia (Morid *et al.*, 2007). The Deciles Index uses long term monthly rainfall records where the records are ranked from the highest to the lowest and then a cumulative frequency distribution constructed. This distribution is then partitioned into ten sections called deciles. One major limitation in using the Deciles Index approach is that it requires long-term rainfall records of 30-50 years if accurate calculations are to be realised (Hayes, 2003). The first decile is the precipitation value not exceeded by the lowest 10% of all precipitation values within

the period under study. This is followed by the second decile that falls between the lowest 10 and 20% in that order. By comparing the amount of precipitation in a certain period with a long term cumulative distribution of precipitation amount in the mentioned period, the severity of the drought can be assessed. Although the data driven indices have been used in other basins in the world, they cannot directly be applied in Kenyan river basins before their calibration and validation are done.

2.9 Drought forecasting models

Development in forecasting and early warning of the drought phenomena is increasingly being applied in many regions in the world. For instance, drought forecasting in Kenya has previously been based on rainfall prediction. To forecast rainfall, the prevailing and expected sea surface temperature anomalies (SSTAs) over the Pacific, Indian and Atlantic oceans are used. The factors of the SSTAs are assessed by applying various tools such as ocean atmosphere models, statistical models, satellite derived information and expert interpretation. The onset, cessation and distribution of rainfall are derived from statistical analysis of previous years that exhibit same characteristics as that of the year under consideration (GoK, 2014). Drought forecasting is being done to help mitigate consequences of drought on vulnerable river basins. Different drought modelling and forecasting techniques are in use today. Some of the commonly used drought forecasting models include; Seasonal autoregressive integrated moving average model (SARIMA), Adaptive Neuro-fuzzy inference system, Markov chain model, Log-linear model and Artificial Neural Network (ANN) model.

2.9.1 Seasonal autoregressive integrated moving average model

The Seasonal Autoregressive Integrated Moving Average model (SARIMA) model is a time series tool. Time series events reoccur in every given number of observations (Chatfield, 2003). For monthly measurements, the recurrence over a year of twelve months, it is expected that the recurring value (x_t) will depend on values that are based on annual lags. These lags are defined by x_{t-12} or x_{t-24}. It may be influenced by recent non-seasonal values. The model has been generalized to deal with seasonality as defined by the relation given as:

$$u_t = x_t - x_{t-\omega} \tag{2.11}$$

Where;

u_t = seasonal value representing seasonality

ω = the period for monthly series, typically of multiples of 12

x_t = recurring value

To achieve stationery conditions, the seasonal difference can be repeated many times defined by *D*. For instance, if *D* = 2 and *D* =12, then the following function which is called SARIMA model results to:

$$\omega_t = u_t - u_{t-12} = (x_t - x_{t-12}) - (x_{t-12} - x_{t-24}) = x_t - 2x_{t-12} + x_{t-24} \quad (2.12)$$

Where;

x_t = recurring value

$x_{t-\omega}$ = annual lag

u_t = seasonal value representing seasonality

2.9.2 Adaptive Neuro-fuzzy inference system model

The Adaptive Neuro- fuzzy logic approach was developed by Zadeh (1965). It is a linguistic uncertainty function that applies the fuzzy inference system (FIS). The adaptive Neuro-fuzzy inference system uses a combination of Artificial Neural Networks (ANNs) and Fuzzy Inference System (FIS), and has greatly been used to come up with engineering solutions. The term fuzzy influence system is a principle that comprises three conventions (Firat and Gungor, 2008) given as:

i) A Rule base that consist of fuzzy 'if-then rules' incorporated in their algorithms
ii) A data-base defining the membership function that converts input value into a value between 0 and 1
iii) An influence system combining fuzzy rules to generate system results

2.9.3 Markov chain model

The Markov chains have greatly been used in stochastic characterization of drought (Cancelliene and Salas, 2004). For instance, an early warning system using Markov chain model in conjunction with PDSI based on probabilistic severity, duration and return period of drought may be developed (Shatanawi *et al*., 2013). Drought has also been characterized in terms of probabilistic occurrence by combining Markov chain model with SPI for short term drought prediction within a period of 1 to 3 months lead time (Paulo *et al*., 2005; Paulo and Pereira, 2007; 2008). The Markov chain model has two main applications; modelling stochastic characteristics of drought and forecasting future series of drought using historical data sets.

The Markov chain model exhibits a discrete stochastic process where a drought state (x) at a future time step (t+1) is dependent upon the present state x_t and independent of previous states X_{t-1}, X_{t-2},..., X_{t-n}. If a system of *n* states is considered, then the relation applies:

$$X = \{S_1, S_2, ..., S_n\} \tag{2.13}$$

Such a system can be transformed from S to S^2, S^3..., S^n according to specified transitional probabilities P_{12}, P_{13}, ..., P_{1n} or remain at state S_1 with a transitional probability of P_{11} (Shatanawi *et al.*, 2013). Therefore, the P_{ij} may be used to denote transitional probabilities from S_i to S_j. The P_{ij} can be represented in form of *nxn* matrix. Entries of such a matrix defined as P may be computed from a number of transitions n_{ij} from state *i* to the next state *j* using the relation:

$$P_{ij} = \frac{n_{ij}}{\sum_{i=1}^{k} n_{ij}} \tag{2.14}$$

Where;

p_{ij} = the nxn matrix

n_{ij} = the entries of the *P* matrix

The following summation holds true for the matrix

$$\sum_{i=1}^{k} P_{ij} = 1 \tag{2.15}$$

The transitional matrix at any given time step is calculated using the function:

$$P^{t+1} = P^t \times P^{t+n-1} \tag{2.16}$$

Where;

P^{t+1}= transition matrix at any given time

P^t = transition matrix at initial time

P^{t+n-1} = transition matrix of previous time step

The Markov chain attains a steady state after several time steps. It is thus possible to define a stationery matrix π as Eigenvector of P^t using the relation:

$$\pi = P^t \times \pi \tag{2.17}$$

Since π_j is a stationery probability for state *j*, thus

$$\sum_{j=1}^{k} \pi_j = 1 \tag{2.18}$$

The persistence and recurrence time can be presented using two main terms of the Markov chain. The first is the probability that the system will retain the same state in a subsequent time step. This is called persistence. The persistence probability P_r is defined using the relation:

$$P_r = \sum_{j=1}^{k} \hat{P}_{ij} \times \pi_j \tag{2.19}$$

On the other hand, the recurrent time is defined as the average time for a system to transit from a certain state *j* and then back to the same state *j* and is computed using the function:

$$t_{ij} = \frac{1-\pi_j}{(1-\hat{p})\times\pi_j} \tag{2.20}$$

The time required for a system to be transformed for the first time from state *i* to *j* is called first passage time (t_{ij}) and is computed using the relation:

$$t_{ij} = 1+\sum_{k\neq j} P_{ik}\times t_{kj} \tag{2.21}$$

2.9.4 Log-linear model

The log-linear model was developed in 1990 and can be used on Poisson-distributed data sets. It is a linear model that can fit in Poisson distribution function. The model is an extension of two dimensional contingency tables where the correlation between two or more discrete and categorical variables is determined by getting the natural logarithm of frequency entries in a contingency table. A contingency table is a type of table in a matrix format that is used to display a frequency distribution of variables. The table can provide basic information on interrelationship between two variables and the interactions between them. The model has been used in forecasting drought in various regions in the world including catchments in Portugal where a twelve month data within SPI was modelled (Moreira *et al*., 2008).

2.9.5 Artificial Neural Network models

The Artificial Neural Network (ANN) model is an information processing system developed with a structure and operation similar to that of a human brain (Maier *et al*., 2010). The model has been improved over time by use of different calibration techniques. With sufficient amount of data and complexity, the ANN model can be adapted to establish any correlation between series of independent and dependent variables (Luk *et al*., 2000). The ANNs have some advantages (Morid *et al*., 2007; Tran *et al*., 2009; Mustafa *et al*., 2012; Beale *et al*., 2014) which include:

i) the ability to process information based on their dynamic response to external input
ii) the ability to capture numerous kinds of relationships including non-linear functions which are not usually detected by other techniques
iii) the ability to provide effective analytical techniques in modelling and forecasting systems
iv) the ability to model dynamic/stochastic time series variables in Water Resources Engineering
v) the ANNs to processes large and complex data sets, including that of drought forecasting

2.10 Description of ANN model

The ANN model processes information through an elaborate network of neurons that are linked together. It simulates outputs based on certain inputs by a working principle resembling that of human brain where in the human brain; the neuron receives a set of input signals and generates outputs. The nervous system of human beings is represented by a number of architectural structures that range from simple to complex structures. Whether the structures are simple or complex, the systems consist of neurons or neural cells as the chief building blocks as shown in Figure 2.3.

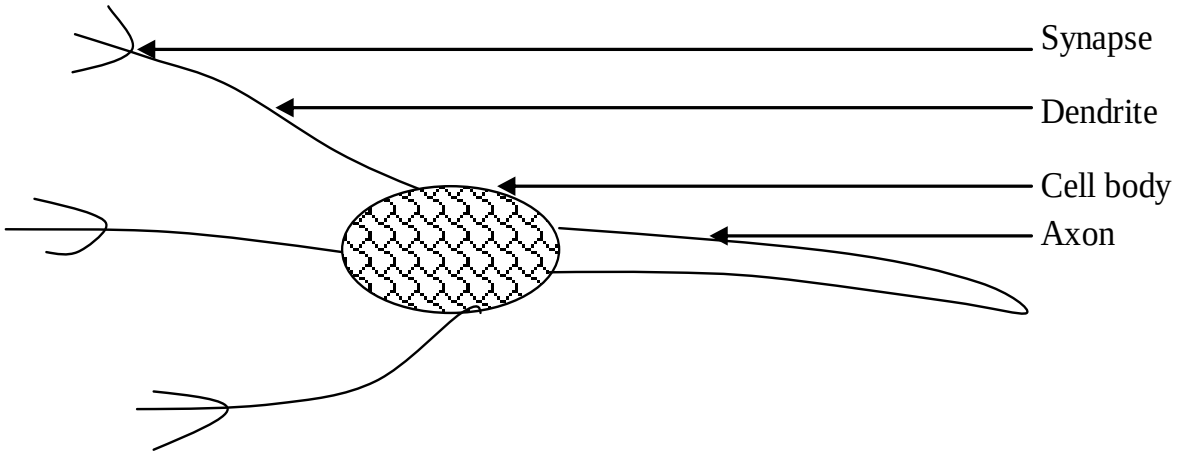


Figure 2. 3: Fundamental structure of a biological neuron

The neuron has four main parts; the cell body, synapse, dendrites and the axon. A neuron receives an input signal and transmits it to the cell body where it is processed and then an output is produced. In a biological neuron, both the inputs and the outputs are electrical signals. The input signals are normally passed to the neuron through the dendrites while the output signal is released through a single channel called axon. On every dendrite contact point, there is a part called synapse which works like a valve to open or close for signal flow. This synapse allows the input signals to pass through or stops some signals based on the conditions of operation of the system (Mishra and Nagarajan, 2012).

The ANN model is similar to a biological neuron in that it has multiple input channels, data processing unit, and output channels called dendrites, cell body and the axon respectively as represented in Figure 2.4. The input signals ($X_1, X_2, \ldots, X_p$) are passed to the neuron through the dendrites that represent different input channels. Each channel has its own weight referred to as

connection weight denoted as $W_1, W_2, \ldots, W_p$. The weights are very critical since they allow for collection and processing of signals based on their magnitude and effects on input functions. If a weight function gives a non-zero value at the synapse, it is allowed to pass through the cell body. Otherwise, if it has a value of zero, it is not allowed to pass the cell body. All the conveyed signals are normally integrated by summing up all the input.

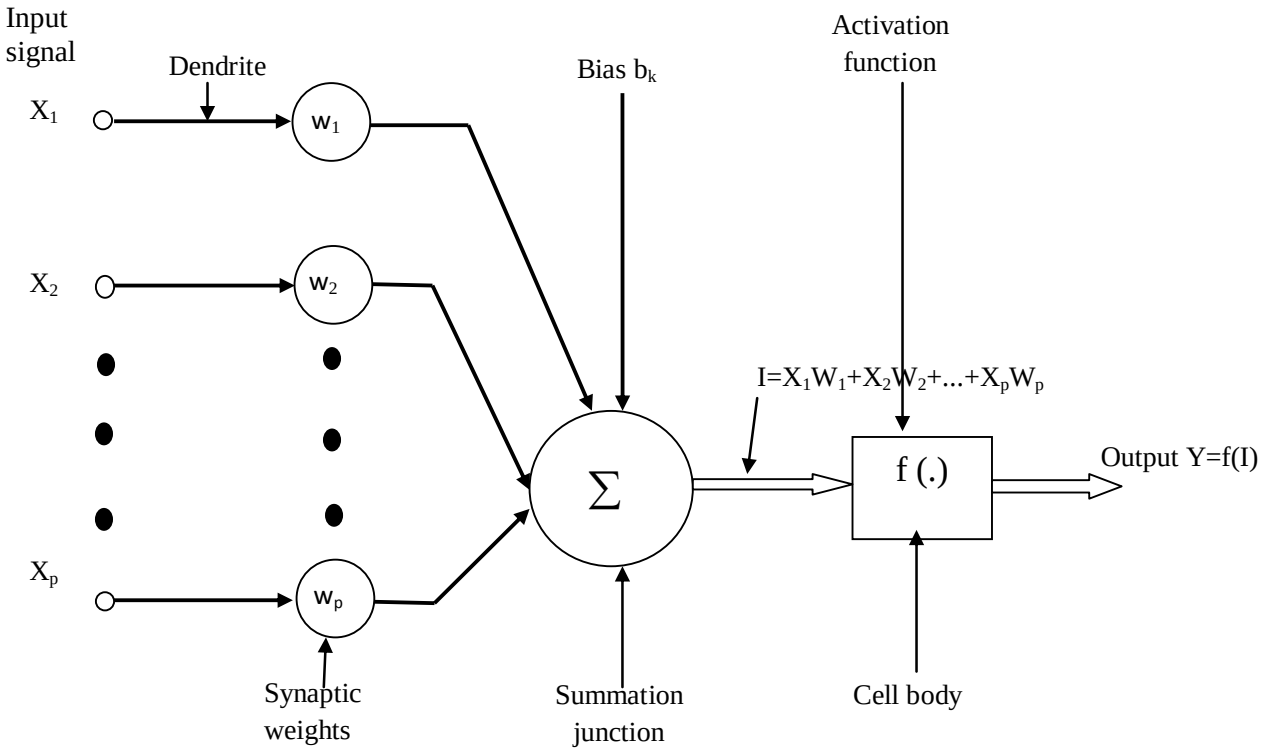


Figure 2. 4: Fundamental parts of a typical neural network

This is achieved by application of a mathematical model referred to as activation function, within the cell body to generate an output signal. According to Barua (2010), the relationship between the input and output signal within an ANN model is represented using the function given as:

$$Y = f(I) = f\left(\sum_{i}^{p} W_i X_i + b_k\right) \tag{2.22}$$

Where;

X_i = the input signal i

W_i = the weight attached to the input signal i

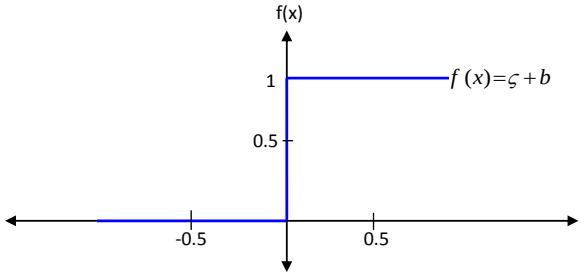
P = the number of input signals

b_k = the bias at the cell of the body

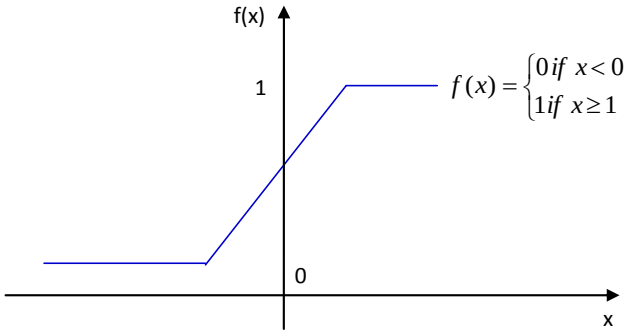
Y = the output

f= activation function

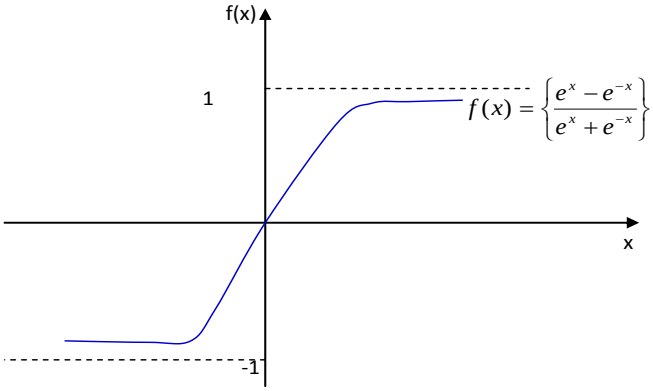
Numerous activation equations or functions can be used within the neurons. The most common functions used in the ANN models include; the step-function, non-linear sigmoidal, hyperbolic tangent and linear activation functions (Mishra and Desai, 2006; Maier *et al*., 2010). These functions are represented in Figures 2.5 *a*, *b*, *c* and *d*.



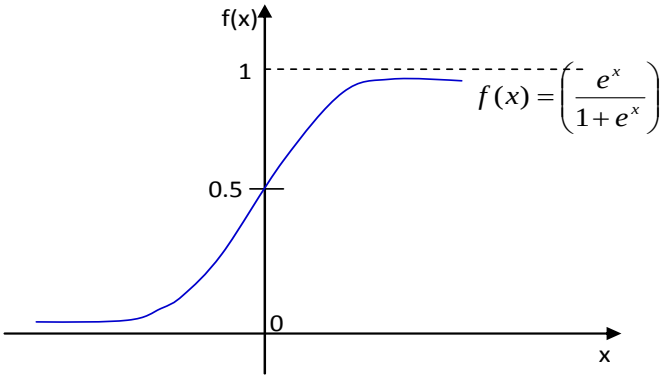
a) ANN step function



b) ANN non-linear sigmoid function



b) **Hyperbolic tangent function**



d) Linear activation function

Figure 2. 5 (a-d): Types of ANN activation functions

2.10.1 Classification of ANN model architectures

Numerous ANN model architectures have been developed and applied in drought research. These ANN architectures are grouped into three broad categories; feed forward, recurrent and hybrid networks. The information flow in feed forward network propagates only in one direction. It moves from input layer to the output layer. The feed forward ANN architecture is further subdivided into Multilayer Perceptron (MLP), Single Layer Feed Forward (SLFF), Support Vector Machine

(SVM), Generalized Regression Neural Network (GRNN), Radial Basis Function (RBF) and Neuro-fuzzy (NF) networks as given in Figure 2.6.

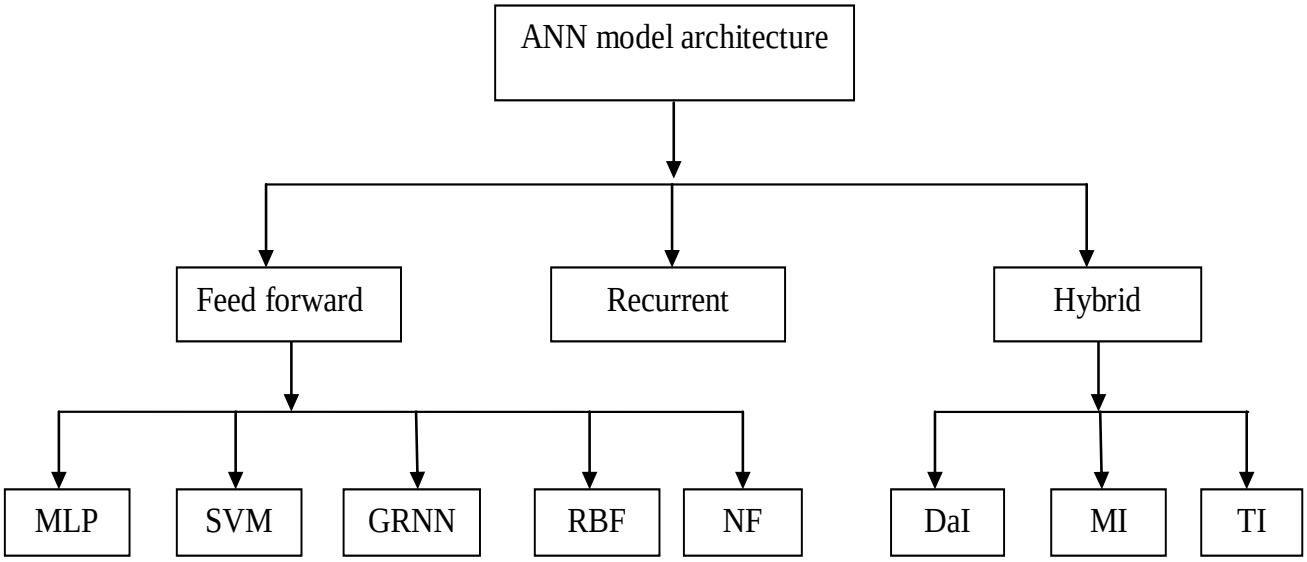
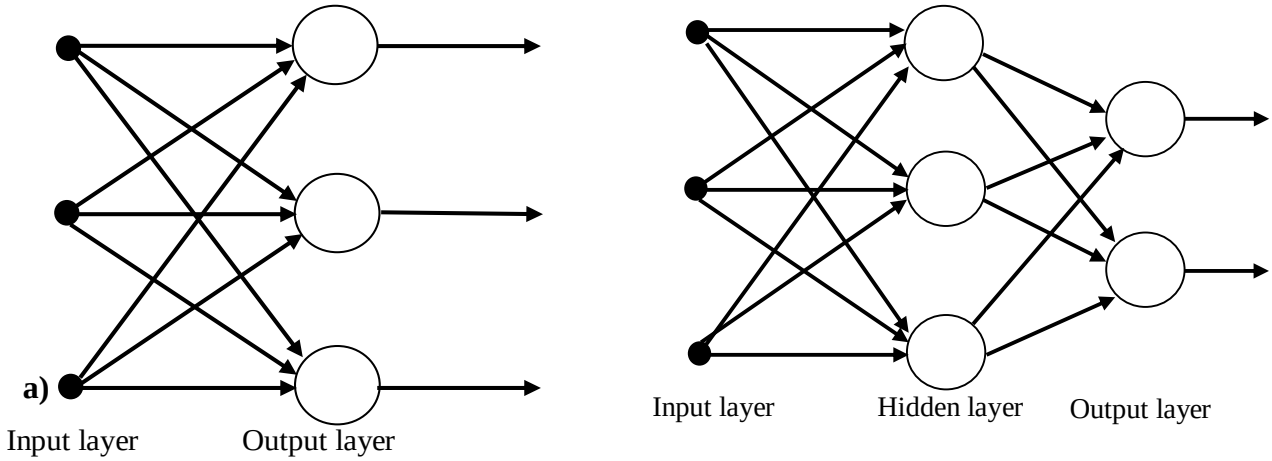


Figure 2. 6: Classification of ANN model architecture

The single-layer network has only one input layer that links directly to the output layer (Figure 2.7a). In the multilayer ANNs, one or more hidden layers are found between input and output layers (Figure 2.7b). By addition of one or more hidden layers, the network can model more complex non-linear and linear functions (Chai, 2010).

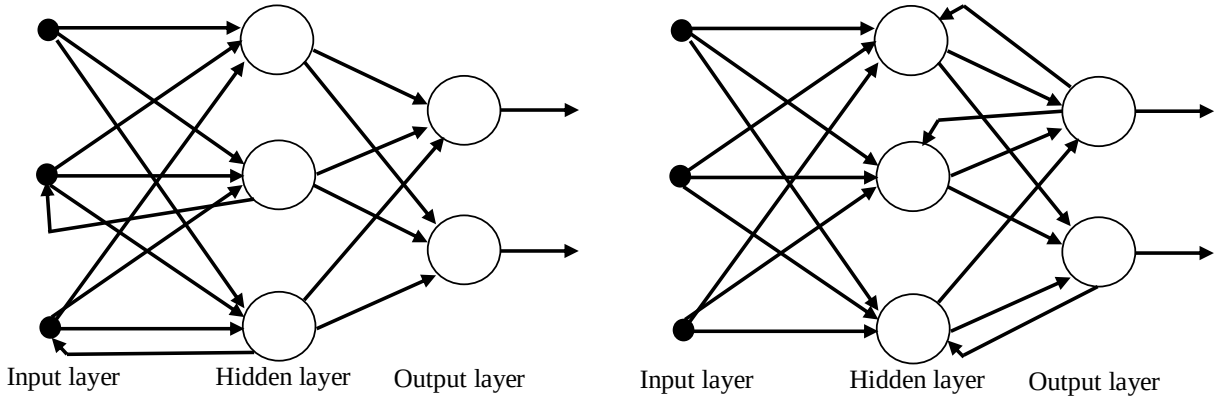


a) Single-layer feed forward ANNs **b) Two-layer feed forward ANNs**

Figure 2. 7 (a and b): Feed forward artificial neural networks

For the recurrent networks, information may propagate in forward and backward directions through feedback loops. In this network, the output layer neurons may transmit back the output to input and/or hidden layer neurons. At least one feedback loop must exist in the network. The hidden

neurons unit activation or the output values are fed back into the network as inputs (Figure 2.8). These models exhibit a dynamic characteristic when modelling data that depends upon different temporal and spatial resolutions (Timm *et al*., 2006). The feedback links layers whose state varies with time, and has adjustable weights. This makes the neuron to depend on both the current input signal and also the previous state of the neuron (Chiang *et al*., 2004).



a) Hidden neuron unit activation feedback **b) Output activation feedback**

Figure 2. 8(a and b): Types of recurrent ANNs

The third hybrid architecture network is subdivided into Data Intensive (DaI), model intensive (MI), and Techniques Intensive (TI) networks normally applied to solve complex modelling of hydrological and environmental systems.

2.10.2 Drought forecasting using ANN models

The ANN modelling has been greatly used for drought forecasting in the world (Ochoa-Rivera, 2008; Cutore *et al*., 2009). For instance, it has successfully been used in India to forecast drought in Kansabati River Basin (Mishra *et al*., 2007). For the Kansabati basin, two ANN models were applied and included the Recursive Multi-Step Neural Network (RMSNN) and the Direct Multi-Step Neural Network (DMSNN). These were used to forecast SPI values across the river basin. The results showed that the RMSNN performed best in one month lead time forecasting while the DMSNN was the best in four month lead time forecasting (Mishra and Desai, 2006). The RMSNN and the DMSNN were also applied in forecasting of Palmer Hydrological Drought Index (PHDI) in Italy. For this case study, the findings showed no significant differences between the outputs of RMSNN and DMSNN when used to forecast the PHSI (Cutore *et al.*, 2009).

2.10.3 ANN data pre-processing

Prior to the use of the data within the ANN environment, it should be pre-processed. This is the initial data transformation processe that converts data into a form compatible with the ANNs (Demuth *et al*., 2009; Beale *et al*., 2014). This process involves data normalization or standardization and is applied to speed up the training process of the ANN. By normalizing the data, the effect of outlier entries are reduced. Normalization transforms the input features of data into the same range of values thus minimizing bias within the ANN.

There are numerous ways of normalizing the data. The most common one is the statistical standard deviation for each feature across input data. The usual way is to obtain a standard deviation and mean from a training data. The output of ANNs is always in the normalized data format. These values are then de-normalized to give meaningful output by reversing the normalization process (Chai, 2010).

2.11 ANN learning processes

ANNs function through a learning process. A learning process refers to the property of an ANN to possess processing units capable of changing its input and or output characteristics as a result of changing environment, values or levels based on historic data (Mustafa *et al*., 2012). These learning processes are categorized into three classes; supervised learning, unsupervised learning and reinforcement learning.

2.11.1 Supervised learning

In supervised learning, there is a set of training data that contains some input values or variables connected with the output values. The output values are commonly referred to as target values. Thus in supervised learning, both inputs and outputs are provided. The network processes the inputs and compares its resulting outputs with the target. Errors are then propagated back through the system, causing the network to adjust the weights which control the network. This ensures that the error is refined every time the weight is adjusted. The supervised networks are broadly classified into two; the feed-forward and feed-back neural networks. In these networks, the training data is used by learning algorithms such as Levernberg-Marquardt (LM), back propagation (BP) (Figure 14B; Appendix B), conjugate gradient (CG), perceptron, multi-layer perceptron (MLP) and generic algorithms. The purpose of a learning algorithm is to create neural network output perfect values for the training data. However, the mission of the algorithm is to give good values for input data that are from the real world and not from the training set.

There are several approaches that are used to effectively avoid over-fitting or under-fitting of data in supervised learning. These include early stopping, model selection, Jiltering, Weight delay and Bayesian estimation.

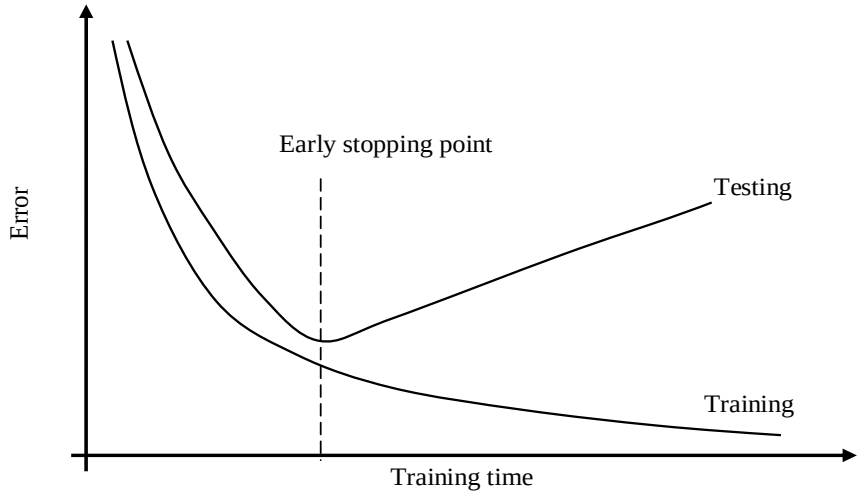


Figure 2. 9: Trend of training error and the point of over-fitting

In early stopping technique for instance (Figure 2.9), the MSE for both training and testing data sets decrease with time. However, at a certain point, the MSE linked to the testing data set begins to rise while that of training data set continues to decrease. At this point, further calibration will result in over-fitting the training data. Thus at such a point, the calibration is stopped in what is referred to as early stopping technique.

2.11.2 Unsupervised learning

This is a form of learning where networks are able to study on their own. They exhibit a kind of self training. In this learning, the network is provided with inputs but not the desired targets. The system decides what features it will use to group the input data in what is called adaptive training. In unsupervised learning, the network is able to learn and recognize patterns in data set whenever the data is introduced to the network. It is achieved though a competitive learning rule (Kalter, 2007). In this rule, a neural network with two layers; input layer and a competitive layer are created. The input layer receives the data introduced to it while the competitive layer has neurons that compete with others on the response to features in the input data. The competition is done in such a way that an input vector is compared with weight vectors that are connected to the competitive layer. Any neuron with weight vectors that closely match the input vector is considered

as the winning neuron. In other words, the network operates in accordance with what is called 'winner-take-all strategy' (Haykin, 1999).

2.11.3 Reinforcement learning

Reinforcement learning in ANNs is a technique acting by trial and error. In this approach, an agent can perceive a given state and perform certain actions. After every action, a numerical output is provided. The purpose of an agent is to maximize the total output it receives over time. Numerous algorithms are used in selecting actions in order to explore the environment and gradually build an approach that gives a maximum output. Such algorithms include model-based algorithm and mode-free algorithm (Yamada, 2011).

2.12 Purpose for ANNs learning process

ANNs are normally subjected to a learning process for two main reasons; classification and function approximation. To achieve the two, the networks are trained. Training a neural network in most cases is an exercise of optimizing a non-linear function. Several methods of non-linear optimization have been developed such as numerical analysis, operations research and statistical computing. However, there is no single best approach for such optimization. The methods are chosen based on characteristics of the problem being solved.

2.12.1 Learning for classification

Learning in ANNs is very useful in classification of information or outputs. In classification, the input is considered as a description of numerous objects recognized by the network. The outputs are normally studied by the network and then their class to which they belong is identified. Thus in the classification process, the target output cannot be organized along a useful band because each output of the network is a separate entity which is discrete from all the other outputs (Chai, 2010).

2.12.2 Learning for function approximation

ANNs have greatly been used for development and modelling of non-linear functions. The ability of an ANN fitting a non-linear function when provided with input data may be demonstrated by use of an example. To demonstrate the fitting ability of ANNs, a set of training data is generated (Table 2.1a) using the function:

$$y = x^3 + x^2 - x \tag{2.23}$$

Where;

y = the output (dependent variable)

x = the input (independent variable)

For such a case, a feed forward back propagation neural network with two layers is created. The transfer function between the input and the hidden layer is sigmoid function while the function between the hidden and output layers is linear. There are ten neurons in the hidden layer. The input of the neuron are x, x^2 and x^3 values while the target output is the y value. These input and output combinations are fed to the network to 'learn' to fit a function to the three inputs and one output. This learning process involves setting of weights and bias values within the neural network architecture. The numerical and predicted values of the y neural network exhibit deviation and can be illustrated using values given in Table 2.1b.

Table 2. 1(a and b): Data from function and predicted y values using neural network

a)

x	x^2	x^3	y
1	1	1	1
2	4	8	10
3	9	27	33
4	16	64	76
5	25	125	145
6	36	216	246
7	49	343	385
8	64	512	568
9	81	729	801
10	100	1000	1090

b)

Value of x	Actual y_a	Predicted y_p
1	1	0.95
2	10	9.8
3	33	32.71
4	76	75.3
5	145	145.44
6	246	245.72
7	385	385.25
8	568	568.08
9	801	800.05
10	1090	1090.36

From the information of predicted values y_p, it can be observed that the resulting values of y are very close to the actual values y_a as given in Figure 2.10.

2.13 Drought assessment and forecasting in river basins

Drought assessment and/or forecasting have been conducted in many river basins across the world. For instance, Barua (2010), developed a Non-linear Drought Index (NADI) for drought forecasting in Yarra river basin in Australia. In this study, average monthly values for rainfall, potential evapotranspiration and one stream flow gauge station were used used to reperesent the entire basin. The shortcomings and gap in this study was that an average value is not a good representative of the whole basin. In another study conducted by Belayneh (2012) on short and long-term standardized precipitation index (SPI) drought forecasts in Awash River basin, Ethiopia. In this study the forecasting ability of the SPI was evaluated using support vector regression (SVR) and

wavelet support vector regression (WA-SVR) neural networks. The results of the study indicated that WA-SVR networks had the best forecasting ability

A tool to assess hydrological drought occurrence in un-gauged catchments in Wabi Shebele river basin, Ethiopia was developed by Awass (2009). In this study, hydrological drought was extracted from reconstructed time-series monthly stream flow models. In another study, Castano (2012) conducted a study on drought monitoring at river basin and regional scale in Sicily, France. The SPI and SSI were compared and yet the two indices do not represent the same type of drught. Yang (2010) applied SPI, PDSI and SWSI to estimate meteorological, agricultural and hydrological droughts respectively in Upper Klamath River basin, California. In this study, drought for both the historical time period (1920-2009) and the future time period (2020-2090) was estimated.

Eden (2012) assessed drought by evapo-trasnpiration (ET) mapping in Twente, the Netherlands. The severity of droughts was quantified by the ET mapping for the period 2003-2010. The study used Evapotranspiration Deficit Index (ETDI) calculated using estimated ET in combination with reference ET from Penman-Monteith method. Evaluation of temperature and precipitation anomalies using SPI was included. Precipitation data from ground measurements were used to calculate SPI and compared with the ETDI. Additionally, Ntale and Gan (2003) assessed drought indices and their application in East Africa (EA). In their study, properties of three drought indices; PDSI, Bhalme-Mooley Index (BMI) and the SPI were analysed and modified. The findings from the study indicate that SPI is more appropriate for monitoring droughts in EA because it is more easily adapted to local climate, has modest data requirements, can be computed for any time scale and is easy to interpret.

Balint *et al* (2013) conducted a study on monitoring of drought with combined drought index in Kenya. In this study, NDVI, Precipitation drought index (PDI) and temperature drought index (TDI) were integrated into one Combined Drought Index (CDI). The CDI is statistical index that compares the present hydro-meteorological conditions with long-term average characteristics. A research by Onyango (2014) analysed meteorological drought in North Eastern province of Kenya. In this study, seasonal drought charateristics for the period 1960-2008 for the region were evaluated. The main characteristics addressed were drought severity, duration, frequency, persistence and probability of occureance. The study gave time series of these characteristics. From these studies, there is scanty research work on drought forecasting, spatial drought assessment and drought trend, and comparison of different drought indices.

2.14 Drought assessment and forecasting in the upper Tana River basin

Some studies on drought have been done in the upper Tana River basin. For instance Agwata *et al* (2014) conducted modelling of hydrological drought events in the baisn. In their study, five frequency distributions were fitted to drought severity and duration based on discharge data for different gauge stations located within the basin. The fitted distributions to two drought events that are severity and duration were Generalized Normal (GN), Generalized Extreme Value (GEV), Genaralized Pareto (GPA), Pearsson type III and Generalized logistic (GL). The findings show that the frequency distribution of the best fit for severity and duration was the GN. The poorest distribution for both severity and duration was the Pearsson type III distribution.

In another study, Agwata (2014) analyzed hydrological drought events using spatial characteristics of drought severity and duration based on discharge records. Drought severity and duration were extacted from discharge records using runs analysis techniques. The stream flow data was subjected to principal component analysis (PCA). The results showed that the drought severity and duration had distinct spatial patterns for the basin. The findings from the study gave anormalies of duration to the eastern and southern areas of the basin. The spatial patterns of drought severity portrayed zonal patterns reflecting differences in relief features in the eastern and western areas. Drought severity and duration were also estimated in four homogeneous regions of the basin (Agwata *et al*., 2015). The homogeneous regions were established using PCA. The findings of the study indicated that mean drought duration varied from 4 to 11 months while the standardized mean severity ranged from 0.63 to 3.89. The southern and eastern parts of the basin were found to more severe drought. However, there is limited research work on formulation of drought forecasting methods and/or tools at different led times, spatial aspects of drought as well as detection of drought trend for upper Tana River basin. In addition, applicability of drought indices, their performance in drought assessment as well as comparison of different drought indices for upper Tana River basin has not been well explored.

2.15 AquaCrop model

The aquacrop model was developed by the food and agriculture organization (FAO) in 2009 as a water-driven crop simulation model (Raes *et al*., 2009; Iqbal *et al*., 2014). The model is based on basic crop yield response to water algorithm. It is useful in various applications such as crop response to environmental changes, estimation of crop yield in the field, identification of

constraints limiting crop production and/or water productivity and maximization of evapo-transpiration (ET) water productivity (Casa *et al*., 2013).

2.16 Kriging interpolation technique

The spatial distribution of drought can be mapped using the Kriging interpolation technique using the point data as input variables. The accuracy of Kriging interpolation has been used in many studies and proved to be within acceptable levels (Robinson and Matternicht, 2006). The Kriging technique may be described using the various parameters and functions that are applied in the interpolation of values. The interpolation technique uses a parameter $z^*(x_0)$ called unbiased estimator where λ_i are chosen to satisfy the unbiased conditions and minimum variance (Kim and Valdes, 2002) in which the following relation is defined:

$$Z^*(x_0) = \sum_{i=1}^{N} \lambda_i \times z(x_i) \tag{2.24}$$

Then a kriging system comprising n+ 1 linear function with undefined values n+ 1 of the form $\lambda_i, ..., \lambda_n$ *given as*:

$$\sum_{j=1}^{N} \lambda_i \gamma_{ij} + \mu = \gamma_{i0},\ i = 1, ...N \tag{2.25}$$

$$\sum_{J=1}^{N} \lambda_i = 1 \tag{2.26}$$

Where;

μ = Lagrange multiplier derived from unbiased condition.

γ = semi-variogram function which is an indicator of spatial correlation of the recognized variables.

For instance, if z(x) is the observed value at station x, then the semi-variogram is estimated using the relation:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[Z(x_i) - Z(x_i + h)\right]^2 \tag{2.27}$$

Where;

N(h) = the number of observed pairs with interval distance of h. by application of a matrix, the kriging approach may be expressed as:

$$Ax = b \tag{2.28}$$

Where x is a vector of the unknown values, b is the right-hand side of Equation (2.28) and A is the coefficient matrix. A, b and x are defined as:

$$x = [\lambda_1 \ \lambda_2 . \ . \ . \lambda_N \mu] \tag{2.29}$$

$$b = [\gamma_{10} \ \gamma_{20} \ldots \ \gamma_{N0} \quad 1] \tag{2.30}$$

$$A = \begin{bmatrix} \gamma_{11} & \gamma_{21} & \gamma_{N1} & 1 \\ \gamma_{12} & \gamma_{22} & \gamma_{N2} & 1 \\ : & : & : & 1 \\ \gamma_{1N} & \gamma_{2N} & \gamma_{NN} & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \tag{2.31}$$

The frequency or percent occurrence of drought can be computed as the ratio of drought occurencs in each period to the total drought occurences in the same period and drought category (McKee *et al*., 1993; Asgari *et al*., 2014).

CHAPTER THREE
MATERIALS AND METHODS

3.1 Study area

The focus of this study was the upper Tana River basin with an area of 17,420 km^2 presented in Figure 3.1. The study area is part of the larger Tana River basin, the largest river basin in Kenya with an area of 126,000 km^2 (Jacobs *et al*., 2004; WRMA, 2010). The upper Tana River basin has forest land resources located along the eastern slopes of Mount Kenya and Aberdares range which have a critical role in regulating the hydrology of the entire basin (IFAD, 2012). The basin was selected because it is located within a fragile ecosystem that represents all agro-ecological zones of Kenya where water resource systems, hydro- power generation (Figure 29B; Appendix B) and food security are negatively impacted by frequent drought occurrences. In addition, it is the area that regulates the hydrology of the larger Tana River basin.

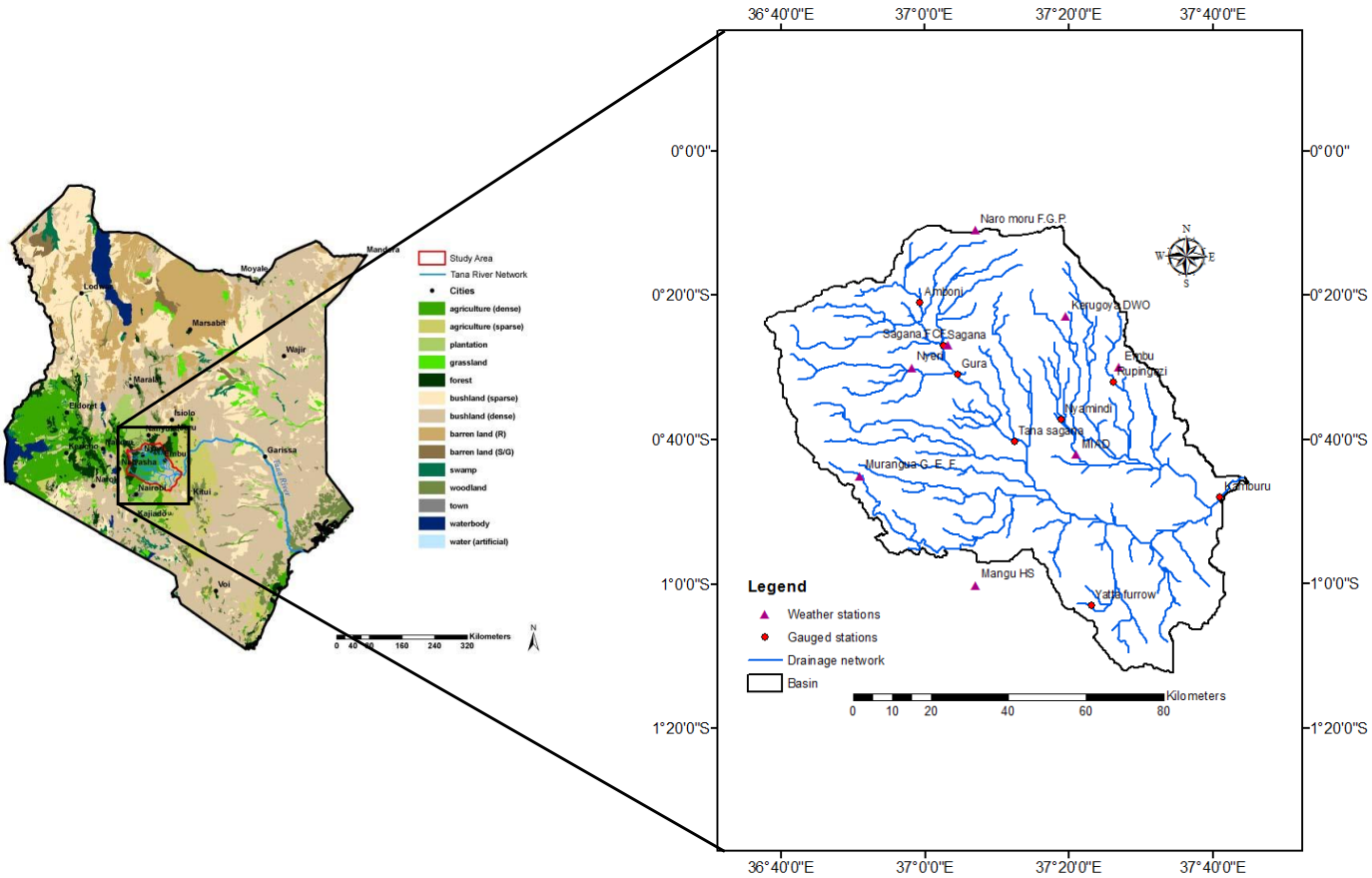


Figure 3. 1: Map of the upper Tana River basin

The upper Tana River basin lies between latitudes 00° 05' and 01° 30' south and longitudes 36° 20' and 37° 60' east. The basin is fundamental in influencing the ecosystem downstream (NEMA, 2004). It drains nine counties namely; Muranga, Nyandarua, Kiambu, Kirinyaga, Laikipia, Machakos, Nyeri, Embu and Kitui (WRI, 2011). The basin comprises major tributaries of Tana River whose total length from the source to the Indian Ocean is approximately 1,000 km (IFAD, 2012). The Tana River tributaries originate from the slopes of Mount Kenya and Aberdares range. The basin forms a principal resource in Kenya. For instance, the upper Tana River basin is critical in water supply, hydro-power generation and agricultural production.

The elevation of the upper Tana River basin ranges from approximately 730 m to 5,190 m above mean sea level (a.m.s.l.). These elevations are at Mount Kenya and adjacent to Kindaruma hydropower dam respectively which fall within the study area. The River basin exhibits heterogeneous soil types. For example, Andosols are the soil types which are predominant at higher elevations, Nitosols are at the middle elevations while Ferrasols and Vertisols are at the lower elevations (Jacobs *et al*., 1998).

Precipitation and temperature vary across the river basin. The Mount Kenya and Aberdares ranges receive approximate annual rainfall of 1800 mm (Otieno and Maingi, 2000). Within the middle elevation of 1200 to 1800 m a.m.s.l., the annual rainfall ranges from 1000 to 1800 mm while the lower elevations that are less than 1000 m, receive annual rainfall of 700 mm (Figure 7B; Appendix B). Although the basin receives significantly high rainfall amounts, it is characterized by seasonal rainfall fluctuations. This translates to seasonal variation of stream flow in Tana River. Generally the basin experiences a bimodal rainfall pattern (Figure 2 B; Appendix B) caused by inter-tropical convergence zone (Wilschut, 2010). The rain seasons are distributed in the months of March to June, and September to December as illustrated in Figure 25B of Appendix B. The precipitation is highly influenced by the orographic forces (Saenyi, 2002). The maximum and minimum mean annual temperature vary between 25.5 to 31.0°C and 21.0 to 24.0°C respectively (Mutua and Klik, 2007). The average catchment evapo-transpiration is around 500 mm in the summit area. Numerous land use types are found within the upper Tana River basin. The major ones include; forests, crop land and range land. The forests and tea plantations dominate the land use activities at the high elevations of the basin.

3.2 Assessment of spatial and temporal drought using selected DIs

This research applied selected drought indices as well as formulation and validation of a non-linear integrated drought index (NDI) that detects critical characteristics of drought for the upper Tana River basin. The drought indices that were used included; the Surface Water Supply Index (SWSI), Stream Flow Drought Index (SDI), Standardized Precipitation Index (SPI), Effective Drought Index (EDI), Palmer Drought Severity Index (PDSI) and Soil Moisture Deficit Index (SMDI). These indices were selected because each index was adopted for different drought types; meteorological, hyrological and agricultural drought types (Table 25A: Appendix A) as reuired in this study. In addition, the input data required was available for the river basin. To compute the drought indices, hydrometeorological data was first acquired as described in Section 3.2.1.

3.2.1 Hydro-meteorological data acquisition

Different data sets were used for this study to compute the DIs and the drought forecasting models. These data include; stream flow, storage dam levels, precipitation, potential evapo-transpiration, soil moisture content and temperature. The hydro-meteorological data from 1970-2010 (41 years) used in this study for defined hydrometric stations (Figure 3.1, Figure 8B; Appendix B). This data was selected because it was consistent for the period. Part of this data was available while the missing data was estimated for the hydrometric stations. The available data was on daily time step but had to be re-organized into monthly average time scales for all the variables to match with the data requirements for this study. The monthly stream flow and monthly flow data was obtained from the Ministry of Environment and Natural Resources, and Water Resources and Manageent Authority (WRMA). Data on dam levels was obtained from Kenya Electricity Generating Company (KenGen).

3.2.2 Stream flow data

There were fourteen hydrometric stations in the upper Tana River basin with complete and incomplete data records. However, only eight stations were selected for this study since they have sufficient long and reliable data for the period 1970-2010. The stations were in addition considered to be representative of the basin as they are located within the low, lower middle, middle and higher elevations for different agro-ecological zones. The station names and gauge identification (ID) numbers, their spatial locations are shown in Table 3.1 and Figure 3.1 respectively. In this study, only the Masinga dam levels were used because of the availability of long-term data records.

Table 3. 1: Stream flow gauge stations

S.No	Hydrometric Name	Gauge ID	Coordinates	
			Easting	Northing
1	Amboni	4AB05	36.989	-0.350
2	Sagana	4AC03	37.043	-0.449
3	Gura	4AD01	37.076	-0.517
4	Tana sagana	4BC02	37.207	-0.672
5	Yatta furrow	4CC03	37.361	-1.094
6	Nyamindi	4DA10	37.317	-0.621
7	Rupingazi	4DC03	37.438	-0.533
8	Kamburu	4ED01	37.683	-0.800

3.2.3 Precipitation data

In the upper Tana River basin, data from twenty four meteorological stations were obtained from the Ministry of Water and Irrigation. The meteorological data included precipitation, temperature and evaporation data. The data were then subjected to exploratory data processing. It was found out that only eight stations had reliable and sufficient data. Where the available data contained less than 20% data gaps, then these data were selected for computation the drought indices (DIs). The eight stations used in the study (Table 3.2) are located within the low (LE), lower middle (LME), middle (ME) and high (HE) elevations. The stations are located at different agro-ecological zones (Figure 13B; Appendix B) and sub-basins (Figure 15B; Appendix B) of the Upper Tana River basin.

Table 3. 2: Meteorological stations

S.No	Station name	Station ID	Coordinates		Elevation (m)
			Longitude (Degrees)	Latitude (Degrees)	
1	MIAD	9037112	37.350	-0.700	1246
2	Embu	9037202	37.450	-0.500	1494
3	Kerugoya DWO	9037031	37.327	-0.382	1598
4	Sagana FCF	9037096	37.054	-0.448	1234
5	Nyeri	9036288	36.970	-0.500	1780
6	Maragua G. E. F.	9036212	36.850	-0.750	2296
7	Naro-moru F.G.P.	9037064	37.117	-0.183	2296
8	Mangu HS	9137123	37.033	-1.100	1630

3.2.4 Consistency test of the hydro-meteorological data

A double-mass curve was fitted for the collected hydro-meteorological data to test for consistency/homogeinity. The homogeinity of stream flow time series data was conducted to detect for any possible errors resulting from the data measurements. In addition, homogeneity was used to

check for the fluctuations due to climate and weather changes. The cumulative total stream flow and precipitation were computed for each station and then plotted against the cumulative total of an adjacent station (Figures 3B and 4B; Appendix B). Any sudden change in the gradient of the double-mass curve was considered to indicate inconsistency in the data. Although there were some changes at some points on the curves for some stations, it was considered insignificant for the present study.

3.2.5 Filling in missing data

The continuity of the data records was first examined. It was found out that the gauge stations; 4AB05, 4AC03, 4AD01, 4BC02, 4CC03, 4DA10, 4DC03 and 4ED01 (Table 3.1) had continuous data records for, 35, 28, 33, 29, 33, 37, 29 and 36 years respectively. The meteorological stations; 9037064, 9037112, 9037031, 9137123, 9037202, 9037096, 9036288 and 9036212 (Table 3.2) had continuous data for 26, 28, 35, 32, 40, 35, 40 and 23 years respectively. The data for each station was partitioned into training and validation data sets comprising 70% and 30% respectively of the total continuously recorded data.

In this study, the ANN structure for each station was obtained by considering different input neurons for different time delays; *t, t-1, t-2,…,t - n,* in the input layer. The number of input variables was equal to the input neurons. The initial number of hidden neurons of the ANN model architecture was determined using the procedure adapted from Belayneh (2012) where the hidden layer neurons were initially set at $2n+1$ and n being the input neurons. The Hidden Neurons (HN) were then increased and decreased through trial and error technique for the data sets at each hydrometric station. This resulted to an output that was taken as the estimated variable. The ANN model at each station was trained using different structures as given in Tables 2A and 3A of Appendix A.

The output layer comprises neurons in all the networks that are equal to the following month's variable value (I_{t+1}). In this study, the Feed Forward Neural Network (FFNN) and Recursive Neural Network (RNN) were applied and tested in the model training. Initially three different training algorithms were applied to train the structures. These were the back-propagation (BP), Levernberg-Marquardt (LM) and Conjugate Gradient (CG) algorithms. From preliminary results, it showed that a three-layer feed forward neural network with different input and hidden neurons was superior in performance, and that the best results were also obtained using the LM training algorithm. Thus the best ANN structure of three-layer feed forward network based on LM training algorithm was

adopted for filling in of missing data in this study. The data was first normalized at each station before exporting it into the graphical user interface (GUI) of the MATLAB. This was done by applying the function given in Equation (3.1) which was adapted from Morid *et al.* (2007).

$$X_n = X_{\min} + \frac{(X_o - x_{\min})}{(x_{\max} - x_{\min})} \times (X_{\max} - X_{\min}) \quad (3.1)$$

Where;

X_n = normalized value

X_{min} = the selected minimum value for standardization

X_{max} = the selected maximum value for standardization

X_o = original value

x_{min} = minimum value present in the original data set

x_{max} = maximum value present in the original data set

All the input and output values for ANN were normalized to range between X_{min} equal to 0.1 and X_{max} of less than 1. According to Morid *et al*. (2007), the values of the X_{min} 0.1 and X_{max} of 0.9 perform best for drought indices such as SPI and EDI. Thus these values were adapted for this study. After normalization, the various drought forecasting ranges were determined.

For each of the ANN model run on the GUI of the MATLAB, the performance was evaluated based on the correlation coefficient R and Mean Square Error (MSE) criteria and the best model results were summarized in a tabular form in the Table 2A (Appendix A). The best ANN models were then adopted for filling any missing data for respective hydro-meteorological stations. The steps that were followed in filling the missing data are summarized in Figure 3.2

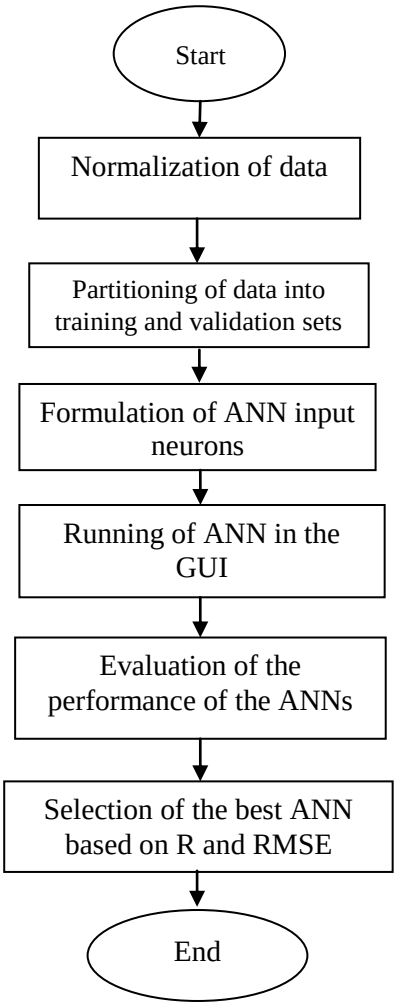


Figure 3. 2: Flow chart of the steps used in filling the missing data using ANN

3.2.6 Surface Water Supply Index

The input data for the Surface Water Supply Index (SWSI) were the monthly precipitation, stream flow, reservoir level and dam inflow. In the present study, the data was first summed up and normalized using the probability of non-exceedance. The probability of non-exceedance refers to the possibility that a random drought magnitude is less than or equal to a defined real value. The values of the SWSI were computed using various input parameters as shown in Equation (3.2):

$$SWSI = \frac{\left[\left(a \times PN_{rn}\right) + \left(b \times PN_{sf}\right) + \left(c \times PN_{rs}\right) + \left(d \times PN_{df}\right) - C_1\right]}{C_2} \tag{3.2}$$

Where;

$SWSI$ = modified surface water supply index (dimensionless)

PN = probability of non-exceedence in percent

rn = the rainfall (mm)

sf = stream flow (m^3/s)

rs = storage reservoir level (m)

df = dam inflow (MCM)

C_1 = 50 and C_2 = 12 (Shafer and Dezman, 1982)

The parameters *a, b, c, d* and *e* are the weights for each component and their sum must be equal to 1 as given in the expression:

$$a+b+c+d+e=1 \tag{3.3}$$

The weighting parameters *a, b, c* and *d* corresponding to the rainfall, stream flow, reservoir levels and dam inflows were estimated through a proportioning objective procedure (Table 4A; Appendix A). Such a procedure is better for SWSI determination since it computes the weighting parameters more accurately compared to the method developed by Shafer and Dezman (1982), where the values by these authors are based on assumptions. The probability of non-exceedance of monthly precipitation, stream flow, storage reservoir, dam inflow and the ground water level were computed using the relation:

$$PN=1-\frac{r}{n+1} \tag{3.4}$$

Where;

PN = probability of non-exceedence in percent

r = the rank of data arranged in ascending order

n = the number of years considered in the analysis

The integration of the rainfall (*rn*), stream flow (*sf*), storage reservoir volume (*rs)* and dam inflow into SWSI is summarized in Figure 3.2. The product of the respective time series weighted parameters and the probability of non-exceededance were computed to get the composite factor $SWSI_f$ which after normalization was used to get the surface water supply index *SWSI*. A regression plot between $SWSI_f$ and *SWSI (*Figure 24B; Appendix B), was then done to develop a simplified SWSI equation for the upper Tana River basin. In the normalization, the *SWSI* values were set at between -4.2 and +4.2 as minimum and maximum values respectively. The procedure that was followed to obtain the time series and spatial SWSI is summarized in Figure 3.3.

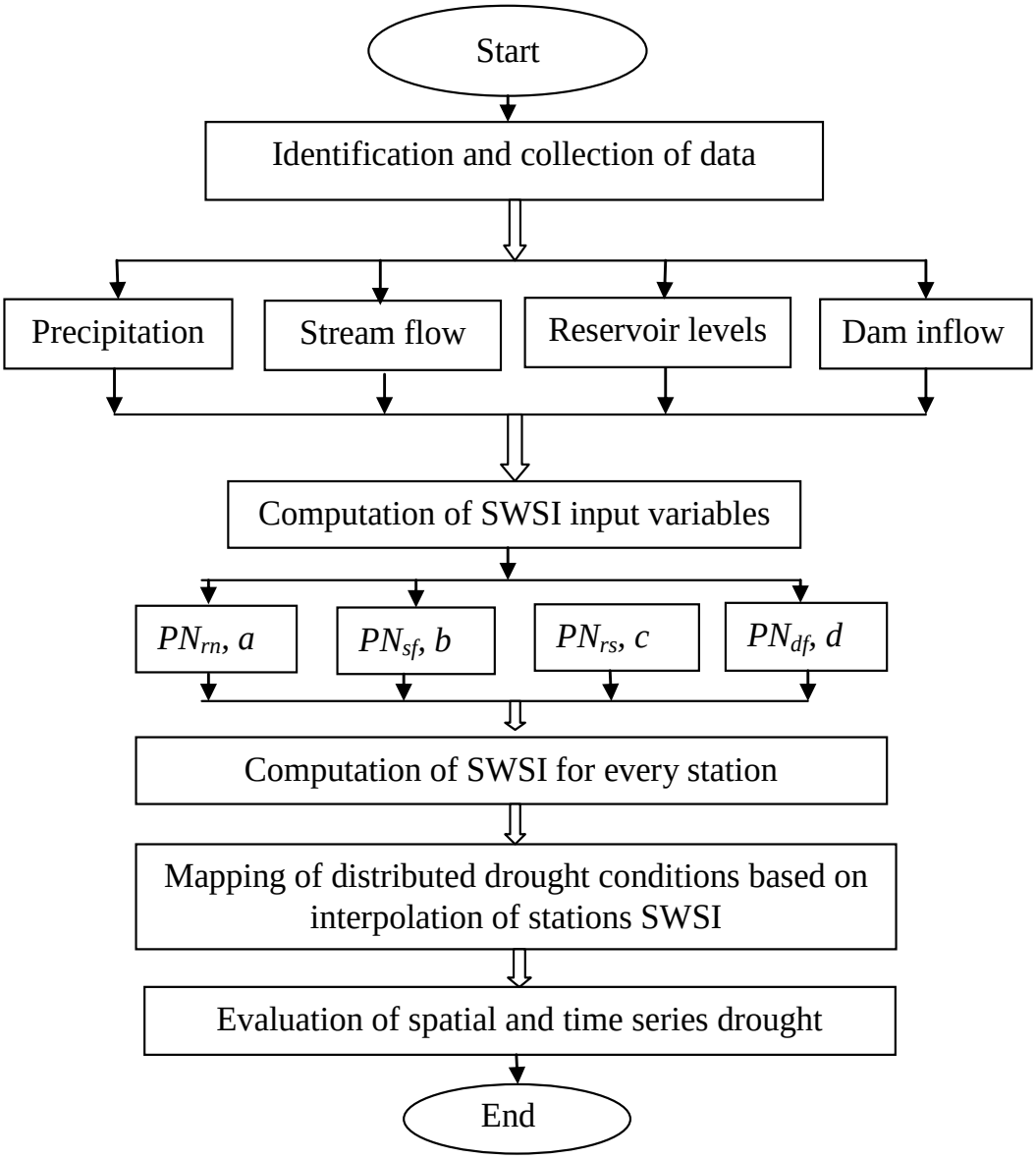


Figure 3.3: Process of evaluating drought using the modified SWSI

The computed SWSI values were then categorized into severity classes based on Table 3.3 adopted from Shafer and Dezman (1982).

Table 3. 3: Drought classification based on SWSI

State	Criterion	Drought description
0	4.00 or more	Abundant water availability
1	3.99 to 1.99	Wet
2	2.00 to-0.99	Near normal
3	-1.00 to -1.99	Incipient drought
4	-2.00 to -2.99	Moderate drought
5	-3.00 to -3.99	Severe drought
6	-4.00 and less	Extreme drought

3.2.7 Stream flow drought index

Stream flow drought index was selected in this research because it is one of the most important types of drought that affects a number of activities depending on surface water resources (Nikbakht *et al*., 2013). Monthly Stream flow drought index (SDI) for each gauged station was determined using the relation adapted from Modarres (2007) defined as:

$$SDI = \frac{\left(Q - \overline{Q}\right)}{\sigma} \tag{3.5}$$

Where;

Q = discharge in time interval (monthly average flow) (m^3/s)

$\overline{Q}$ = mean discharge of the series (m^3/s)

σ = the standard deviation of the discharge for the reference period defined as:

$$\sigma = \sqrt{\frac{1}{n-1}\sum\left(Q - \overline{Q}\right)} \tag{3.6}$$

Where;

n = total number of months in the series

The computed SDI was then defined using a criterion presented in Table 3.4.

Table 3. 4: Definition of states of drought based on SDI

State	Drought description	Criterion
0	Nodrought	$SDI \geq 0.0$
1	Mild drought	$-1.0 \leq SDI < 0.0$
2	Moderate drought	$-1.5 \leq SDI < -1.0$
3	Severe drought	$-2.0 \leq SDI < 1.5$
4	Extreme drought	$SDI < -2.0$

3.2.8 Standardized precipitation index

The Standardized Precipitation Index (SPI) was used to quantify rainfall deficit within the basin as a representation of drought condition as defined by Mckee *et al*. (1993). The first step involved fitting the rainfall data into a probability distribution function and then computation of the SPI values.The computed SPI values were used in drought assessment and classification. In the first step, the gamma distribution function was adapted since it fits well in time series rainfall data (Cassiamani *et al*., 2007). The gamma distribution is expressed in terms of its probability density function as:

$$f(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad for\ x,\alpha,\beta > 0 \tag{3.7}$$

Where;

α = the shape parameter

β = scale parameter

x = the rainfall amount (mm)

$\Gamma(\alpha)$ = the value taken by gamma function

$\bar{x}$ = mean rainfall (mm)

The $\Gamma(\alpha)$ is the value defined by the Gamma function which is determined by applying an integral function according to Cacciamani *et al*. (2007) expressed as:

$$\Gamma(\alpha) = \int_0^{\alpha} x^{\alpha-1} e^{-y} dx \tag{3.8}$$

Where;

$\Gamma(\alpha)$ = the value taken by gamma function

x = the rainfall amount (mm)

α = the shape parameter

The Gamma function in Equation (3.8) was evaluated both by the numerical method and use of tabulated values using the selected shape parameter α. A maximum probability was then used to estimate the optimal values of α and β using Equations (3.9) and (3.10):

$$\alpha = \frac{1}{4A}\left(1+\sqrt{1+\frac{4A}{3}}\right) \tag{3.9}$$

$$\beta = \frac{\bar{x}}{\alpha} \tag{3.10}$$

Where;

α = the shape parameter

β = scale parameters

$\bar{x}$ = mean rainfall (mm)

A = sample statistic

The sample statistic is defined as:

$$A = \ln(\bar{x}) - \frac{\ln x}{n} \tag{3.11}$$

Where;

$\bar{x}$ = the rainfall average (mm)

n = the number of observations

The calculated values were in turn used to compute the cumulative probability for non-zero rainfall using Equations (3.12) and (3.13) respectively:

$$f(x;\alpha,\beta)=\int_0^x f(x,\alpha,\beta)dx = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^x x^{\alpha-1}e^{-x/\beta}\,dx \tag{3.12}$$

Where;

α= the shape parameter

β =scale parameter

x =the rainfall amount (mm)

The Equation (3.12) above was reduced to:

$$f(x;\alpha,\beta)=\frac{1}{\Gamma(\alpha)}\int_0^x t^{\alpha-1}e^{-t}dt \qquad for\ \ t=\frac{x}{\beta} \tag{3.13}$$

Where;

$\Gamma(\alpha)$ = the value taken by gamma function

x = the rainfall amount (mm)

β = scale parameter

t = the time period

The Gamma function was applied for values of rainfall $x > 0$ for the rainfall time series of the upper Tana River basin. In case of non-zero values, cumulative probability of both zero and non-zero values were computed. This probability is represented by a function $H(x)$ defined as:

$$H(x)=q+(1-q)F(x;\alpha,\beta) \tag{3.14}$$

Where;

$H(x)$ = Cumulative probability

q = probability of zero rainfall

When m was taken as the number of zero entries in the time series rainfall data, then the q value was estimated by the ratio m/n. The cumulative probability was then transformed into a standard normal distribution function. This gave values of the mean and variance of the *SPI* as zero and one respectively. This step was carried out using approximate transformation functions adapted from Mishra and Desai (2006). These functions given in Equations (3.15) and (3.16) are expressed as:

$$SPI = -\left(k - \frac{c_0 + c_1 k + c_2 k^2}{1 + d_1 k + d_2 k^2 + d_3 k^3} \right) \quad for \ 0 < H(x) \leq 0.5 \tag{3.15}$$

$$SPI = +\left(k - \frac{c_0 + c_1 k + c_2 k^2}{1 + d_1 k + d_2 k^2 + d_3 k^3} \right) \quad for \ 0.5 < H(x) < 1 \tag{3.16}$$

Where;

c_0 =2.515517

c_1= 0.802853

c_2 = 0.010328

d_1 =1.432788

d_2 = 0.189269

d_3 = 0.001308

The parameters were used to compute the SPI and were adapted from Cassiamani *et al*. (2007). The value of *k* in Equations (3.15) and (3.16) was determined from the functions given as:

$$k = \sqrt{\ln\left(\frac{1}{H(x)^2} \right)} \qquad for \ 0 < H(x) \leq 0.5 \tag{3.17}$$

$$k = \sqrt{\ln\left(\frac{1}{1 - H(x)^2} \right)} \qquad for \ 0.5 < H(x) < 1 \tag{3.18}$$

In this study, the *SPI* values were calculated using a monthly time step and the threshold ranges adapted from Mckee *et al*. (1993) given in Table 3.5 were used to define drought conditions. The SPI computatuion porcess is summarized in Figure 3.4.

Table 3. 5: Drought conditions based on SPI

Threshold value (s)	Drought classification
2.00 or more	Extremely wet
1.50 to 1.99	Very wet
1.00 to-1.49	Moderate wet
0.99 to -0.99	Near normal
-1.00 to -1.49	Moderate drought
-1..50 to -1.99	Severe drought
-2.00 or less	Extreme drought

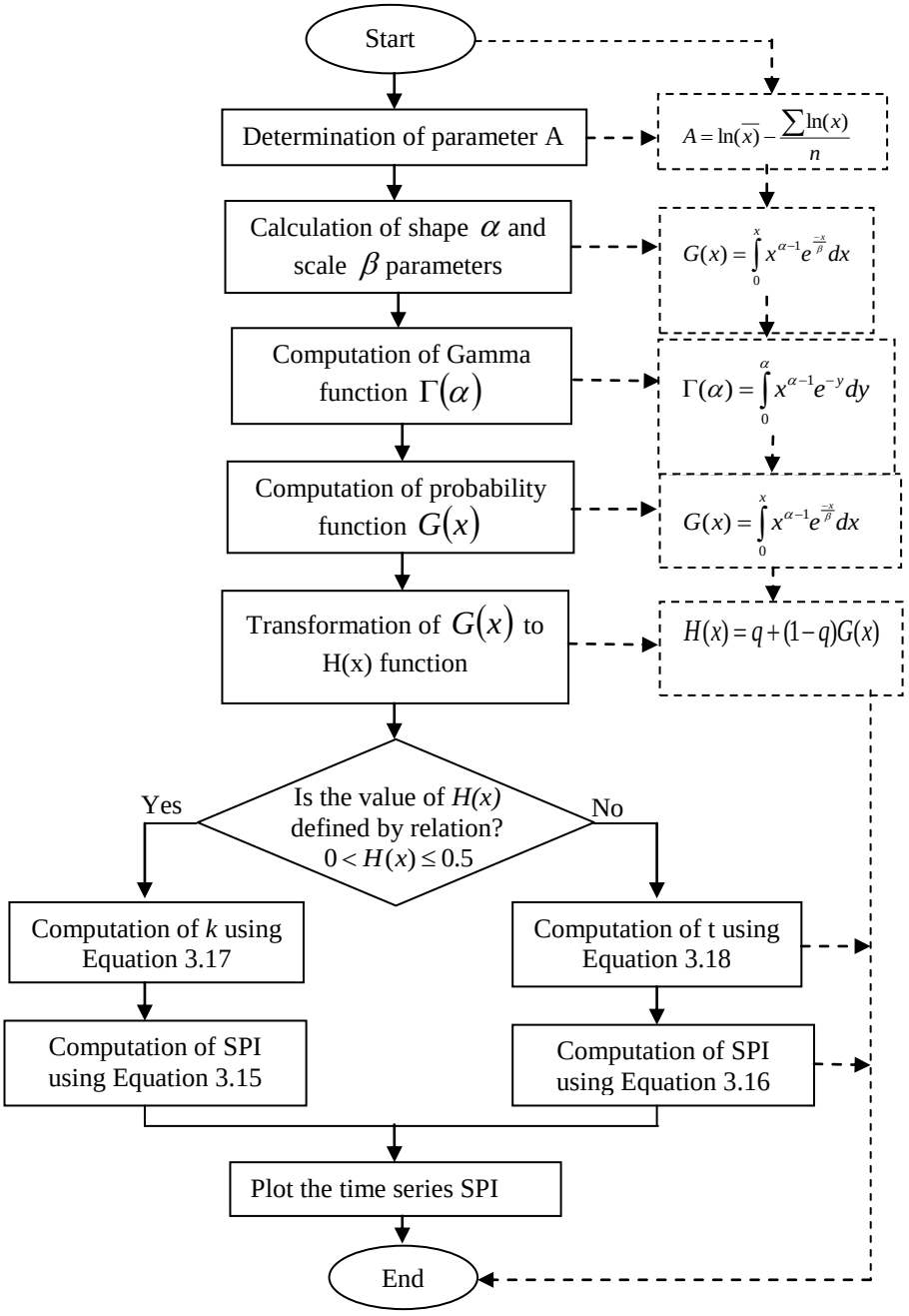


Figure 3. 4: Process for computation of the time series SPI

3.2.9 Effective drought index

The effective drought index (EDI) was computed using monthly time step data for the weather stations within the study area according to Smakhtin and Hughes (2007). The computation of the EDI was done through four steps. The first step involved the calculation of the effective precipitation parameter EP_p of the current month using the relation:

$$EP_p = \sum_{m=1}^{N} \left(\frac{\sum_{i=1}^{m} PE_m}{m} \right) \quad (3.19)$$

Where;

EP_p = effective precipitation parameter (mm)

m = total period before the current month

PE_m = the precipitation in m-1 months before the current month (mm)

N = duration of summation of the precipitation

The mean EP is computed annually to represent the climatological characteristics of water resources. For practical application of MEP a 5-months running mean is applied in this computation (Bhun and Whilhite, 1999). Then the deviation time series EP from the mean EP was computed using the relation:

$$DEP = EP - MEP \quad (3.20)$$

Where;

DEP=deviation of time series EP_p from mean effective precipitation parameter (mm)

MEP=mean effective precipitation parameter (mm)

From the EP_p, both the mean and the standard deviations of the monthly values were determined. The resulting time-series *EP* was used as inputs to calculate its deviation from the mean. Then the return to normal precipitation (RNP) values was determined using the relation adopted from Roudier and Mahe (2010):

$$RNP = \frac{DEP}{\sum\left(\frac{1}{N}\right)} \quad (3.21)$$

Where;

RNP=return to normal precipitation (mm)

N=previous period (months)

From the calculated *RNP*, the *EDI* was be derived from the relation:

$$EDI = \frac{RNP}{Std\left(RNP\right)} \quad (3.22)$$

Where;

Std (RNP) = Standard deviation of a particular months RNP values

Using the computed *EDI* values, the severity of the drought was categorized based on the thresholds and classification (Table 3.6) adopted from Morid *et al.* (2006 and 2007).

Table 3. 6: Drought severity based on EDI

Threshold value (s)	Drought classification
2.5 or more	Extreme wet
1.50 to 2.49	Very wet
0.7 to 1.49	Moderately wet
-0.99 to 0.99	Near normal
-1.0 to -1.49	Moderate drought
-1.5 to -1.99	Severe drought
Less than -2.00	Extreme drought

3.2.10 Soil Moisture Deficit Index

The first step was the preparation of all the AquaCrop input data. The data set consisting of monthly precipitation, maximum and minimum temperature, relative humidity, mean wind speed and sunshine hours for the period 1970-2010 were used for the selected meteorological stations. In addition, soil texture data and soil depth profile of 60-cm were used. The 60-cm soil depth was chosen in this study because it represents the most active soil zone that contributes to evapo-transpiration processes of plants. The 60-cm soil depth represents the potential of most agricultural crops to extract water from rooting depth. The quantity of water extracted depends upon stage of growth and crop type (FAO, 2013).

To effectively assess the agricultural drought, data was segregated into dry and wet seasons for specific agricultural applications. This is because different seasons exhibit different drought characteristics as stated by Nalbantis (2008), and Tsakiris and Nalbantis (2009). Each season had two distinct period with the dry season represented by the period from the months of July to September (J-S), and January to March (J-M). The wet season was represented by the period from October to December (O-D), and March to June (M-J).

3.2.11 Smulation of Soil Water (SW) content using AquaCrop model

The Input climatic data (average monthly rainafall, maximum temperature (T_{max}), Minimum temperature (T_{min}) and radiation were prepared in excel. First the T_{max}, T_{min} and the radiation data was imported into AquaCrop model where potential evapotranspiration (ET_o) was computed using Hargreaves Equation 3.28. This technique was chosen over the other methods of estimating ET_o because it requires only temperature and radiation as input data which were available in the study area. After computation of the ET_o, the monthly data sets; the ET_o, dominant soil type and the 60-cm soil depth profile were imported to FAO developed AquaCrop model Version 4.0. The monthly soil water content (SW_i) was then generated from the AquaCrop model. The steps used in computation of SW is summarized in a flow chart (Figure 6B; Appendix B)

The resulting SW_i was then used to compute *SMDI* by first computing the monthly Soil Defict (SD_i) and then $SMDI_i$. Equations (3.23), (3.24) and (3.25) adapted from Narasimhan and Srinivasan (2005) and Ramazani *et al*. (2012) were used to estimate the *SMDI* for the upper Tana River basin.

$$SD_i = \left(\frac{SW_i - MSW_i}{MSW_i - SW_{\min i}} \right) \times 100, \quad for\ SW_i \le MSW_i \tag{3.23}$$

$$SD_i = \left(\frac{SW_i - MSW_i}{SW_{\max i} - MSW_i} \right) \times 100, \quad for\ SW_i > MSW_i \tag{3.24}$$

Where;

SD_i = soil water deficit (percentage)

SW_i = mean monthly available soil water in a soil profile (mm)

MSW_i = long-term median available soil water in the soil profile (mm)

SW_{maxi} = long-term maximum available soil water in the soil profile (mm)

SW_{mini} = long-term minimum available soil water in the soil profile (mm)

$$SMDI_j = 0.5 \times SMDI_{j-1} + \frac{SD_j}{50} \tag{3.25}$$

Where;

$SMDI_i$ = soil moisture deficit index for j^{th} month

i = i^{th} month of the year

$SMDI_i$ = the soil moisture deficit for i^{th} month

SD_i = soil moisture deficit (%) for i^{th} month of a particular year

The integration of all the steps used in computing the values of SMDI are summarized in Figure 3.5.

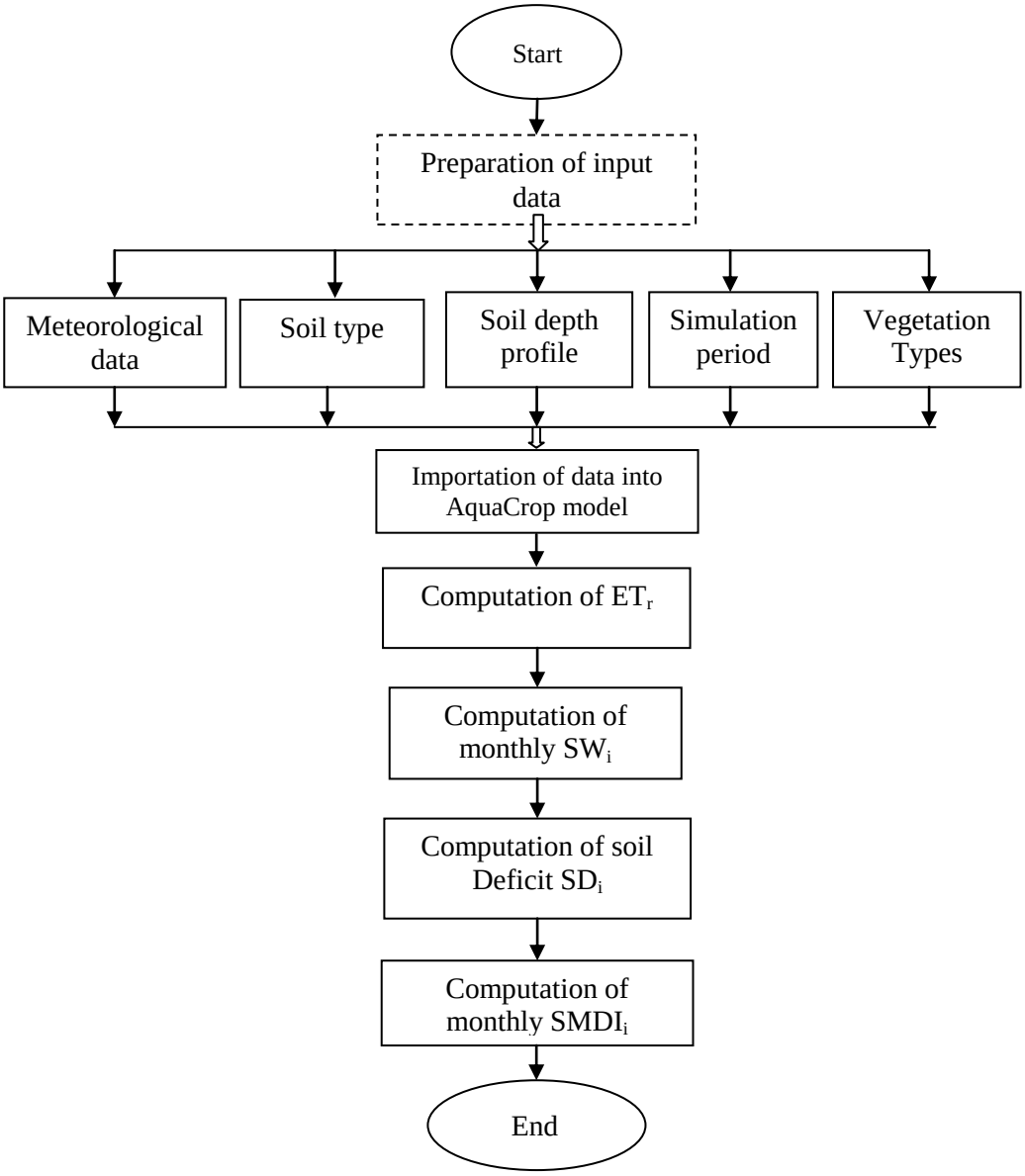


Figure 3.5: Flow chart showing the steps in computation of SMDI

3.2.12 Palmer Drought Severity Index

The Palmer Drought Severity Index (PDSI) was computed using precipitation, temperature and the local Available Water Content (AWC) of the soil as the input variables. The available water capacity (AWC) was estimated based on the dominant soil characteristics for the each elevation band of the upper Tana River basin. For the gauge stations within the four partitions of elevation bands, the AWC values adapted for PDSI computation were 172, 98, 74 and 82 mm which were based on values given in Table 3.7, for defined dominant soil types.

Table 3. 7: Dominant soils for the upper Tana River basin

Elevation	Dominant soil type	MC at saturation %	MC at field capacity %	MC at wilting point %	AWC (%)	TAW (mm)
HE	Andosols	60	40	24	16	172
ME	Nitosols	53	31	22	9	98
LME	Cambisols	48	28	14	14	74
LE	Ferrasols	53	26	17	9	82

HE, ME, LME, LE means highest elevation, middle elevation, lower middle elevation and lowest elevation respectively

Source: Hunink *et al*. (2009)

The PDSI was determination by getting the difference between actual precipitation and water deficiency or surplus in any given month *i*. This was achieved by applying the relation:

$$d_i = P_i - \hat{P}_i \tag{3.26}$$

Where;

d_i = difference between actual precipitation and p_i and the climatically appropriate for existing conditions (mm)

p_i = actual precipitation (mm)

$\hat{P}_i$ = an indicator of water deficiency or surplus in month *i*

The water deficiency or surplus was estimated from the relation:

$$\hat{P} = \alpha PE_i + \beta PR_i + \gamma PRO_i + \delta PL_i \tag{3.27}$$

Where;

$\hat{P}_i$ = an indicator of water deficiency or surplus in month *i* (mm)

PE_i = potential evapo-transpiration of month *i* (mm)

PR_i = potential recharge that gives the quantity of water required to bring the soil to its water holding capacity (mm)

PRO_i = the potential runoff (which is defined as the difference between the precipitation and potential recharge (mm)

PL_i = potential loss or the amount of soil moisture that could be lost from soil by evapo-transpiration during a zero precipitation period (mm)

The potential evapotranspiration was estimated using Hargreaves method adapted from Sivaprakasam *et al*. (2011) given as:

$$PE = 0.0023 \times R_a \times (T_{mean} + 17.78) \times (T_{\max} - T_{\min})^{0.5} \tag{3.28}$$

Where;

PE = potential evapotranspiration (mm/month)

R_a = solar/extra-terrestial radiation (MJ m^{-2} $month^{-1}$)

T_{mean} = mean monthly temperature (^oC)

T_{max} = maximum monthly air-temperauture (^oC)

T_{min} = minimum monthly air-temperature (^oC)

The α, β, γ and δ are climatic coefficients which provide mean value averaged within the base period. These coefficients were computed from the following relations:

$$\alpha = \frac{\overline{ET}}{\overline{PE}} \tag{3.29}$$

$$\beta = \frac{\overline{R}}{\overline{PR}} \tag{3.30}$$

$$\gamma = \frac{\overline{RO}}{\overline{PRO}} \tag{3.31}$$

$$\delta = \frac{\overline{L}}{\overline{PL}} \tag{3.32}$$

Where;

$\overline{ET}$ = mean actual evapo-transpiration (mm)

$\overline{PE}$ = mean potential evapo-transpiration (mm)

$\overline{R}$ = mean actual recharge (mm)

$\overline{PR}$ = mean potential recharge (mm)

$\overline{RO}$ = mean actual runoff (mm)

$\overline{PRO}$ = mean potential runoff (mm)

$\overline{L}$ = mean water loss due to evapo-transpiration when precipitation is zero (mm)

$\overline{PL}$ = mean potential water loss (mm)

The values of monthly PR_i, PRO_i and PL_i were derived from the generated results of soil water content for every month i using the technique given by Yan *et al*. (2013). These variables were calculated from the following relations:

$$PR_i = AWC - SW_{i-1} \tag{3.33}$$

$$PRO_i = SW_{i-1} = AWC - PR_i \tag{3.34}$$

$$PL_i = \min\left(PE, SW_{i-1}\right) \tag{3.35}$$

The d_i was then converted into indices of moisture anomaly z_i which was calculated using the equation:

$$z_i = k_1 \times d_i \tag{3.36}$$

Where;

k_c = climatic characteristic that was estimated using the relation:

$$k_1 = \frac{\left(\overline{PE} + \overline{R}\right)}{\left(\overline{P} + \overline{L}\right)} \tag{3.37}$$

The PDSI function that was used in this study is of the form:

$$PDSI_i = 0.897\, X_{i-1} + \frac{Z_i}{C_1} \tag{3.38}$$

Where;

$PDSI$ = The PDSI for the i^{th} month

X_{i-1} = previous months PDSI

Z_i = Palmer Moisture Anomaly Index ($PMAI$)

The value of PDSI for the initial month was taken as equal to $\frac{Z_i}{C_1}$.

Z_i (PMAI) is expressed as:

$$Z_i = \frac{C_2\, D}{\sum_1^{12} D k_2} \times d_i \tag{3.39}$$

Where;

k_2 = weighting factor

d = water deficiency (mm)

c_2 = conceptual parameter

D = absolute value of d

In this study, a C_2 value of 438.91 adapted from Yan *et al*. (2013), was used. The k_2 which is a function of average water demand and supply (Barua, 2010; Zoljoodi and Didevarasl, 2013) was estimated using the relation:

$$k_2 = C_3 \log_{10}\left(\frac{\left(\overline{PE} + \overline{R} + \overline{RO}\right)}{\left(\overline{P} + \overline{L}\right)\overline{D}}\right) + C_4 \tag{3.40}$$

Where;

$\overline{D}$ = mean of the absolute values of *d*

The conceptual parameters C_3 and C_4 were equated to 1.2459 and 3.3684 respectively adapted from Yan *et al*. (2013). The computed *PDSI* values were used to classify drought conditions based on the threshold levels given in Table 3.8 which was adapted from Palmer (1965) and Castano (2012).

Table 3.8: Classification of drought based on PDSI

Value of index	Drought classification
4.00 or more	Extremely wet
3.00 to 3.99	Very wet
2.00 to 2.99	Moderately wet
1.00 to 1.99	Slightly normal
-0.50 to -0.99	Incipient wet
0.49 to -0.49	Near normal
-0.50 to -0.99	Incipient drought
-1..00 to -1.99	Mild drought
-2.00 to -2.99	Moderate drought
-3.00 to -2.99	Severe drought
-4.00 or less	Extreme drought

3.2.13 Evaluation of Spatial distribution of drought severity

The sum of drought severity DI_d values below zero during each year for the study period was calculated. The probability *P* of drought occurrence was determined by dividing the number of months that had *DI* values less than zero by 12 months of the year. The drought severity was then computed at each station using the relation:

$$S = \sum_{N=1}^{N} DI_d \times P \tag{3.41}$$

Where;

S = annual drought severity for a defined year

DI_d =The *sum* of drought severity values below zero during a particular year

P = probability of drought occurrence for the defined year

The resulting data was then used to estimate spatial distribution of drought severity using the Krigging estimator in the ArcGIS 10.1. In this study, sixteen hydrometric stations within and

adjacent to the upper Tana River basin were used for hydrological evaluation. These stations have unique geographical location and their spatial extent was created through the application GIS. The GIS tool was used to compute and present the spatial distribution, variation and trends of droughts for different drought indices (DIs).

3.2.14 Mann-Kendall trend test for drought conditions

To test for the trend in drought severity, a non-parametric Mann-Kendall trend test was applied. This method has the capacity to test for increasing, decreasing or no trend (Kendall, 1962). The data for the upper Tana River basin was evaluated using ordered time series. The data sets were organized in form of $x_1, x_2, x_3, \ldots, x_j$ *n*-data points where x_i represent data point at time *j*. Then the Mann-Kendall statistical trend *S* was determined using the relation:

$$S = \sum_{k=1}^{n-1}\left[\sum_{j=k+1}^{n} sign\left(x_i - x_k\right)\right] \tag{3.42}$$

The right hand side of the Equation (3.42) was simplified using Equation (3.43) given as:

$$sign\left(x_j - x_k\right) = \begin{cases} 1 \; if \; (x_j - x_k) > 0 \\ 0 \; if \; (x_j - x_k) = 0 \\ -1 \quad if \; (x_j - x_k) < 0 \end{cases} \tag{3.43}$$

The probability linked to the Mann-Kendall statistic *S* and the selected n-data were determined to quantify the level of significance of the trend. The *VAR(S)* was calculated and then the normalized test statistic *Z* was computed using the following equations:

$$VAR(S) = n(n-1)(2n+5) - \sum_{t} \frac{t(t-1)(2t+5)}{18} \tag{3.44}$$

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} \; if \; S > 0 \\ 0 \; if \; S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} \; if \; S < 0 \end{cases} \tag{3.45}$$

Where;

VAR(S) = the variance of the data set

n = the number of data points

Equation (3.42) which was adapted from Mahajan and Dodamani (2015) was used to qualify the drought trend in the basin as: no trend, increasing trend and decreasing trend when $S = 0$, $S > 0$ and $S < 0$ respectively. In order to determine whether or not the drought trend in the upper Tana River basin was significant or insignificant, significance levels at 90% and 95% were used. At these significance levels, the null hypothesis of no trend was rejected when $|Z| > 1.645$ and $|Z| > 1.96$ respectively where the values of Z were adapted from Sneyers (1990).

3.3 Drought forecasting using DIs and ANNs

Evaluation of the performance of the DIs for short, medium and long-term drought forecasting was carried out in two steps. The first step involved forecasting the short-term drought conditions followed by the medium-term and lastly the long-term drought conditions. To forecast the future values of the indices, the inputs into the forecasting networks were presented by various combinations of their present and previous values. All the input and output values were standardized using the normalization Equation (3.31) to range between 0.1 and 1, taking the minimum value (X_{min}) of 0.1 and maximum value (X_{max}) of 0.9 since previous research by Morid et al. (2007) confirmed this range to give the best results.

3.3.1 Drought forecasting

In this study, drought forecasting was objectively done by partitioning the selected drought indices into those that represent hydrological, meteorological and agricultural droughts. Table 3.9 shows the indices adapted for each drought forecasting category.

Table 3. 9: Different categories of drought forecasting

Drought index	Input variables	Drought type
SWSI	*P,Q rs, df*	Hydrological
SDI	*Q*	Hydrological
SPI	*P*	Meteorological
EDI	*P*	Meteorological
PDSI	*P, T, AWC*	Agricultural
SMDI	*Sm*	Agricultural
NDI	*P,Q, rs, sm*	Hydrological, Meteorological, Agrcicultural

P, Q, rs, df, T, AWC, sm are precipitation, stream flow, reservoir levels, dam inflows, temperature, available water content, soil moisture

3.3.2 Temporal drought forecasting using DIs

To forecast the drought using DIs, the forecasting range was partitioned into short, medium and long-term time ranges. This means that the drought was forecasted for up to 3, 9, and 24 months

respectively. To implement the forecasting of the DIs using the ANN models, the input and output parameters were first defined and the appropriate ANN model architecture designed. The ANN models were then trained and validated using selected perforamance criteria. The processs that was followed in drought forecasting using the ANN is summarized in Figure 3.6.

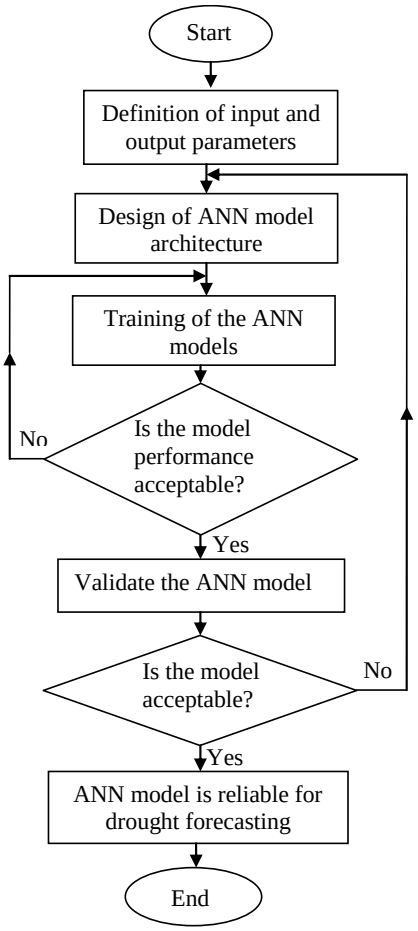


Figure 3. 6: Flow chart of the applied ANN-based drought forecasting model

3.3.3 Short-term drought forecasting

The hydro-meteorological data for different stations within the upper Tana River basin were used by applying the DIs modelling programme within the Matrices Laboratory (MATLAB). The first step involved defining the input neurons in the input layer. In this step, different input neurons with different time delays of *t, t-1, t-2, . . . , t-n,* for each hyro-meteorological station were applied. Secondly, the number of Hidden Neurons (HN) was set as equal to 2n+1 in the architecture of the ANN. By trial and error method, the hidden neurons were increased or decreased by a value of one and this was used to evaluate the value of forecasts using the following relation.

$$E_F = \left| 1.0 - \frac{DI_{Obs} - DI_{For}}{DI_{Obs}} \right| \times 100\% \quad (3.46)$$

Where;

E=Efficiency of forecasting

DI_{Obs}=Observed magnitude of drought index

DI_{For}=Forecasted magnitude of the drought index

An example of this process is given in Figure 3.7 that was used to forecast the SWSI.

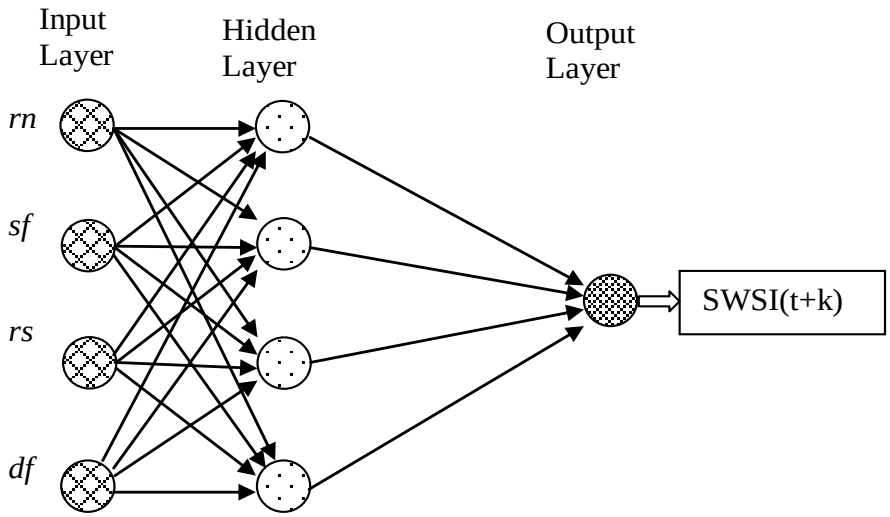


Figure 3. 7: ANN Architecture used for the forecasting of SWSI

For each hydro-meteorological station, the FFN and RNN structures of the ANN nodels were trained using the Levernberg-Marquardt (LM) algorithm (Figure 5B; Appendix B). Preliminary results at each station showed that a three-layer FFN with different input and HN perfomed best in terms of efficiency using MSE and R. In this case, the FFN were considered as the best for the detailed drought studies.

The ouput layer neurons in all the networks were equal to the forecast of the respective lead short-term forecast of DIs. For instance, a DI_{t+1} and DI_{t+3} into the future was calculated for 1 and 3 months lead times using the Neural Networks within the GUI of the MATLAB toolbox. In this research, the DIs values of 1 and 3 were considered to reflect relatively short-term conditions. On a monthly temporal domain, the DI_{t+1} and DI_{t+3} were calculated as the drought forecasts for lead times of 1 and 3 months respectively. Table 3.10 presents a summary of the forecasting inputs for various DIs as modified from Belayneh (2012).

In this study, the initial number of HN was determined using the relation adapted from Belayneh (2013) given as 2*n*+1, where n is the number of input neurons. The trial and error method was applied in which the number of HN was altered by either decreasing or increasing by 1 the computed HN, and testing the performance efficiency for each trial using Equation 3.46. The best ANN was selected by picking the model architecture that gave the best level of performance in forecasting efficiency. This process was repeated for all the other gauge stations, meteorological stations and different drought indices.

Table 3. 10: Key variables for short-term drought forecasting

DI	Lead time (months)	Drought forecast/Output
SWSI	1	$SWSI_{t+1}$
SDI	1	SDI_{t+1}
SPI	1	$SWSI_{t+1}$
EDI	1	EDI_{t+1}
PDSI	1	$PDSI_{t+1}$
SMDI	1	$SMDI_{t+1}$
SWSI	3	$SWSI_{t+3}$
SDI	3	SDI_{t+3}
SPI	3	SPI_{t+3}
EDI	3	EDI_{t+3}
PDSI	3	$PDSI_{t+3}$
SMDI	3	$SMDI_{t+3}$

3.3.4 Medium-term drought forecasting

The steps followed for medium-term drought forecasting were similar to those of short-term drought forecasting, except in the ouput layer of the ANN structure. In the ouput layer, neurons in all the networks are equal to the forecast of the respective lead medium-term forecast of DIs. DI_{t+6} and DI_{t+9} into the future were calculated for 6 and 9 months lead times using the Neural Networks within the GUI of the MATLAB toolbox. For this forecast, the DIs values of 6 and 9 were considered to represent medium-term conditions. On a monthly temporal domain, the DI_{t+6} and DI_{t+9} were computed as drought forecasts for lead times of 6 and 9 months respectively for the selected indices as shown in Table 3.11.

Table 3. 11: Variables for medium-term drought forecasting

DI	Lead time (months)	Drought forecast/Output
SWSI	6	$SWSI_{t+6}$
SDI	6	SDI_{t+6}
SPI	6	SPI_{t+6}
EDI	6	EDI_{t+6}
PDSI	6	$PDSI_{t+6}$
SMDI	6	$SMDI_{t+6}$
SWSI	9	$SWSI_{t+9}$
SDI	9	SDI_{t+9}
SPI	9	SPI_{t+9}
EDI	9	EDI_{t+9}
PDSI	9	$PDSI_{t+9}$
SMDI	9	$SMDI_{t+9}$

3.3.5 Long-term drought forecasting

For the long-term forecasting, the ouput layer neurons in all the networks are equal to the forecast of the respective long-term forecast of DIs. A DI_{t+12}, DI_{t+18} and DI_{t+24} into the future were computed for 12, 18 and 24 months lead times using the Neural Networks within the GUI of the MATLAB toolbox. The DI values of 12, 18 and 24 represent long-term drought conditions.

Table 3. 12: Variables for long-term drought forecasting

DI	Lead time (months)	Drought forecast/Output
SWSI	12	$SWSI_{t+12}$
SDI	12	SDI_{t+12}
SPI	12	$SWSI_{t+12}$
EDI	12	EDI_{t+12}
PDSI	12	$PDSI_{t+12}$
SMDI	12	$SMDI_{t+12}$
SWSI	18	$SWSI_{t+18}$
SDI	18	SDI_{t+18}
SPI	18	$SWSI_{t+18}$
EDI	18	EDI_{t+18}
PDSI	18	$PDSI_{t+18}$
SMDI	18	$SMDI_{t+18}$
SWSI	24	$SWSI_{t+24}$
SDI	24	SDI_{t+24}
SPI	24	SPI_{t+24}
EDI	24	EDI_{t+24}
PDSI	24	$PDSI_{t+24}$
SMDI	24	$SMDI_{t+24}$

On a monthly temporal domain, the DI_{t+12}, DI_{t+18} and DI_{t+24} were computed as drought forecasts for lead times of 6 and 9 months respectively for the selected indices as shown in Table 3.12. In these cases, results were presented in tabular forms.

3.4 Formulation of Nonlinear-Integrated Drought Index (NDI)

The Nonlinear-Integrated Drought Index (NDI) for the upper Tana River basin was formulated using the precipitation, stream flow, reservoir level and soil moisture content as parameters. The computation of the principal components, NDI time series values, and determination of threshold values were done. Figure 3.8 presents a summarized process that was used after the modification of the approach given by Barua (2012).

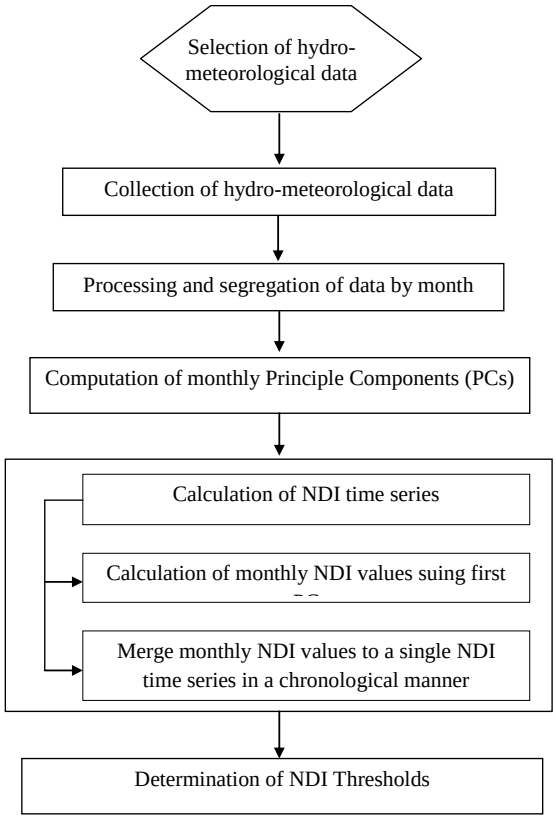


Figure 3. 8: Flow chart showing the process for computation of NDI

3.4.1 Computation of principal components (PC)

The 41 years data used to integrate four hydro-meteorological variables in formulating of the NDI were organized into (41x4) matrix of observations. The principal components (PCs) for each month of the year were generated and then the NDI values were computed for each month for all the years. The NDI values were computed and normalized using the relation:

$$NDI_{i,k} = \frac{Y_{i,l,k}}{\sigma_k} \quad (3.47)$$

Where;

$NDI_{i,k}$ = NDI value for month k in year i

$Y_{i,l,k}$ = PC1 during year i for month k

σ_k = standard deviation of $Y_{i,l,k}$ overall years for month k

All the steps in sections 3.41 and 3.42 involved writing of computation codes for the MATLAB as summarized in Table 3.13.

Table 3. 13: Principal Component Analysis in MATLAB

S.No.	Description of the operation	Formulated code of the MATLAB program
(i)	Definition of input matrix with hydro-meteorological variables	>> A=[2.54 1055.44 695.94 2.986…]
(ii)	Definition of matrix size	>> [n m]=size(A)
(iii)	Computation of mean vector	>> AMean=mean(A)
(iv)	Derivation standard deviation	>> AStd=std(A)
(v)	Standardization of the matrix data	>> B=(A-repmat(AMean,[n 1]))./repmat(AStd,[n 1])
(vi)	Computation of coefficients and variances for the principal components	>> [C S L]=princomp(B)
(vii)	Computation of first principal component (PC1)	>> PC1=S*L
(viii)	Calculation of NDI from the PC1	>> NDI=PC1/2.857
(ix)	Exportating of the results and plotting on MATLAB/Excel	

The matrices defined as Q = SCORE, E = LATENT and Y = Q x E as indicated in the equations above

After computing the *NDI* values for each year and each month, these values were combined into a single unit time series data.

3.4.2 Assessment of drought characteristics using the formulated NDI

After the computation of the NDI values, a time series plot was created based on truncation levels. The NDI values were then categorized into different drought conditions as given in Table 3.14 which was adapted from Barua (2010).

Table 3. 14: Classification of drought based on NDI

Value of index(s)	Drought classification
>1.63	Extreme wet
>1.20 to ≤1.63	Severe wet
>0.88 to ≤1.20	Moderate wet
>-0.84 to ≤0.88	Near normal
>-1.64 to ≤-0.84	Moderate drought
>-2.27 to ≤-1.64	Severe drought
≤-2.27	Extreme drought

From the time series plot, the duration, maximum and median magnitude of each drought condition were extracted. The model was also used to detect the on-set and end months of the drought period by selecting the times in months when the NDI value was below or above the truncation level. The NDI values below and above the truncation level were considered as the on-set and end time of the drought respectively. The duration, magnitude and severity of the drought in the upper Tana River basin was then computed using the relation:

$$S_{NDI} = \sum_{N=1}^{N} NDI_d \times P_{NDI} \tag{3.48}$$

Where;

S_{NDI} = annual drought severity based on NDI for a defined year

NDI_d = sum of drought severity values below zero during a particular year

P = probability of drought occurrence for the defined year

3.5 Drought forecasting using NDI

The formulation of the forecasting model involved selection of ANN structure, the drought input variables, calibration and validation of the selected variables and models. The process is summarized in Figure 3.9.

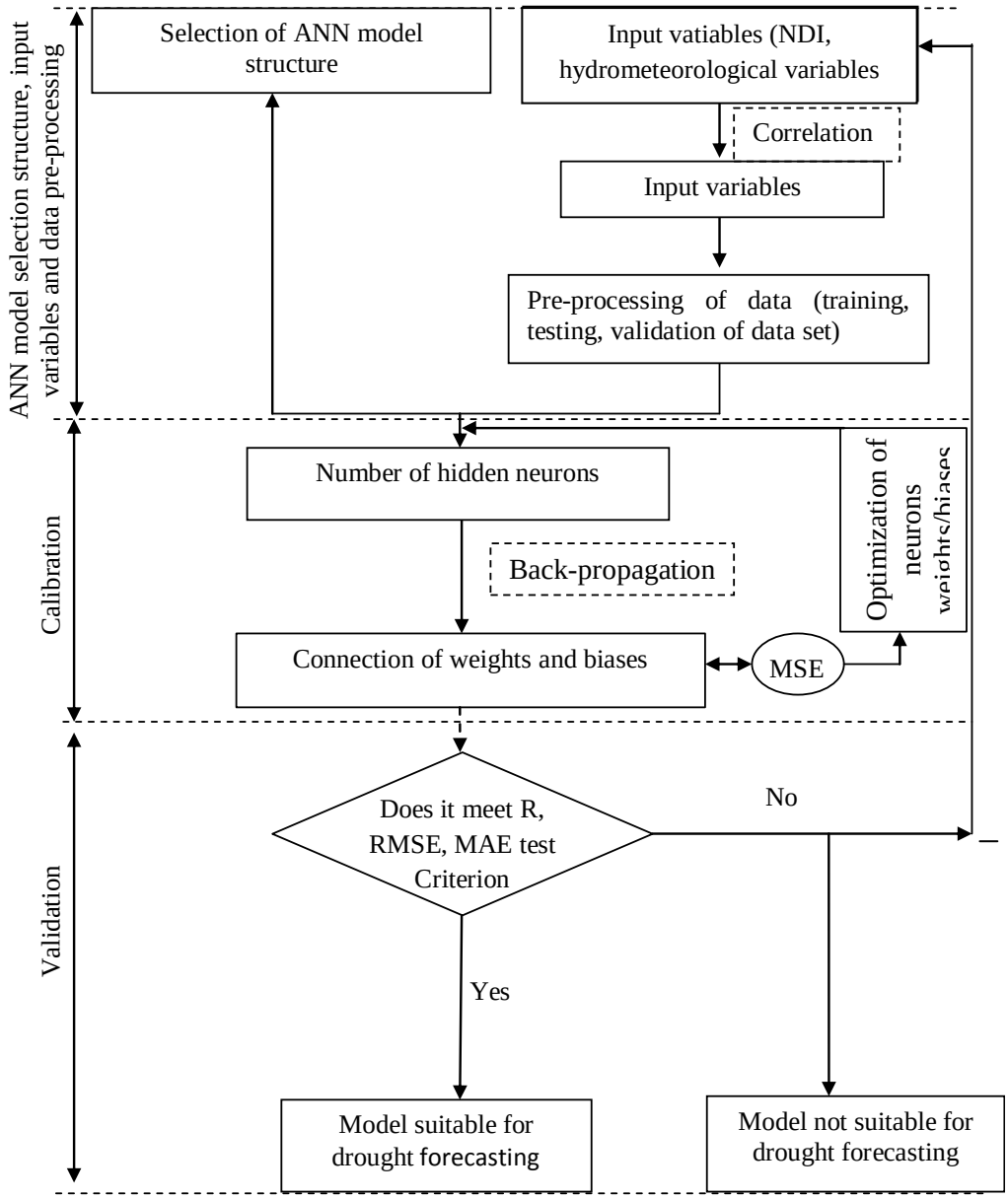


Figure 3. 9: Flow chart of ANN-based drought forecasting model for NDI

3.5.1 Identification of ANN model structure

The NDI values with several lead times for short, medium and long term droughts were forecasted using a three-layer Direct Neural Network (DNN). A three-layer ANN network model was adopted in this study because it can approximate any function if sufficient connection weights and biases are provided. The network had multiple neurons in the input and hidden layers with a single neuron in the output layer. The single output neuron represents a defined lead time forecast as shown in the Figure 3.10.

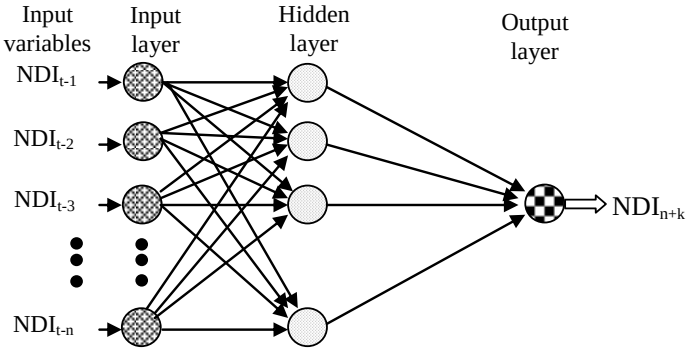


Figure 3. 10: A three-layer DNN

The multiple neurons in Figure 3.10 represent different combinations comprising NDI. The DNN was designed and calibrated for forecasting drought using present and several months of the past NDI values as inputs. Figure 3.10 shows *k* lead time forecasts from the present (*t*) month defined as NDI_{t+1}, NDI_{t+2},…, NDI_{t+3}.

The network was then used for forecasting NDI values for multiple lead times of *k*=1, 3, 6, 9 and 12 for short, medium and long-term drought forecasting. A direct Neural Network was adapted from Mishra and Desai (2006) and Barua (2010). The inputs for different forecast ranges given in Table 3.15 were used**.**

Table 3. 15: NDI input variables for drought forecasting

Model NDI	Lead time (months)	Output	Forecast range
NDI-LT$_1$	1	NDI$_{t+1}$	Short-term
NDI-LT$_3$	3	NDI$_{t+3}$	Short-term
NDI-LT$_6$	6	NDI$_{t+6}$	Medium-term
NDI-LT$_9$	9	NDI$_{t+9}$	Medium-term
NDI-LT$_{12}$	12	NDI$_{t+12}$	Long-term
NDI-LT$_{18}$	18	NDI$_{t+18}$	Long-term
NDI-LT$_{24}$	24	NDI$_{t+24}$	Long-term

3.5.2 Drought projection using NDI and Recursive Multi-Step Neural Networks

Recursive multi-step neural networks (RMSNN) with multiple neurons in the input and hidden layers adapted from Mishra and Dessai (2006) was used for drought projection from the year 2010 to 2099. The RMSNN consist only of a single neuron in the output layer which represent a one month lead time projection and done for k months (k=1068) . Based on the available data period of

1970 to 2010, the RMSNN was fisrt designed and calibrated considering only one month lead time based on present and several months of the past NDI values as inputs. The resulting network with the same number of input combination/variables was then used for projecting NDI values for multiple lead times recursively as shown in the Figure 3.11.The drought projection was conducted at month *t* for *k* time steps from (*t+1*) to (*t+k*). The projection (*t+1*) was first calculated based on n months of the past NDI values including the NDI at *t*. This projected value $NDI_{(t+1)}$ was then used with past NDI values of *t, t-1,…,t-n* months to project $NDI_{(t+2)}$. This process was repeated recursively to obtain the drought projection for 1068 months to represent the total period from the year 2011 to 2099. By using the projection it is assumed that the drought trend and associated hydro-meteorological variables within the basin is the same both for the data period (1970 to 2010) and the projection period (2011 to 2099)

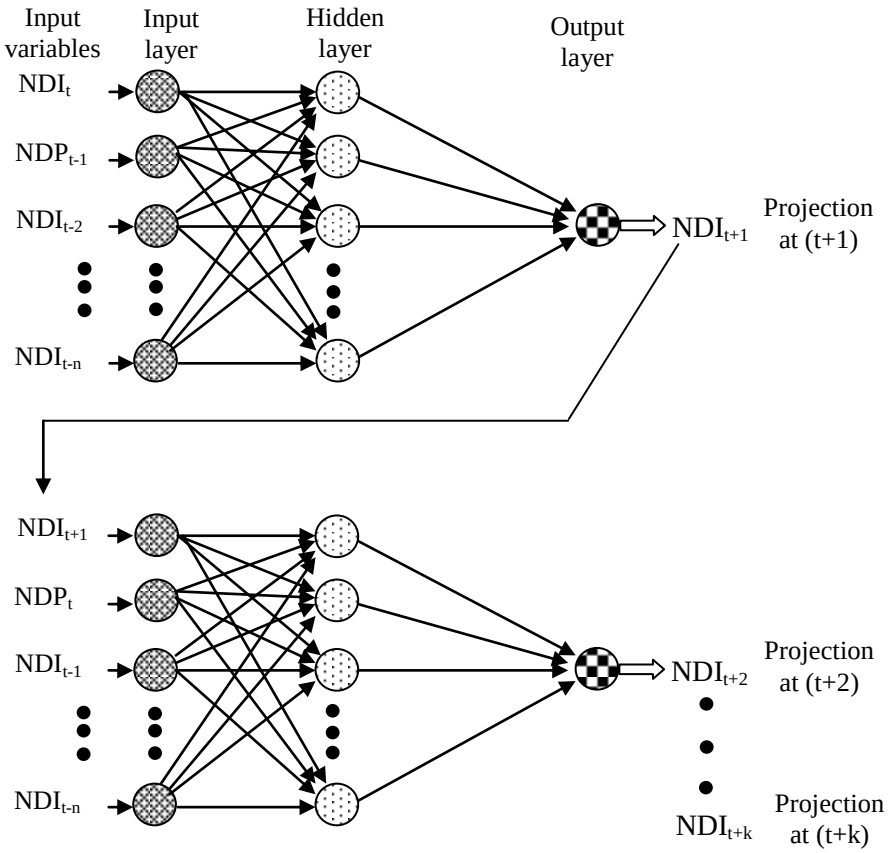


Figure 3. 11: Three Layer RMSNN used for drought projection

3.6 Sensitivity analysis of drought indices

Sensitivity of the indices to input parameters were carried out prior to calibration and drought assessment. This sensitivity analysis was used to identify parameters that are fundamental in influencing the index output and in detecting and quantifying interaction effects among the

parameters. The sensitivity analysis was conducted for drought indices for input parameters which included precipitation, stream flow, storage reservoir volume and dam inflow and soil moisture. The procedure involved computing the mean values of each parameter and then used as the initial value. Each of the mean parameter value was altered at a rate of $\pm 10\%$ and the corresponding change in both the parameter and *DI* values determined as the difference between the new values and the previous ones. Then, the sensitivity of the *DI* to parameter alterations was done by computating the absolute sensitivity. This was based on the concept of the ratio of of the relative change of a state variable to the relative change in parameter according to the procedure given by Jorgensen and Bendoricchio (2001) and Raude (2013). The Absolute Sensitivity relation adapted for drought in this study is defined by the relation:

$$S_A = \left| \frac{\delta DI / DI}{\delta P / P} \right| \tag{3.49}$$

Where;

S_A = Absolute sensitivity

δDI = change in the *DI* value

DI = drought index value before the change

P = value of parameter being evaluated before the change

δP = change in parameter value

3. 7 Time series drought characterization

The drought time series was characaterized by selecting the annual drought index with maximum severity. The resulting values were then arranged in ascending order and ranked. The return period of each event was then computed using the Weibull plotting position function given in Equation (3.49).

$$T = \frac{n+1}{r} \tag{3.50}$$

Where;

T = return period (years)

N = the number of data sets

r = the rank of an extreme drought event arranged in ascending order

The probability of any drought magnitude/cumulative severity was calculated as a reciprocal of its return period. The computed return period and the probability values of drought were then plotted against the DI values as fitted using Equation (3.50).

The drought was also charcaterized by developing Severity-Duration-Frequency (SDF) Curves by adopting and applying the Gumbel extreme value type 1 (EV1) approach. The Gumbel's method was originally developed for flood estimation. However, it has previously been adapted in drought studies for instance by Dalezios *et al*. (2000). The form of Gumbel technique adapted for the present study for estimating the extreme drought event is expressed as:

$$DI_T = \overline{DI}\left(1 + KC_v\right) \tag{3.51}$$

Where;

DI_T = probable cumulative droughtmagnitude/ severity with a return period of T years

C_v = coefficient of variation

$\overline{DI}$ = mean cumulative drought magnitude/severity (m^3/s)

K= frequency factor

In this study, drought severities for different durations were calculated and then the return period (years) and probalility of each drought severity computed. The computed values were then fitted in the Gumbel extreme value type 1 (EV1) distribution function to obtain severities corresponding to each of the drought durations. The frequency factor and the coefficient of variation were determined from the relation:

$$K = \frac{\left(y_T - y_n\right)}{\sigma_n} \tag{3.52}$$

$$C_v = \frac{\sigma}{\overline{Q}} \tag{3.53}$$

$$y_T = -\ln\ln\left(\frac{T}{T-1}\right) \tag{3.54}$$

Where;

y_n, = expected mean value

σ_n = standard deviation of reduced drought extremes estimated from Gumbel's Table

Equation (3.53) adopted from Asad *et al*. (2013) used an equation developed to estimate K using the relation:

$$K = \left[\frac{\left(\sqrt{6}\right) \times \left(0.5772 + \ln \ln T / T - 1\right)}{\pi} \right] \tag{3.55}$$

For stream flow drought index for example the following relation was applied modelling drought:

$$Y = \frac{Q_T}{\overline{Q}} = 1 + \frac{c_v \left(y_T - y_n\right)}{\sigma_n} \tag{3.56}$$

Where;

Q_T=Probable hydrological drought discharge with a return period of T years (m^3/s)

c_v=coefficient of variation

$\overline{Q}$=mean hydrological drought discharge (m^3/s)

σ_n=standard deviation of drought extremes estimated from Gumel's table

$y_T = -\ln \ln \left[T / \left(T - 1\right)\right]$

y_n=expected mean of extremes found from Gumbel's table

Although the above function (equation 3.55) improved the method of estimating the frequency factor *K*, which can now be computed based on return period *T* and not number of years of record, the method for calculating coefficient of variation (*Cv)* still remains as suggested by Gumbel.

The corresponding cumulative severity (CS) for different drought durations was computed for selected return periods of 2, 5, 10, 20, 30, 50, 100, 500 years and the using the resulting data, plots of the return period of different drought durations on the abscissa and cumulative severities (CS) on the ordinate were generated to yield the Severity-Duration-Frequency (SDF) curves.

3.8 Model calibration

The measured data was partitioned into two data sets. The first set of data consisting of 70% of the total data set was used for calibration while the remaing 30% was used for validation. The Calibration data set was further divided into training and testing data taking 70% and 30% respectively. The calibration involved determination of suitable number of neurons in the hidden layer, connection weights and biases in the neural network. This was achieved by minimizing the mean square error of the training data set using the approach given by Tran *et al*. (2009).

The number of neurons and hidden layers were determined to establish the ANN architecture. To establish the different function combinations of the variables adopted are log (N), 2n+1 to determine initial number of neurons in the hidden layer (Wanas *et al*., 1988; Mishra and Desai, 2006). The n is the number of input layers. The optimal number of neurons was then obtained through trial and error procedure by incresing or decreasing the initial number of neurons, while at the same time evaluating the value(s) of the performane measures. To effectively calibrate the models, training and testing data sets were used in parallel.

In this study, the Levernberg-Marquadt (LM), Back-propagation (BP) and CG training algorihms described in section 3.2.5 were adopted. The preliminary calibration showed that the feed-forward neural network with LM training algorithm was the best model for SWSI, SDI, SPI, EDI, SMDI PDSI and NDI. The feed-forward neural network (FFNN) trained using LM algorithm was therefore used to present the final results for the selected drought indices that were considered in the upper Tana River basin.

The Sigmoid and linear equations were used as the activation functions within the hidden and output neurons respectively. The two functions were selected because of the following advantages; the combination of sigmoid and linear functions has the ability to extrapolate values beyond the range of training data (Barua, 2010), and that the standardized range within the extreme limits of activation function prevents the weights from adjusting to extremely small values during calibration. The function used is of the form:

$$DI_{forecast} = f_o \times \left[\sum_{j=1}^{m} w_{jk} \times f_h \left(\sum_{i=1}^{n} w_{ij} x_i + b_j \right) + b_k \right] \tag{3.57}$$

Where;

x_i = input at the i^{th} neuron i output layer

w_{ij} = weight connecting the i^{th} neuron in input layer and j^{th} neuron in hidden layer

b_j = bias for the j^{th} hidden neuron

f_h = activation function of the hidden neuron

w_{jk} = weight connecting j^{th} hidden neuron in hidden layer and k^{th} neuron in output layer

b_k = bias for the k^{th} output neuron

f_o = activation function of the output neuron

To determine the model performance, the MSE was minimized using training data through monitoring the output for different iterations. For each iteration number, initial weights and biases

were taken as adjusted weights and biases of the previous iteration. The pattern of changes in MSE for both the training and testing data sets were plotted using the MATLAB GUI.

The calibration stopped at a point when the MSE for training data decreased while the MSE for testing data was observed to increase according to an approach given by Bishop (1995) and Barua (2010). At such a point, the correct set of weights and biases constitute optimal values. The number of HN was changed sequentially by adding or subtracting a new neuron to create another ANN model. Again, LM algorithm was then applied in determination of connection weights and biases for the current ANN. The same early stopping technique was again employed to select best number of HN of the ANN.

3.9 Model validation

A validation data set was used as input into the performance measures; correlation coefficient, mean absolute error, Nash-Sutcliffe efficiency and modified index of agreement to validate the models. A performance criterion for each performance measure was adapted as presented in the following sections.

3.9.1 The correlation coefficient

The correlation coefficient (R) was used to determine the statistical relationship between the observed and the predicated drought conditions within the upper Tana River basin. The fundamental function was customised to the respective DI values using the relation:

$$R = \frac{\sum_{i=1}^{n}\left(DI_{Obs} - \overline{DI}_{Obs}\right)\left(DI_{For} - \overline{DI}_{For}\right)}{\sum_{i=1}^{n}\left(DI_{Obs} - \overline{DI}_{Obs}\right)\left(DI_{For} - \overline{DI}_{For}\right)^{2}} \tag{3.58}$$

Where;

R = correlation coefficient

DI_{Obs} = observed value of the drought index

$\overline{DI_{Obs}}$ = mean of the observed values of the drought index

DI_{For} = forecasted value of the drought index

$\overline{DI_{For}}$ = mean of the forecasted values of the drought index

n = number of data points considered

The *R* is a measure of the strength of the linear relationship between the observed and forecasted *DI* values. It varies from 0 to 1. The values of 0 and 1 indicate a poor and perfect forecasting capability of the model respectively.

3.9.2 Mean absolute error

The mean absolute error (*MAE*) was determined from the relation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|DI_{Obs} - DI_{For}\right| \tag{3.59}$$

Where;

MAE = the mean absolute error

DI_{Obs} = observed value of the drought index

DI_{for} = forecasted value of the drought index

n = number of data points

The *MAE* was used to measure the average magnitude of errors of the set of forecasted values. The values of *MAE* increase from zero to large positive values. The higher the value of *MAE*, the higher the discrepancy between forecasted and observed values (Kim and Valdes, 2003). The lower values of *MAE* were considered to give the satisfactory results.

3.9.3 Mean square error

The Mean square error (*MSE*) is a measure of the difference between the observed and forecasted drought values from different indices. It measures the average of the squares of the errors between two values being compared. In this study, Equation (3.58) was adopted for calculating *MSE* associated with drought forecasting:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(DI_{iFor} - DI_{i\,Obs}\right)^2 \tag{3.60}$$

Where;

MSE = mean square error

DI_{Obs} = observed value of the drought index

DI_{for} = forecasted value of the drought index

n = number of data points

The *MSE* ranges from 0 to 1. The smaller the MSE value the better the forecasting capability of the model.

3.9.4 Nash–Sutcliffe efficiency

The Nash–Sutcliffe Efficiency (*NSE*) statistical approach has been used effectively to evaluate measured and predicted hydrologic data including drought (Nash and Sutcliffe, 1970; Biamah *et al*., 2002). The *NSE* was used to indicate how well the plot of observed versus simulated data fits the 1: 1 line: Its value ranges from $-\infty$ to 1.0. Mathematically *NSE* is given as:

$$NSE = 1.0 - \left[\frac{\sum_{i=1}^{n} \left(DI_i^{obs} - DI_i^{for} \right)^2}{\sum_{i=1}^{n} \left(DI_i^{obs} - \overline{DI}_{Obs} \right)^2} \right] \tag{3.61}$$

Where;

NSE = Nash–Sutcliffe Efficiency

DI_i^{obs} = observed value of the drought index

DI_i^{for} = forecasted value of the drought index

$\overline{DI}_{\text{Obs}}$= mean value of the drought index

n = total number of observations

The resulting values of NSE were compared with those given as acceptable levels of the efficiency as per Nash-Sutcliffe (1970) criterion.

3.9.5 Modified index of agreement

The modified index of agreement (d_1) was applied in performance testing to supplement the other methods used because it is more sensitive to the differences on forecasted and observed hydro-meteorological values than correlation coefficient (R) (Krause *et al*., 2005). This index gives the ratio of mean square error and the potential error (Wilmot, 1984) and is mathematically expressed as:

$$d_1 = 1.0 - \frac{\sum_{i=1}^{n} \left| DI_{Obs} - DI_{For} \right|}{\sum_{i=1}^{n} \left(\left| DI_{For} - \overline{DI}_{Obs} \right| + \left| DI_{Obs} - \overline{DI}_{Obs} \right| \right)} \tag{3.62}$$

Where;

d_1 = modified index of agreement

DI_{obs} = observed value of the drought index

DI_{for} = forecasted value of the drought index

$\overline{DI}_{Obs}$ = mean of the measured values

n = number of observations.

CHAPTER FOUR
RESULTS AND DISCUSSIONS

4.1 Temporal and spatial drought conditions

The results of the spatial and temporal drought conditions assessed using Surface Water Supply Index (SWSI), Stream Flow Drought Index (SDI), Standardized Precipitation Index (SPI), Effective Drought Index (EDI), Palmer Drought Severity Index (PDSI) and Soil Moisture Deficit Index (SMDI) are discussed in the following sections.

4.1.1 Time series SWSI

The results of drought assessement based on Surface Water Supply Index (SWSI) are presented in Table 4.1 for the Gura gauging station (ID. 4AD01) as an illustration for all the stations used in the upper Tana River basin. From Table 4.1, the highest and lowest values of weighted parameters for SWSI are the reservoir level and rainfall as denoted by *c* and *a* respectively.

Table 4. 1: Weighted Parameters for SWSI at Gura gauging station for 1970

Month	Weighted parameters for SWSI				Probability of non-exceedance			
(1970)	*a*	*b*	*c*	*d*	PN_{rn}	PN_{sf}	PN_{rs}	PN_{df}
Jan	0.096	0.126	0.605	0.173	0.371	0.116	0.605	0.173
Feb	0.035	0.039	0.703	0.222	0.773	0.515	0.703	0.222
Mar	0.061	0.034	0.673	0.232	0.592	0.576	0.673	0.232
Apr	0.079	0.167	0.548	0.205	0.424	0.051	0.548	0.205
May	0.252	0.097	0.475	0.176	0.049	0.119	0.475	0.176
Jun	0.277	0.029	0.522	0.172	0.047	0.517	0.522	0.172
Jul	0.011	0.036	0.739	0.214	0.945	0.594	0.739	0.214
Aug	0.087	0.057	0.686	0.171	0.475	0.367	0.686	0.171
Sep	0.007	0.061	0.771	0.161	0.961	0.381	0.771	0.161
Oct	0.163	0.118	0.598	0.119	0.213	0.124	0.598	0.119
Nov	0.221	0.044	0.597	0.138	0.138	0.412	0.597	0.138
Dec	0.097	0.029	0.692	0.180	0.442	0.651	0.692	0.180

Key*: a,b,c* and *d* ; PN_{rn}, PN_{sf}, PN_{rs} and PN_{df} are weight parameters; probability of non-exceedance for rainfall, stream flow, reservoir levels and dam inflows respectively

The sum of weighted parameters *a, b, c,* and *d* for each row data entry was found to be unity. This sum is in conformity with SWSI results obtained for other river basins such as those given by

Shafer and Dezman (1982), Kwon and Kim (2010). The results in Table 4.1 also indicate that the probability of non-exceedance was highest for the streamflow (PN_{sf}), followed by rainfall (PN_{rn}), reservoir level (PN_{rs}) and dam inflow (PN_{df}) respectively.

4.1.2 Sensitivity of SWSI to weighting parameters

The results of the absolute sensitivity analysis show that different weighting parameters of SWSI influence the index output at varying magnitudes. The results show that the sensitivity of SWSI to the four main weighting parameters increased in the order of c, *b*, d and *a* where these parameters correspond to reservoir level, stream flow, dam inflow, and precipitation respectively. The SWSI was found to be most sensitive when weighting variables related to precipitation (*a*) was reduced. On the other hand, the SWSI was least sensitive to both decrease and increase in reservoir levels (*c*). Generally, the SWSI was found to be more sensitive to a decrease than an increase in the weighted parameter values as given in Figure 4.1. The SWSI high sensitivity of weighting parameter *a* is attributed to the fact that precipitation is the most significant hydrological component that contributes to the runoff and the stream flow. Quantifying the SWSI sensitivity to parameter changes helped to understand its response to errors in parameter estimation prior to model calibration.

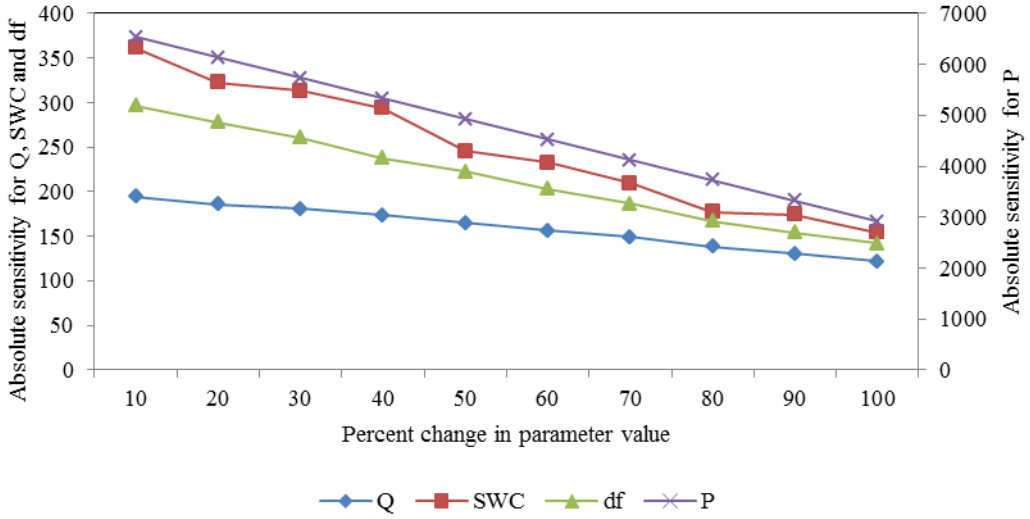


Figure 4. 1: Sensitivity of SWSI to decrease in weighted parameters

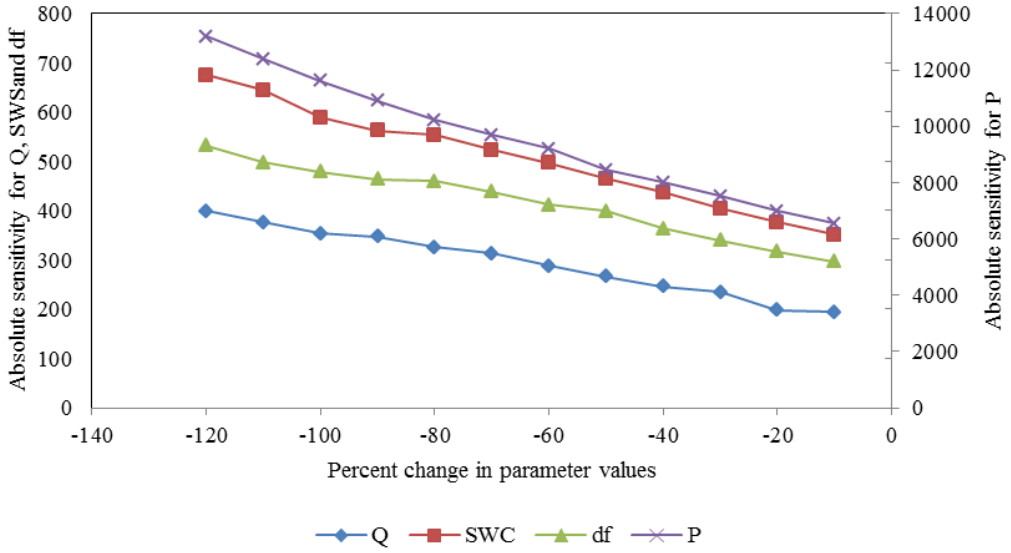


Figure 4. 2: Sensitivity of SWSI to increase in weighted parameters

4.1.3 Development and modification of SWSI equation

The time series SWSI was computed based on the equations formulated for respective gauging stations. For instance, the results for the Kamburu (ID 4ED01) gauging station were obtained at regression coefficient "*R"* of 0.863 using the modified SWSI Equation (4.1) given as:

$$SWSI_m = \frac{\left\lfloor 2.04 \times \left(a \times PN_{rn} + b \times PN_{sf} + c \times PN_{rs} + d \times PN_{df}\right) - 1 \right\rfloor}{0.1849} \qquad (4.1)$$

Where;

$SWSI_m$ = the modified surface water supply index,

This procedure was repeated for all the selected gauging stations. Equation (4.2) that presents a general mathematical function of SWSI was formulated for all stations in the upper Tana River basin.

$$SWSI_m = \frac{\left\lfloor k_1 \times \left(a \times PN_{rn} + b \times PN_{sf} + c \times PN_{rs} + d \times PN_{df}\right) - k_2 \right\rfloor}{k_3} \qquad (4.2)$$

The parameters k_1, k_2 and k_3 in the equation are conceptual parameters that are dependent upon the basin characteristics. As an illustration, these conceptual values for the Kamburu gauge station (ID 4ED01) are 2.04, 1.0 and 0.1849 respectively. In addition, the results give time series of the SWSI values between -4.22 to +4.22. This equation resulted into values that ranged between -4.22 and

+4.22 and which are in agreement with the values given by Shafer and Dezman (1982), and supported by Kwon and Kim (2010).

The time series SWSI results of the four gauging stations namely; Yatta furrow (ID 4CC03), Nyamindi (ID 4DA10), Tana sagana (ID 4BC02) and Amboni (ID 4AB05) are summarized in Figures 4.3 to 4.6.

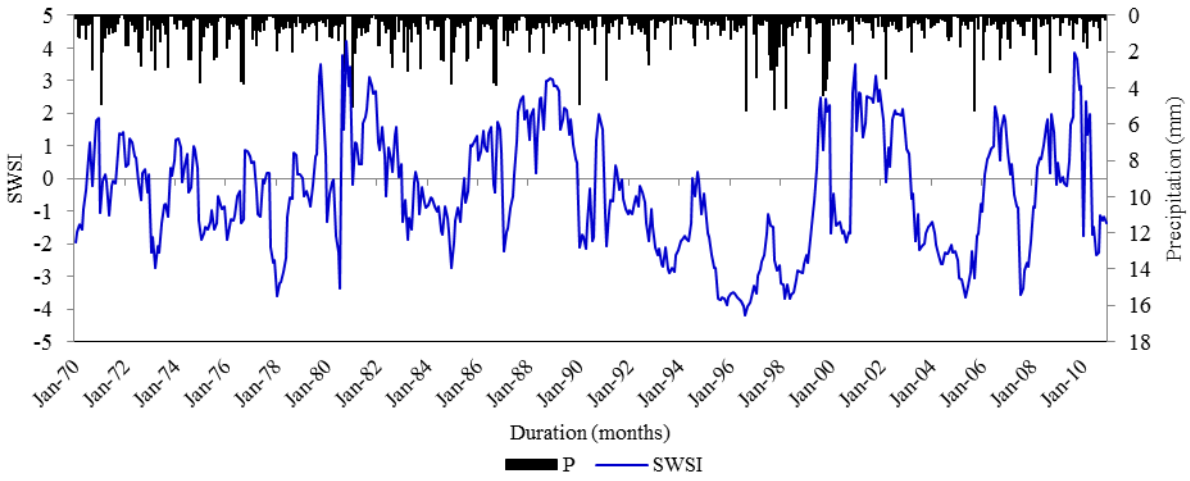


Figure 4. 3: Time series of SWSI for Yatta furrow gauging station

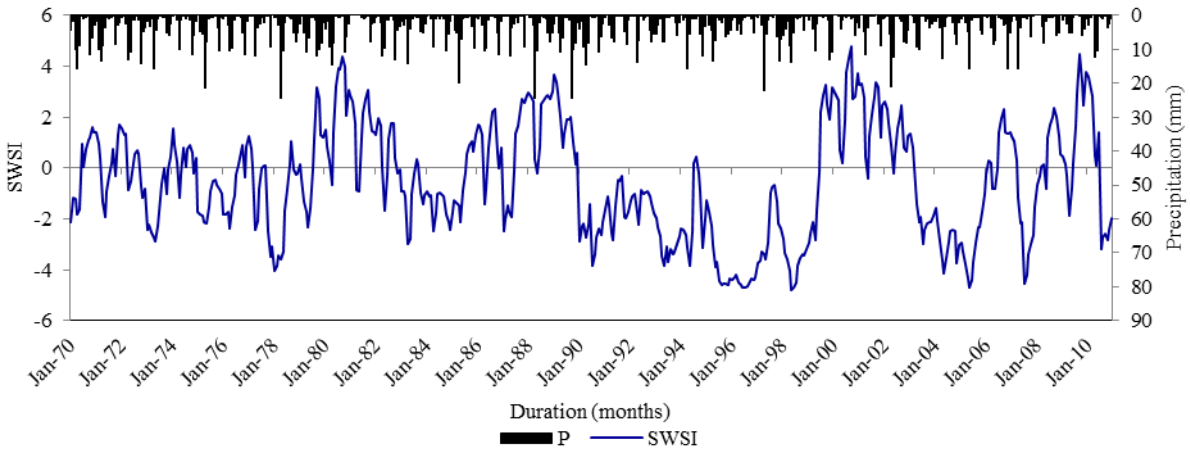


Figure 4. 4: Time series of SWSI for Nyamindi gauging station

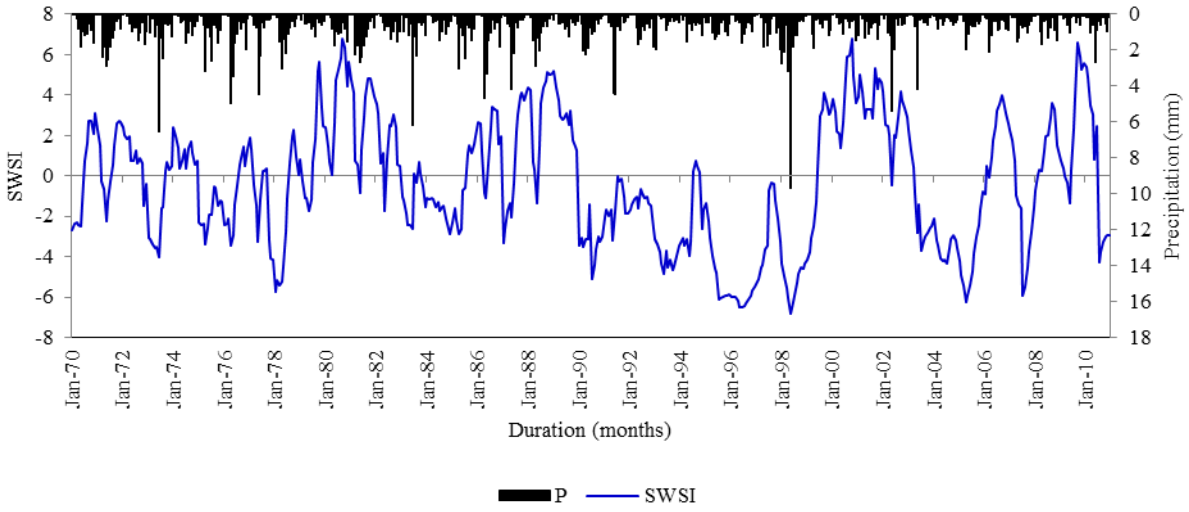


Figure 4. 5: Time series of SWSI for Tana sagana gauging station

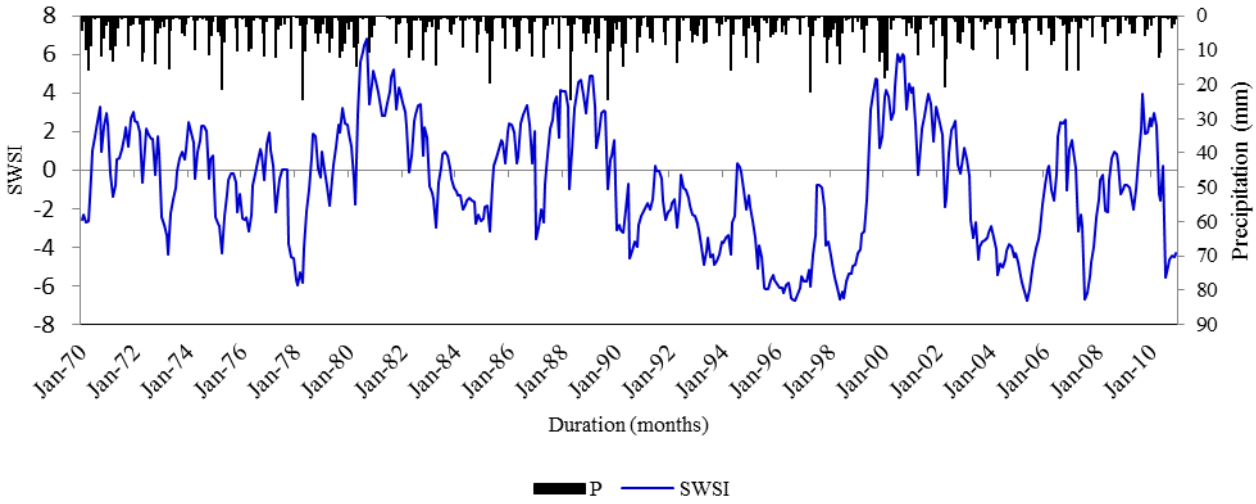


Figure 4. 6: Time series of SWSI for at Amboni gauging station

From these results, it is observed that the drought magnitude varied from station to station with major droughts of different durations occurring between 1990 and 1997 for all the gauging stations. The results show that the SWSI values were consistently below -3.0 in all the stations for the period 1990 to 1997. This means that most areas in the basin experienced severe drought in the stated years. The results also indicate that severe or moderate droughts at varying magnitudes as detected by the SWSI were experienced in 1973, 1978, 1984, 1994, 2000 and 2010 with the SWSI values equal or less than -2.0.

The results presented in Figure 4.7 show that the values of $SWSI_m$ for the months of April and September are -2.40 and -3.5 respectively.

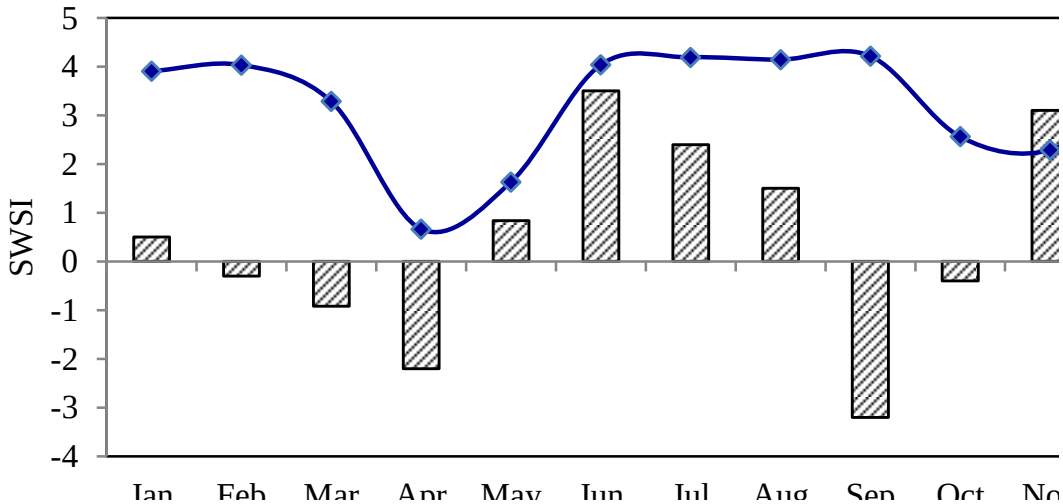


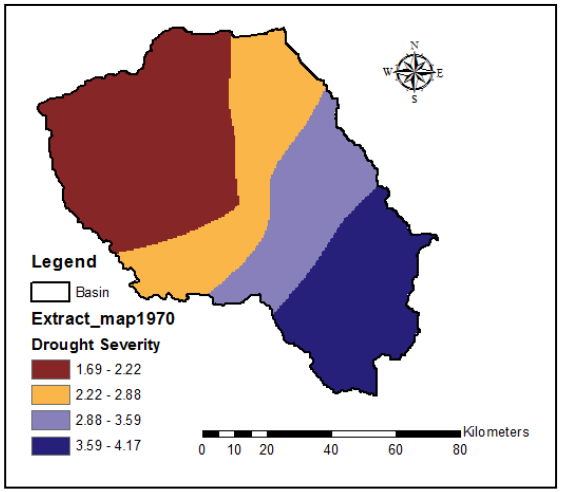
Figure 4. 7: Trend of the mean monthly SWSI and precipitation

These values represent moderate (-2.00 to 2.99) and severe (-3.00 to 3.99) drought conditions in the river basin for the specified months. However, the $SWSI_m$ for the months of June (Jun), November (Nov) and December (Dec) are 3.3, 3.5 and 4.2 respectively. These values lie between 1.99 to 3.99 and 4.00 or greater representing wet and abundant water supply condition for the months of Jun, Nov and Dec for the basin. These findings are are consistent with the results presented by Kwon and Kim (2010).

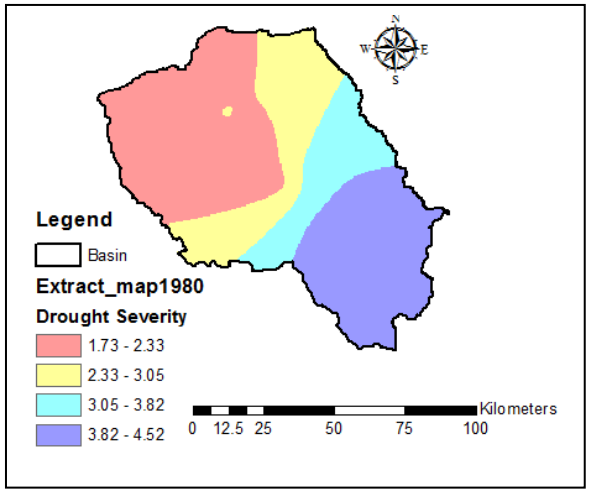
4.1.4 Spatially distributed drought severity based on SWSI

In terms of severity area coverage of the basin, the results of the computed drought severities for 1970, 1980, 1990, 2000 and 2010 show that the temporal and spatial drought severity has been increasing over the years with values ranging between 1.69 to 2.22 and 3.59 to 4.17 (Figure 4.8a) for 1970 and 3.74 to 6.29 and 4.37 to 4.96 (Figure 4.8e) for 2010. On a spatial extent, the areas in south-eastern parts of the basin have generally the highest drought severity over all the years. These areas are within the Arid and Semi Arid Lands (ASALs) at low elevations and receive annual average precipitation of between 700 and 1000 mm. These south-eastern areas which exhibit the highest decadal drought severities are considered to be the most drought-risk prone areas. The areas north-west of the basin are within the humid zones at high elevations. These areas receive annual precipitation amount ranging from 1000 to 2000 mm and exhibit low drought severities as given in

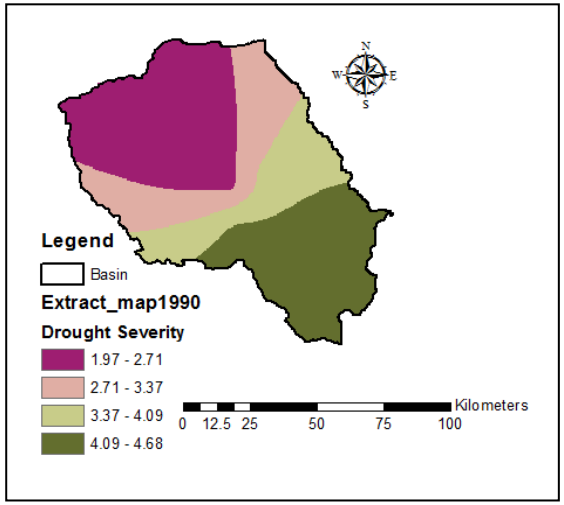
Figure 4.8h. The SWSI results were compared with similar findings given by Kwon and Kim (2010) where 32 and 112 sub-basins for a South Korean river basin were used; results from both the 32 and 112 sub-basin scenarios indicated variation of SWSI for different months of the year and spatial domain.



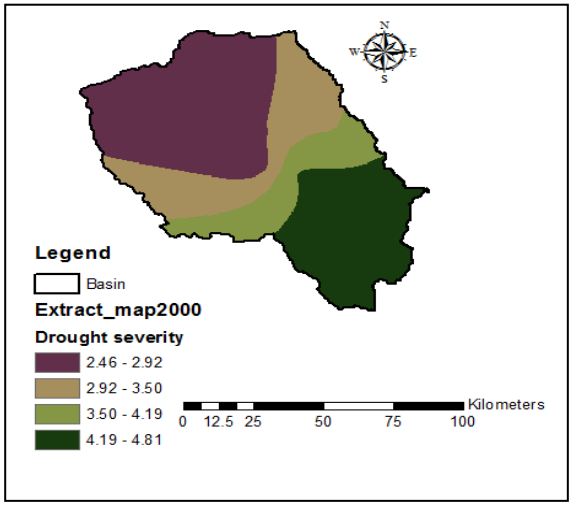
(a)



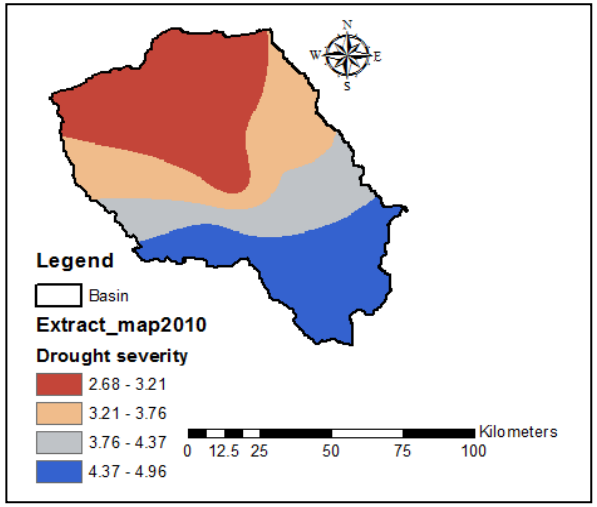
(b)



(c)



(d)



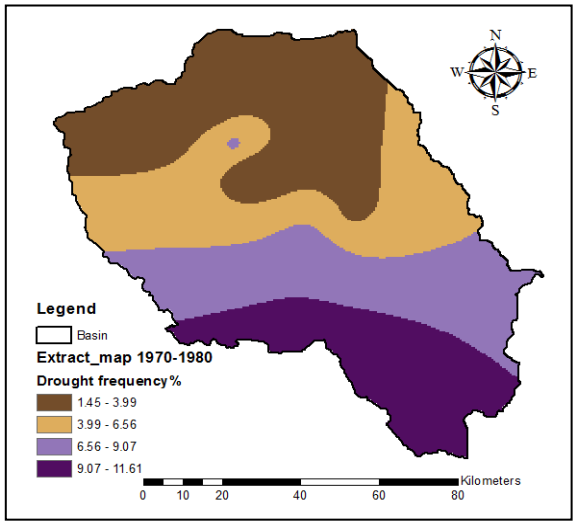
(e)

Figure 4. 8(a-e): Spatially distributed drought severity in the upper Tana River basin

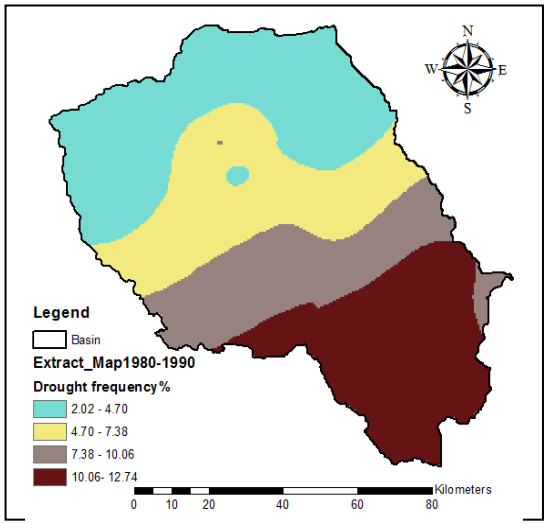
The results in Figure 4.9 show that the drought frequency and its variation for every decade and for the entire study period have been increasing. Due to varaiability of hydro-meteorological factors, and the non-homogeinity in land use/cover systems within the basin, the results show that the drought characteristics vary spatially.

Decadal drought frequency distributions for the upper Tana River basin are presnted in Figure 4.9 (a -d). From the results, it is observed that drought in the basin has become more frequent since the 1970s reaching severe levels in 1990s when the frequency over the entire basin was greater than 4.8%. However, there is notable frequency decline from 2000 to 2010 compared with that of 1990s (Figure 4.9(c) and (d). From Figure 4.9 (e), the the results show that minimum and maximum average drought frequency for the 41 years (1970 to 2010) range from 3.74 to 6.29 and from 10.61 to 13.76% respectively. The results of decadal frequency indicate that the highest and lowest drought frequencies are in the south-eastern and north-western parts of the basin respectively.

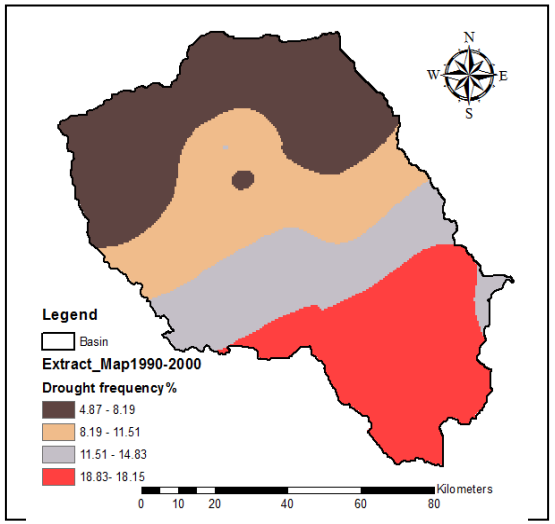
Similar research conducted by Shen *et al.* (2015) showed that the source areas of Waihe River are more drought-prone than the areas downstream of the river. The results of Shen *et al.* (2015) showed that the highest frequency of drought was more than 60% for source areas while the downstream areas exhibited drought frequency of less than 30% during the period 2000 to 2012. However, the results of frequency analysis for the upper Tana River basin are in contrast with these findings. Instead, the downstream parts of the basin located in south-eastern areas indicate higher drought frequencies (Figure 4.9) meaning that the south-eastern areas are more drought-prone than the source areas which are within the north-western and north-eastern parts. The difference in results of upper Tana Tana river basin and the waihe river basin is due to the fact that latter basin underwent more rapid land use/cover change in the source areas than the downstream parts for the period the research was conducted. The occurrence of drought events has significantly changed in most parts of the upper Tana River basin for the period 1970-2010. For the purpose of illustrating the drought frequency, the Severe Drought (SD) is shown using Figure 4.9. The highest drought frequency of between 10.61 and 13.76 occur in the south-eastern areas while the northern parts exhibit the lowest frequency ranging from 3.74 to 6.29 as shown in Figure 4.9. This drought frequency results again indicate that the areas in south-eastern parts of the basin are most prone to severe drought conditions.



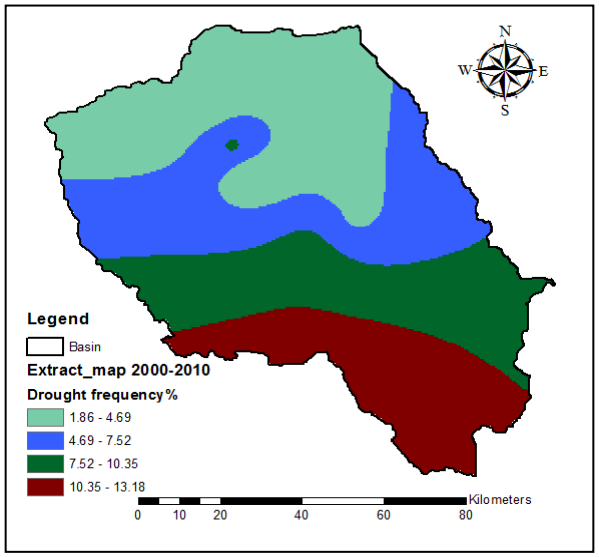
(a)



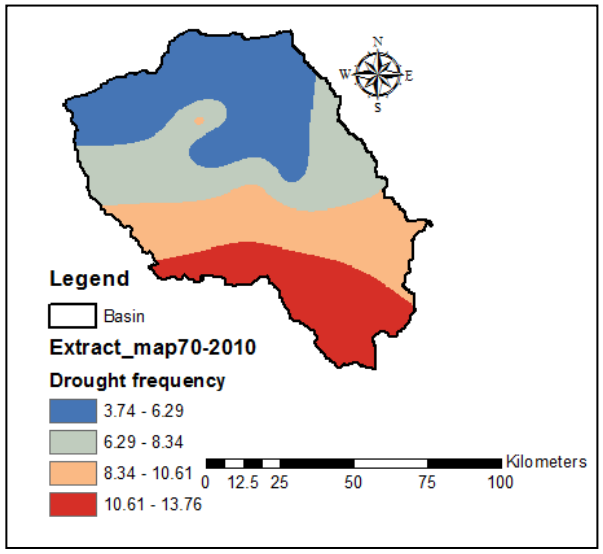
(b)



(c)



(d)



(e)

Figure 4. 9(a-e): Spatially distributed drought frequency of severe drought for SWSI

Results from the Mann-Kendall trend test criteria described in section 3.2.13 show that there was an increase in drought trend in the south-eastern parts of the basin at 90% and 95% significant levels while no significant trend was detected in the north-western areas. An increase in trend in drought which is significant at 95% significant level and insignificant at 90% significant level for the middle elevations as given in Figure 4.10 was obtained.

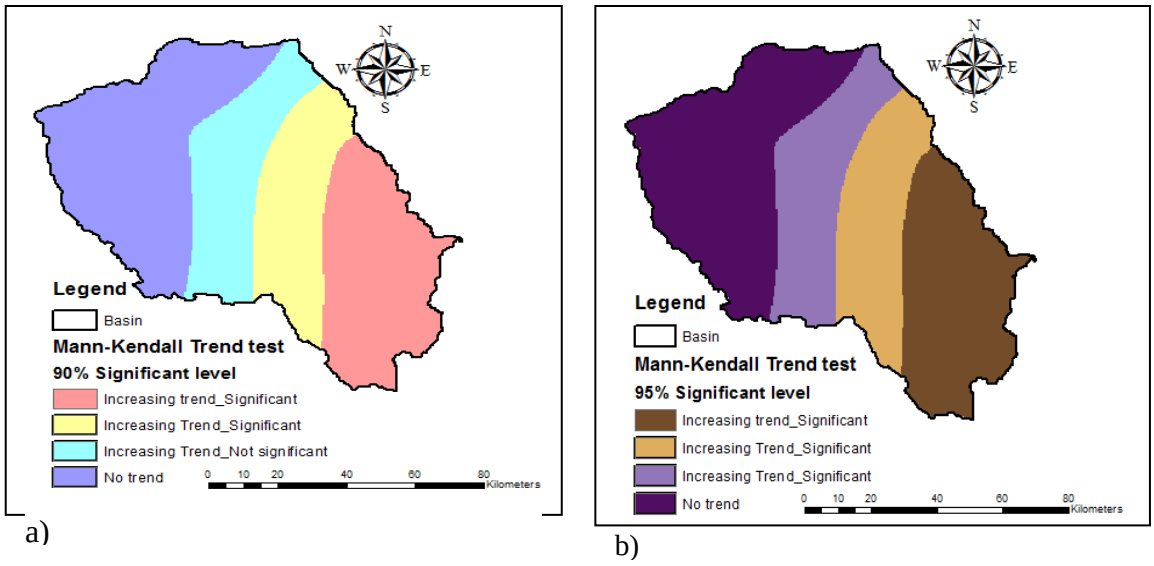


Figure 4. 10: Spatially distributed Mann-Kendall trend of severe drought for SWSI

Variation of the drought frequency and coverage are influenced by the basin meteorological factors and climate change (Zhou *et al*., 2012). Some of the critical meteorological factors include humidity and temperature (Yao *et al*., 2013). According to Wu *et al*. (2015), temperature is a major climatic factor that influences exchange of substances such as water and energy within a basin. Based on the hydrometric stations data plots, there has been a significant increase in yearly and decadal mean temperature in the upper Tana River basin as shown in Figures 4.11 and 4.12. This increase leads to increase in maximum potential evapotranspiration, reduction of precipitation and reduction of terrestrial surface humidity.

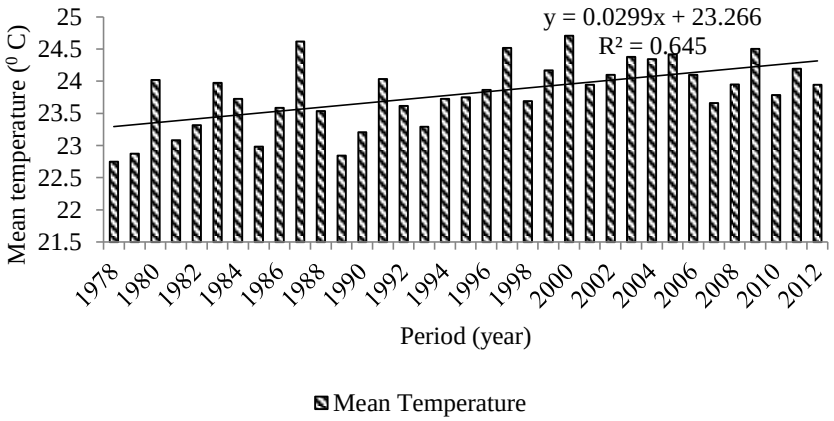


Figure 4. 11: Mean yearly temperature at Nyeri hydrometric station for 1978-2012

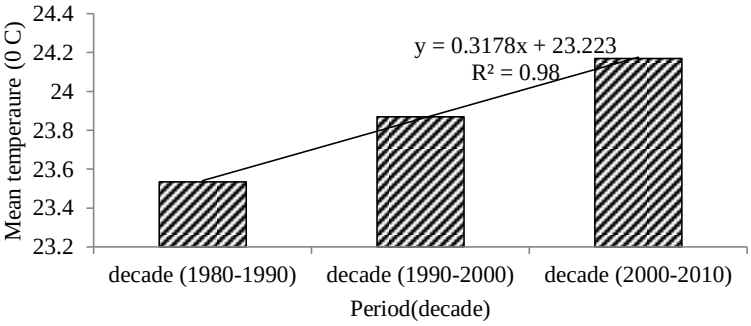


Figure 4. 12: Mean decadal temperature at Nyeri hydrometric station for 1970-2010

4.1.5 Time series SDI

The results of of monthly time series of SDI are presented in Figures 4.13 to 4.16 for gauging stations Amboni (ID 4AB05), Tana sagana (ID 4BC02), Nyamindi (ID 4DA10) and Kamburu (ID 4ED01) respectively.

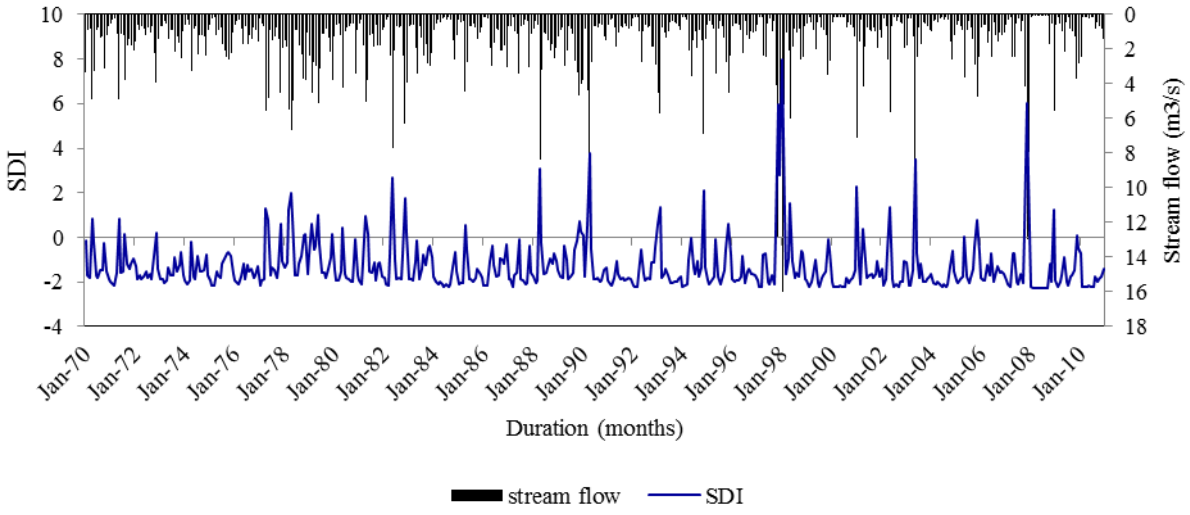


Figure 4. 13: Time series of SDI and stream flow for Amboni gauging station

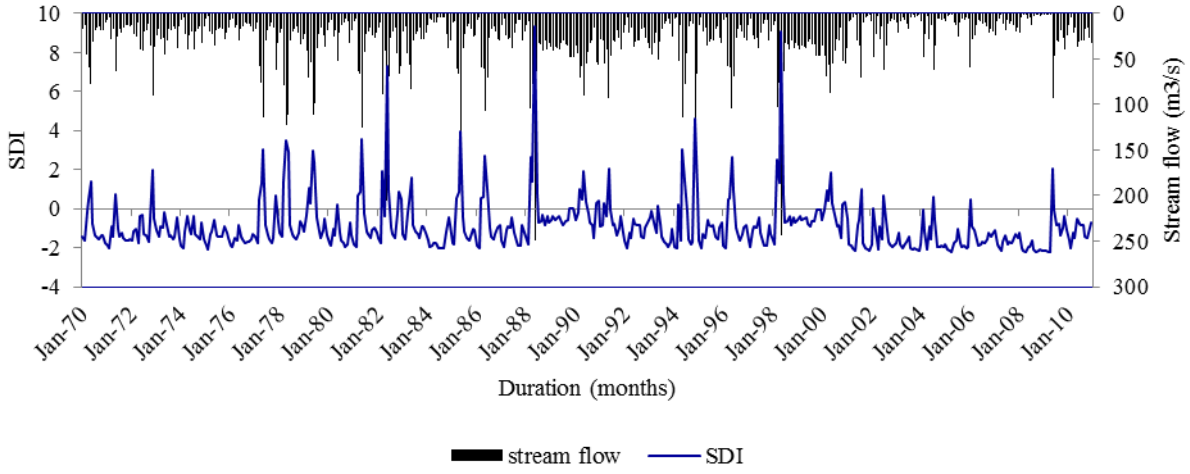


Figure 4. 14: Time series of SDI and stream flow for Tana sagana gauging station

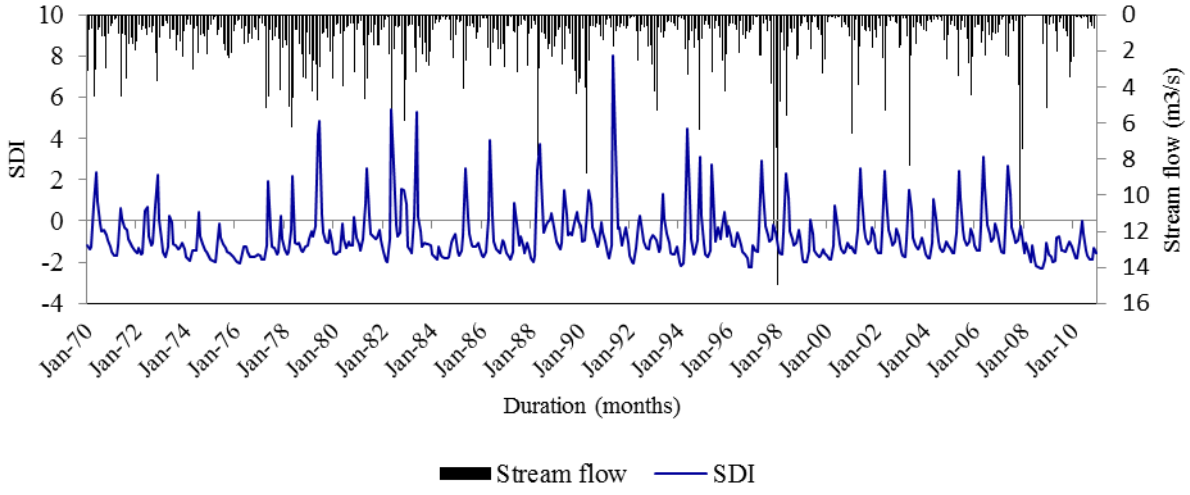


Figure 4. 15: **Time series of SDI and stream flow for Nyamindi gauging station**

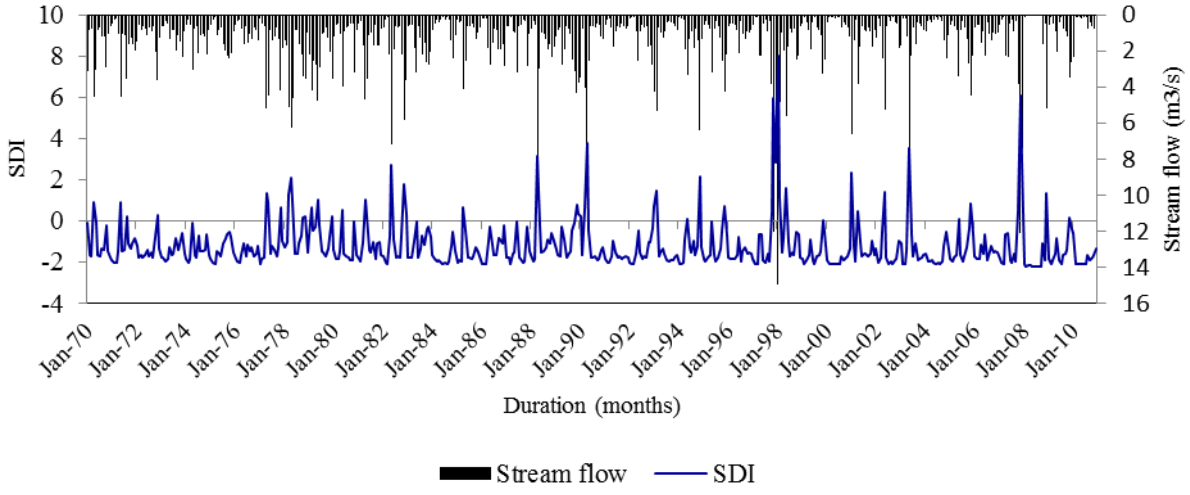


Figure 4. 16: Time series of SDI and stream flow for Kamburu gauging station

4.1.6 Time series SPI

The results for monthly time series SPI and the spatial characteristics of droughts in the upper Tana River basin are presented. The results spatial maps are based on the partitioned basin into four elevations bands; low, lower-middle, middle and high elevations. The results of plotted drought conditions on monthly time series graphs are illustrated using the graphs for meteorogical stations Sagana FCF (ID 9037096), Kerugoya DWO (ID 9037031), Nyeri (ID 9036288) and Naro-moru (ID 9037064) as presented in Figures 4.17 to 4.20.

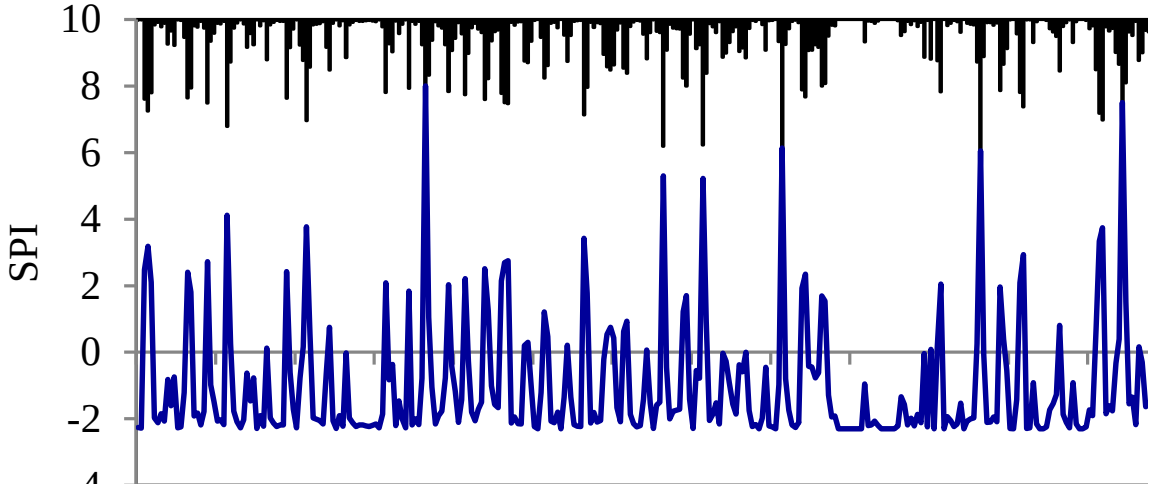


Figure 4. 17: Time series SPI and precipitation for Sagana FCF meteorological station.

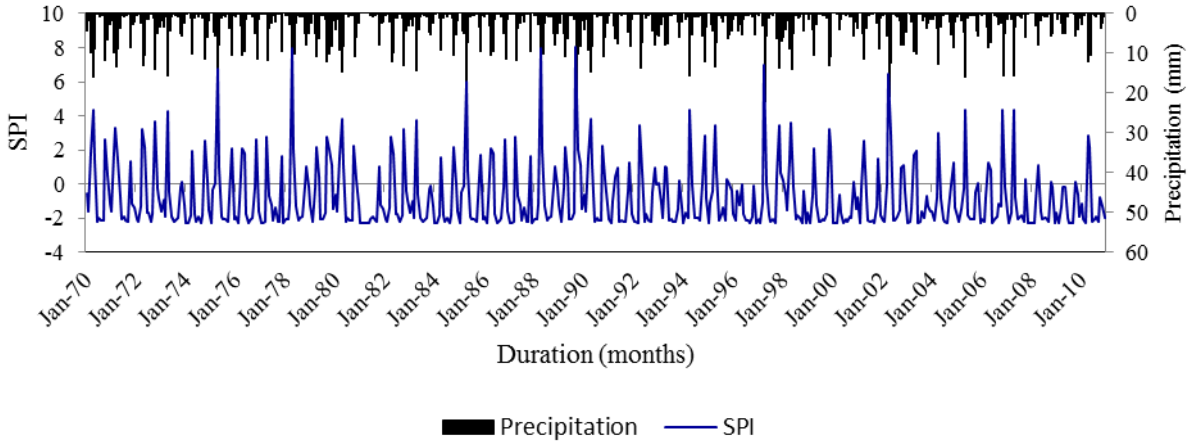


Figure 4. 18: Time series SPI and precipitation at Kerugoya DWO meteorological station

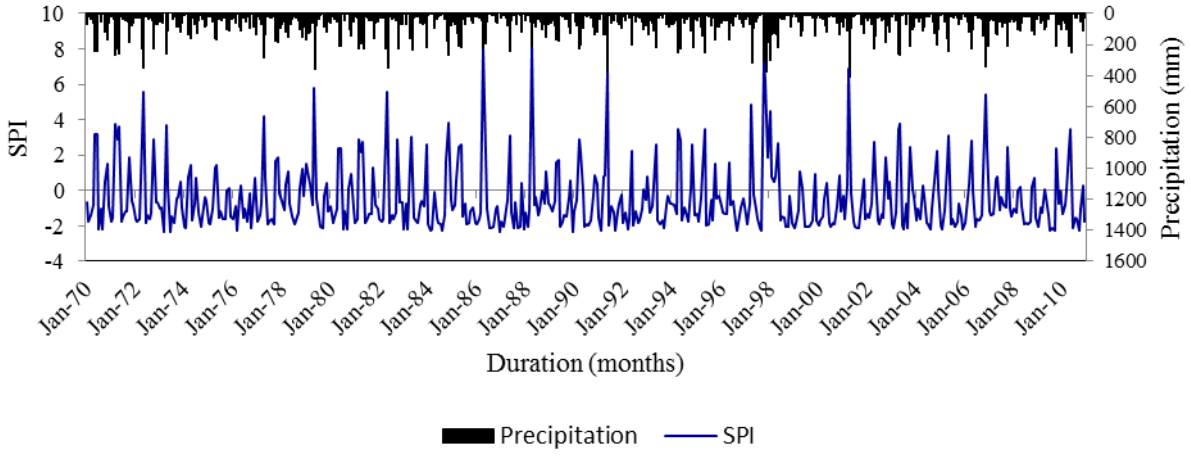


Figure 4. 19: Time series SPI and precipitation for Nyeri meteorological station

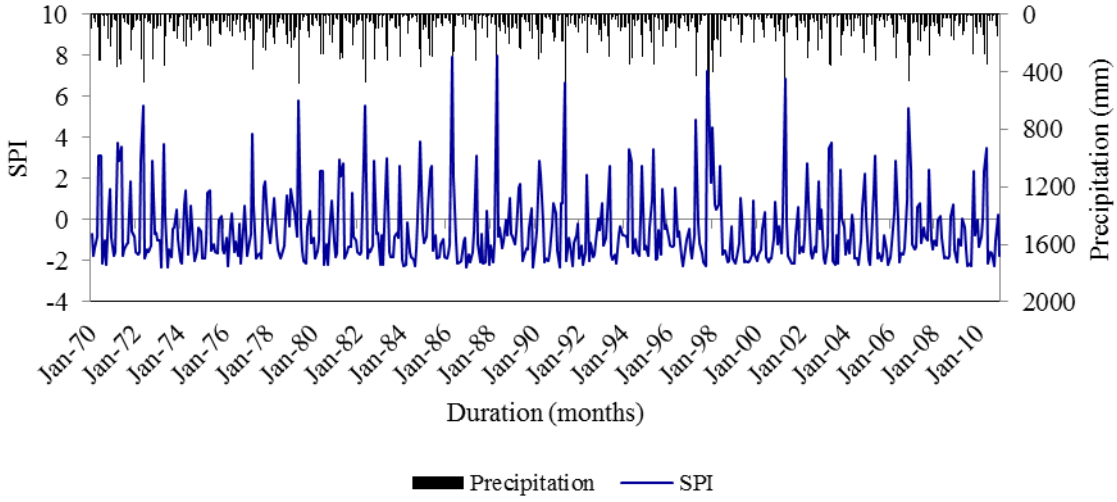


Figure 4. 20: Time series SPI and precipitation for Naro-moru meteorological station

Both time series SPI and precipitation were plotted for ease of comparison as given in Figures 4.17-4.20 for the four meteorological stations. The results show that the SPI varies with the monthly precipitation within the study period and across the river basin. For all the stations, extreme drought events based on SPI were detected using SPI for the periods 1972-1974, 1983-1984, 1987-1988, 1999-2000 and 2011 within which the monthly SPI values were consistently below -2.00. The SPI is used to detect the occurrence of drought (negative values of SPI) or the wetness (positive values of SPI) in a river basin. The other drought conditions detected by SPI for the upper Tana River basin as defined in the SPI criterion presented in Table 3.5 include: severe drought, moderate drought, near normal, moderate wet, very wet and extremely wet conditions. Results of SPI time series within the upper Tana River basin show extreme wetness for 1985-1886, 1992, and 1998 with SPI values being relatively above +2.00.

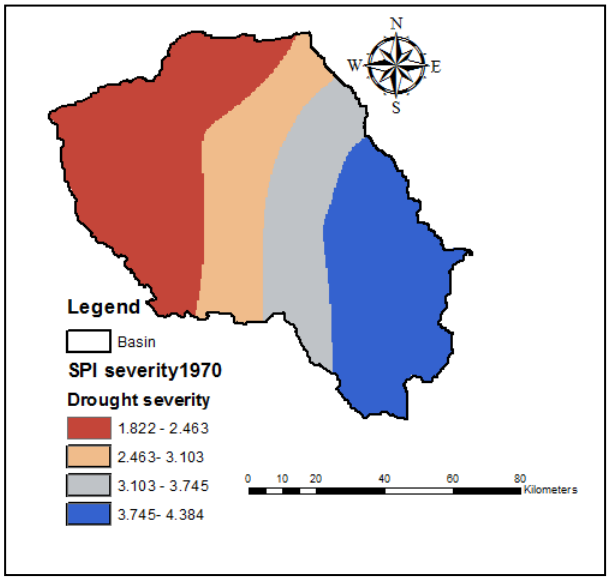
4.1.7 Spatially distributed drought severity based on SPI

Drought severities for the upper Tana River basin were computed and mapped using the Kriging approach for the selected years; 1970, 1980, 1990, 2000 and 2010. From Figure 4.21, it is observed that the spatial drought distribution in the south-eastern areas of the basin exhibit drought severitities ranging from 2.044 to 2.835 and from 4.416 to 5.207. In addition, the results show that the north-western parts of the basin experienced drought severity values of 1.822 to 2.463 and 3.745 to 4.384 for 1970 and 2010 respectively. These results indicate that the south-eastern parts of the basin exhibit the highest drought severities while the north-western areas have the lowest. The spatial variation of drought is comparable with the drought distribution generated in other river

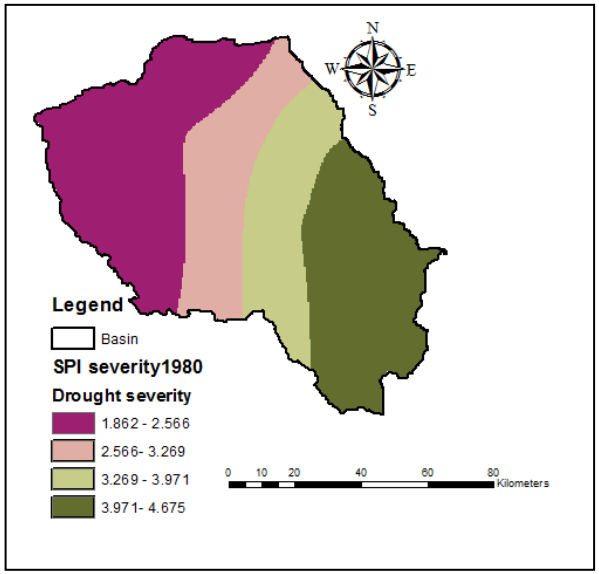
basins for instance by Mishra and Nagarajan (2011) in the Tel river basin and Rajput *et al*. (2014) in the upper Seonath sub-basin.

The difference in drought severities is attributed to the variation of agro-ecological zones found across the basin. Most south-eastern parts of the basin are arid and semi-arid lands (ASALs) that fall in zones V and IV of Kenya's agro-climatic zones. These areas are at low elevations (700-2700 m a.s.l.) and receive low rainfall. On the other hand, the north-western parts which are at higher elevations (2700-4700 m a.s.l.), are humid and fall within agro-ecological zones III to I**.** According to FAO (2006), the agro-ecological zones of the corresponding precipitation to evapo-transpiration (R/E0) Ratio values of the south-eastern and north-western areas are 25-50 and 50-80 respectively (Table 6A, Appendix A).

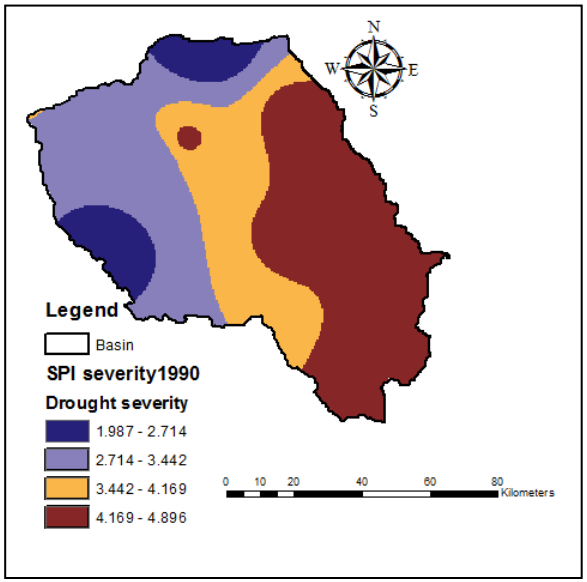
Based on the SPI, the areal-extend of drought severities increased in both the south-eastern and north-western areas from 4868.7 km^2 to 6880 km^2, and 6163.9 km^2 to 6985.5 km^2 from 1970 to 2010 respectively. Between 1970 and 1980, the drought areal-extend is almost the same but a significant increase occurred between 1980 and 2010.



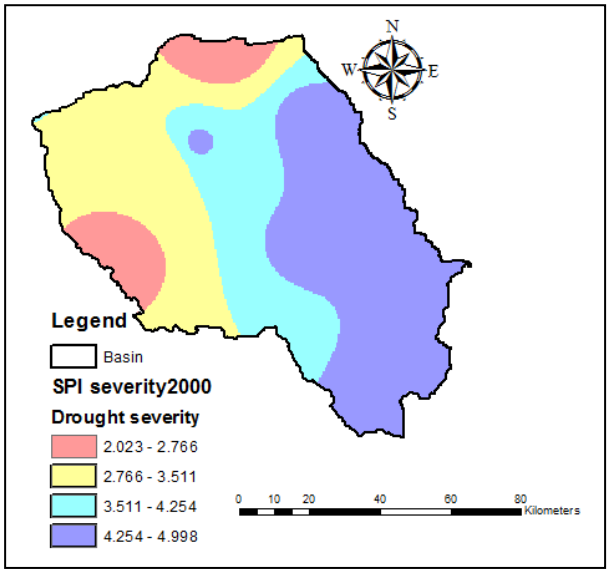
(a)



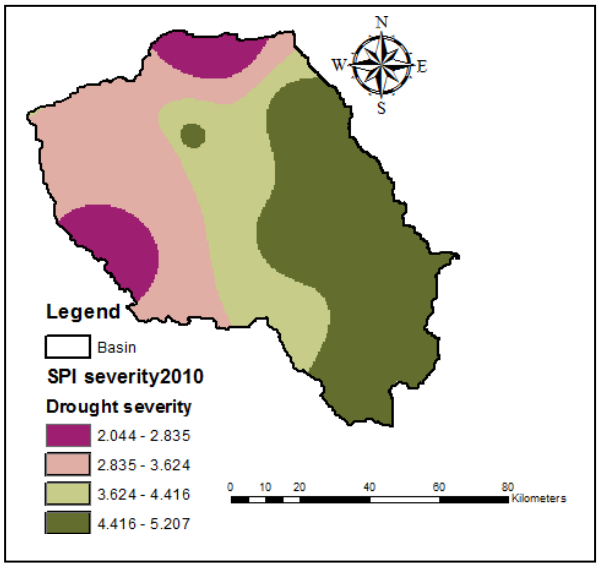
(b)



(c)



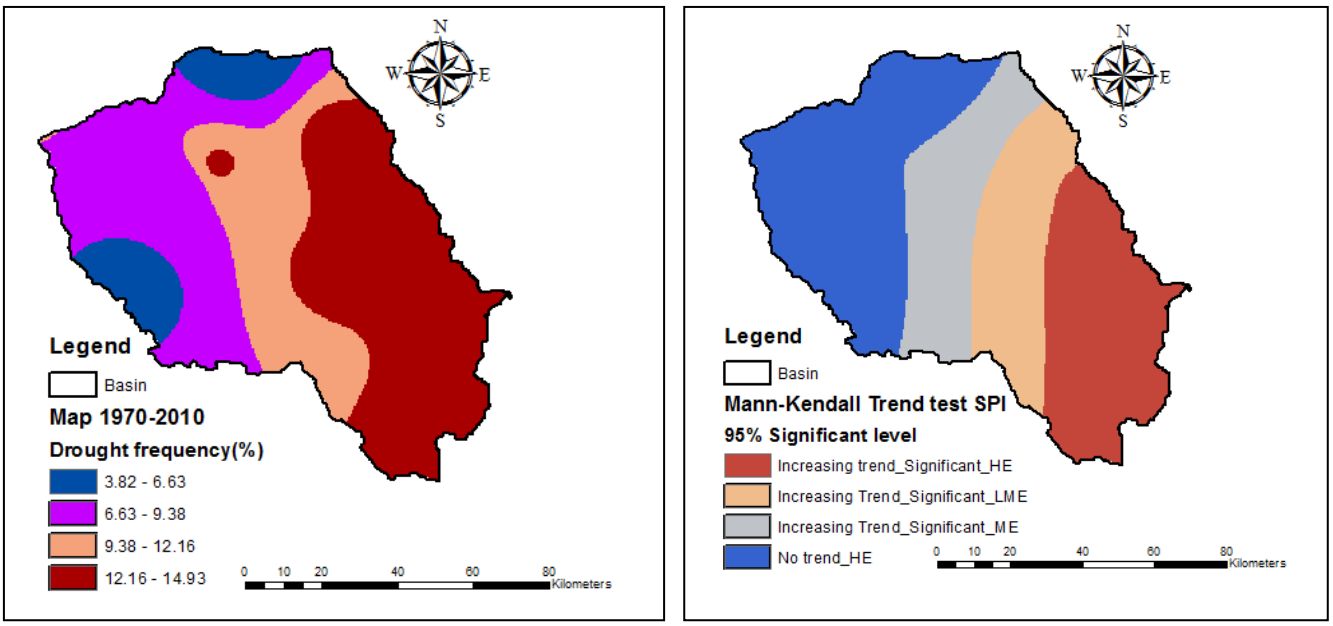
(d)



(e)

Figure 4. 21(a-e): Spatially distributed drought severity based on SPI

From Figure 4.22(a), the results show that the average drought frequency between 1970 and 2010 for the south-eastern and north-western areas ranged from 12.16 to 14.93 and 3.82 to 6.63 respectively. The drought characteristics were also subjected to Mann-Kendall trend test across the basin. Results of the Mann-Kendall test show that drought trend increased in the south-eastern parts of the basin at 90% and 95% significant levels. However, the results given in Figure 4.22(b) shows that there was no significant trend that was detected in the north-western areas. This is an indication that the south-eastern parts are more prone to drought risks than the north-wetsern areas of the upper Tana River basin.



(a) (b)

Figure 4. 22(a and b): Frequency of severe drought and its trend based on SPI

4.1.8 Monthly time series EDI

Monthy time series of EDI for meteorological stations Nyeri (ID 9036288), Kerugoya DWO (ID 9037031), Sagana FCF (ID 9037096) and Naro-moru (ID 9037064) are presented in Figures 4.23 to 4.26.

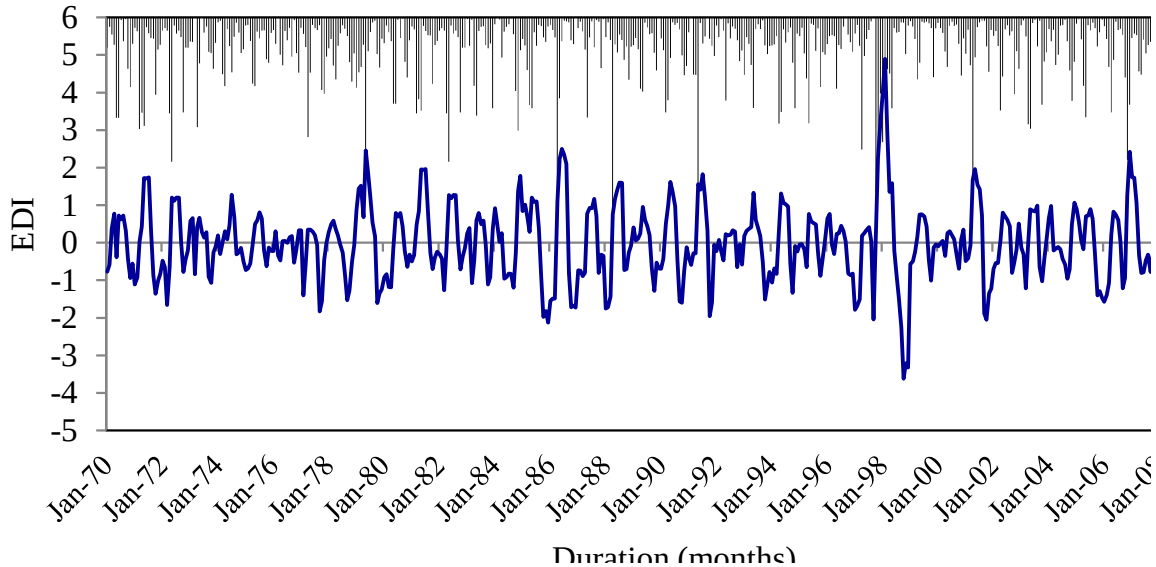


Figure 4. 23: Time series EDI and precipitation for Nyeri meteorological station

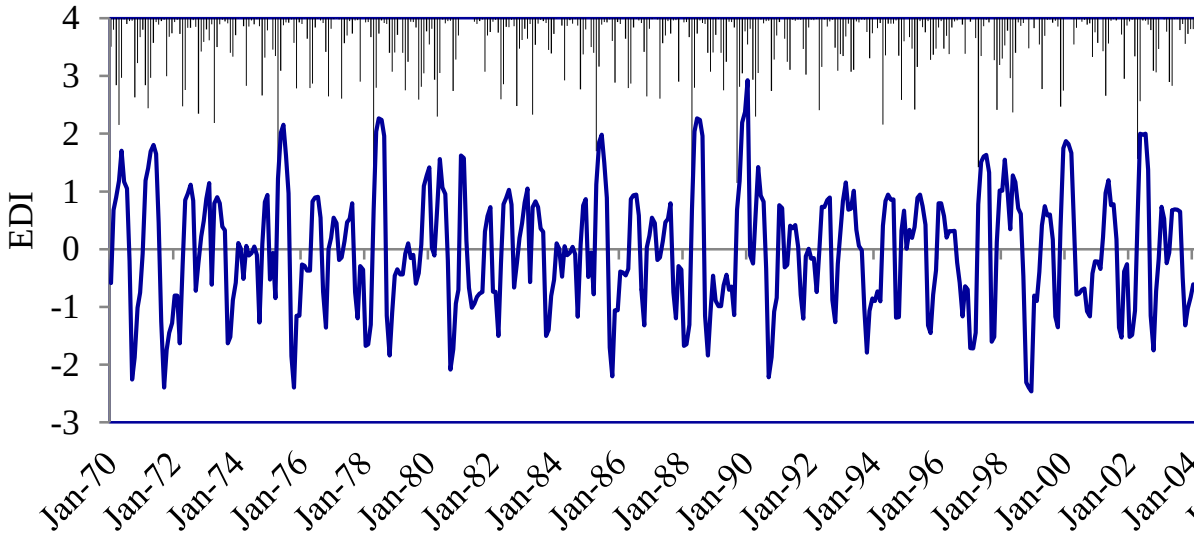


Figure 4. 24: Time series EDI and precipitation forKerugoya DWO meteorological station

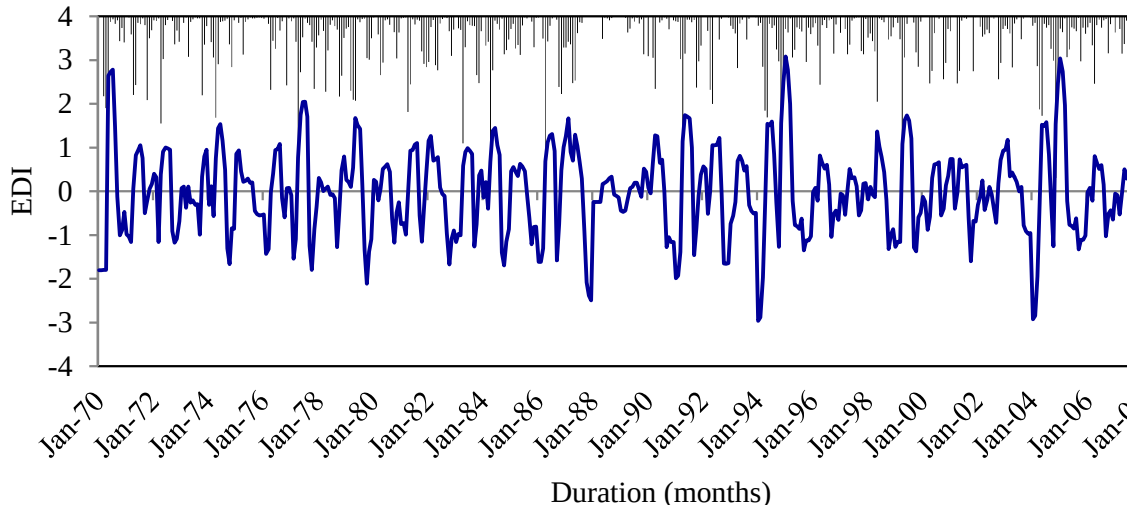


Figure 4. 25: Time series EDI and precipitation for Sagana meteorological station

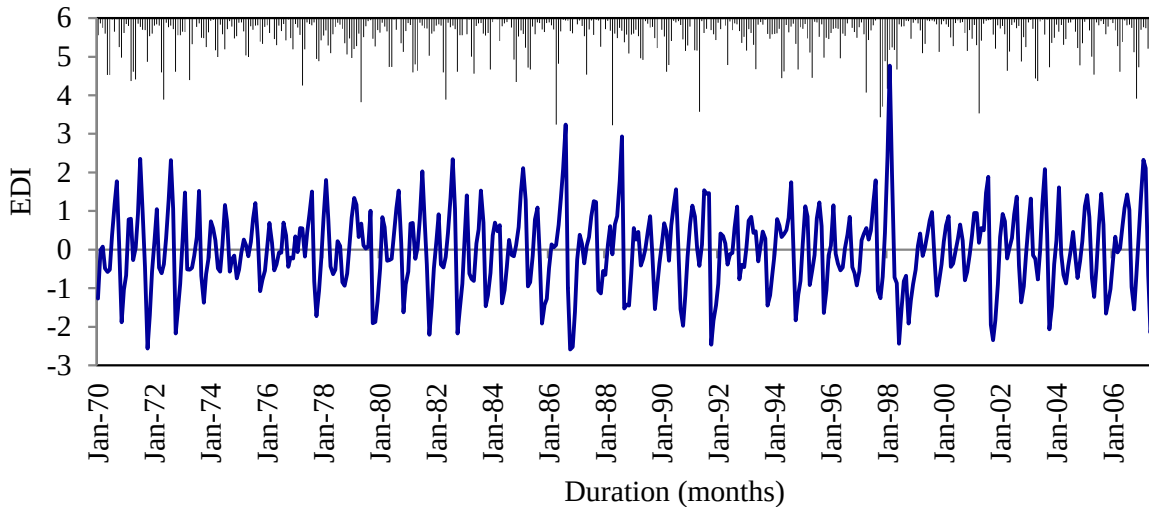
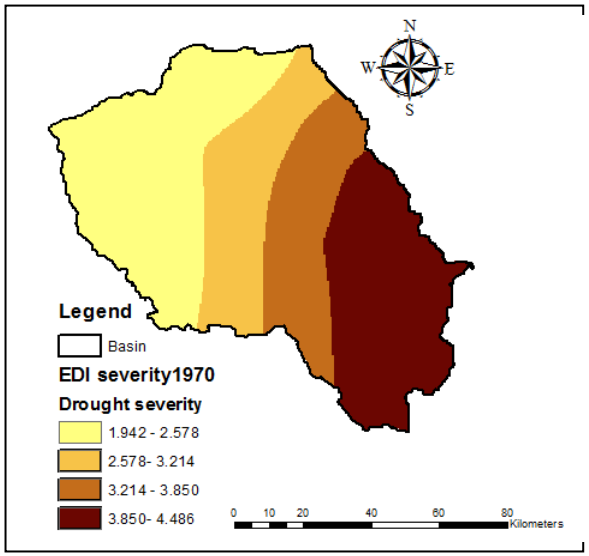


Figure 4. 26: Time series EDI and precipitation for Naro-moru meteorological station

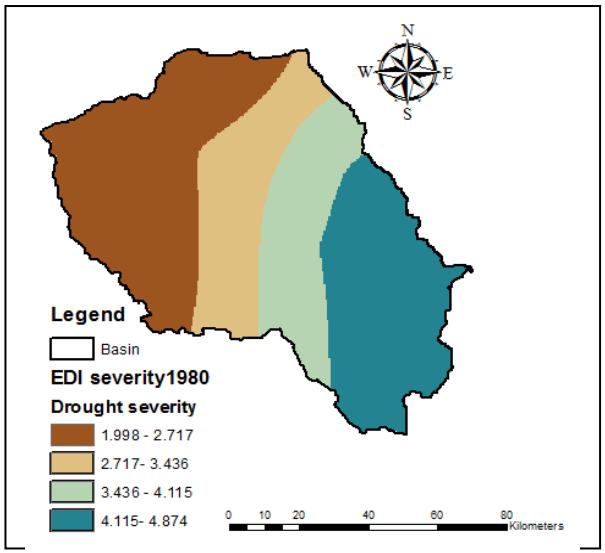
The results of the monthly time series EDI show that this index can be used to detect both the drought and wetness for different years. Typical droughts as presented by this index include the extreme droughts represented by the negative values of -2.5, -2.2, -2.2, -2.5, -2.5, and -2.5 for the years 1972, 1973, 1992, 1994, 2000 and 2010 respectively. At the same time, the index was used to detect the wet conditions of the basin where positive values of +3.0, +3.0 and 4.3 for the years 1986, 1989 and 1998 respectively as illustrated by Figures 4.23-4.26 indicate wetness.

4.1.9 Spatially distributed drougt severity based on EDI

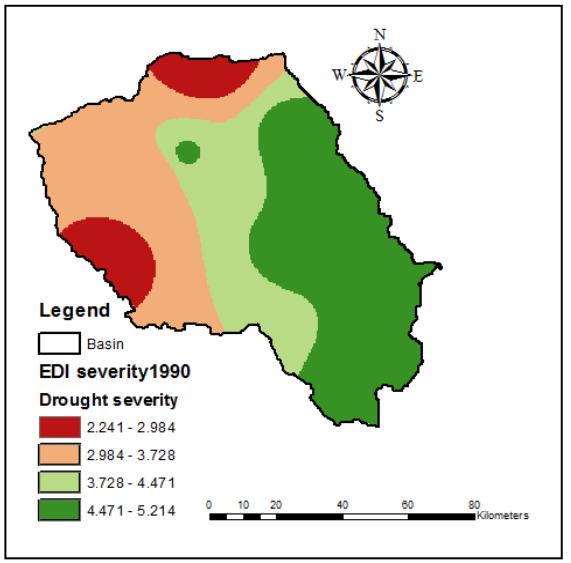
From the results of spatial distribution of drought based on EDI shown in Figure 4.27(a-e), it is observed that the drought severity values differ slightly from those determined using the SPI. It is also noted that the drought severity values in south-eastern areas of the basin range from 3.850 to 4.486 and 4.804 to 5.584 in 1970 and 2010 respectively.



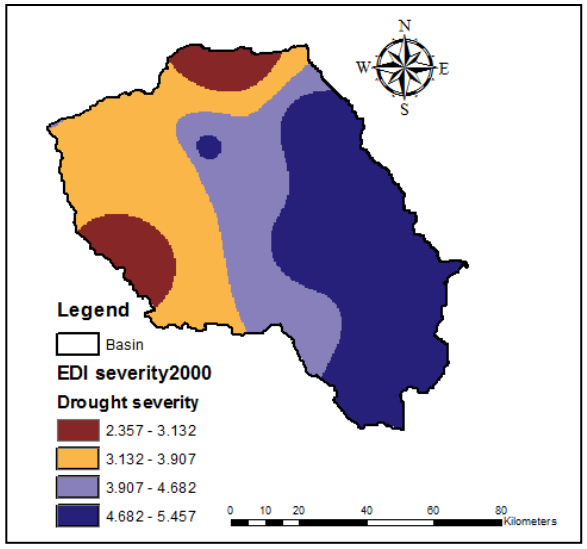
(a)



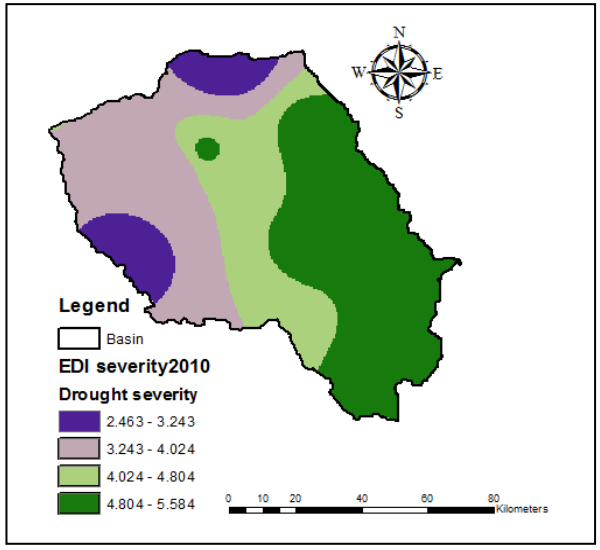
(b)



(c)



(d)



(e)

Figure 4. 27(a-e): Spatially distributed drought severity based on EDI

For the north-western parts, these values range from 1.822 to 2.463 and 3.745 to 4.384 for the years 1970 and 2010 respectively. Although the drought severity based on EDI is generally higher than the SPI, both indices exhibit similar trends in terms of spatial distribution, frequency and Mann-Kendall trend test as given in Figures 4.28 and 4.29.

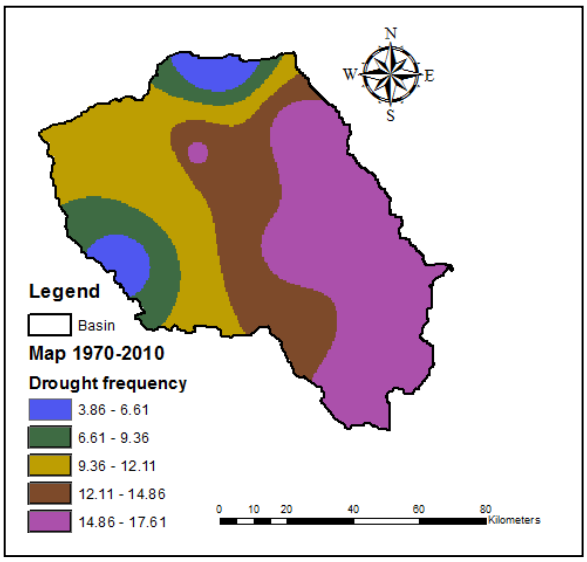


Figure 4. 28: Spatially distributed drought frequency based on SPI

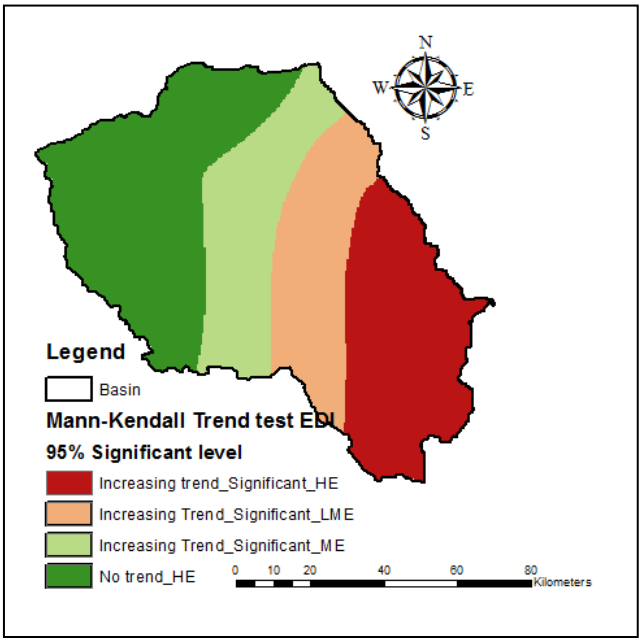


Figure 4. 29: Spatially distributed Mann-Kendall trend test of drought based on SPI

4.1.10 Time series Soil Moisture Deficit Index (SMDI)

The SMDI monthly series data was grouped into dry and wet periods for the computed values of SMDI at 60-cm soil depth that represents the most active Plant rooting zone for plant

evapotranspiration according to FAO (2013). Analyzing the seasonality of drought required definition of the distinct seasons. The seasonal occurrence of droughts was evaluated by analyzing drought events separated into each season (Birkel, 2005). The dry period includes the months of July to September (J-S) and January to March (J-M) while the wet period includes the months of October to December (O-D) and March to June (M-J). The average SMDI was computed for each season to describe the agricultural drought condition for the upper Tana River basin. From the results given in Figures 4.30 to 4.33, it is observed that the seasonal drought magnitudes vary from year to year. Such a trend is consistent with that presented by Nalbantis (2008) and Tsakiris (2009) in Evinos river basin in Greece. The results also show that the dry seasonal SMDI values in the months of January to March are consistently higher than the ones for July to September. This significant discrepancy between the SMDI values for the months of January to March, and July to September is due to the influence of the normal hydrological regime of the basin manifested as long dry period and short dry period.

By comparing the time series results of SMDI for dry and wet seasons for all meteorological stations and using the meteorological station MIAD (ID 9037112) for illustration, it is observed that the values of SMDI for the dry season are consistently lower than those for the wet seasons as given in Figures 4.30 and 4.31. From Figures 4.30 and 4.32, it is evident that the SMDI time series values for MIAD (ID 9037112) located at lower elevations of the basin is less than those for the meteorological station Naro-moru F.G.P (ID 9037064) which is at higher elevations. Thus, it can be deduced that the areas within the lower elevations are more prone to drought risks than those in the higher elevations.

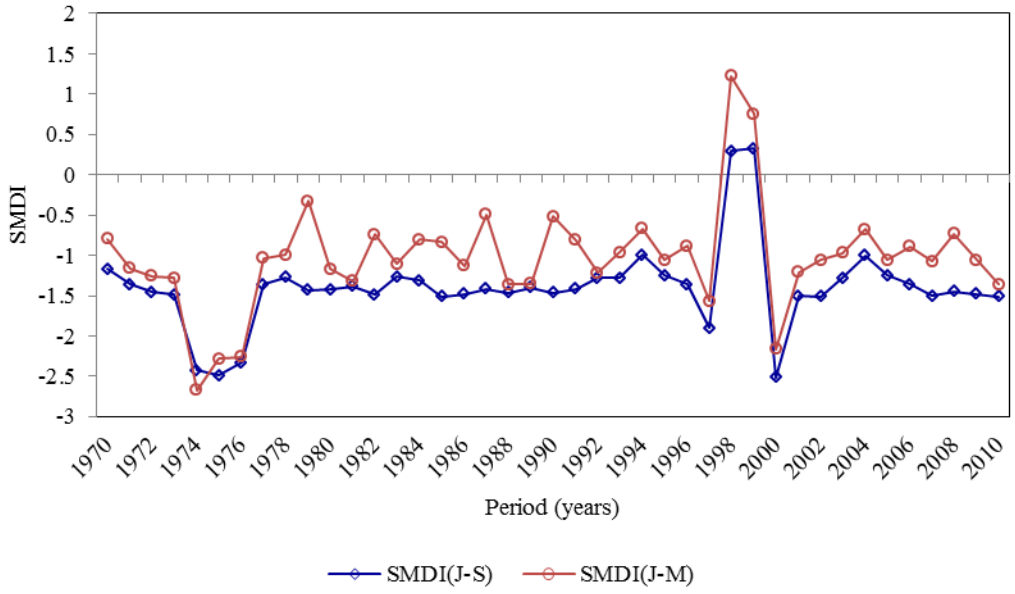


Figure 4. 30: Time series of SMDI for dry season of at MIAD meteorological station

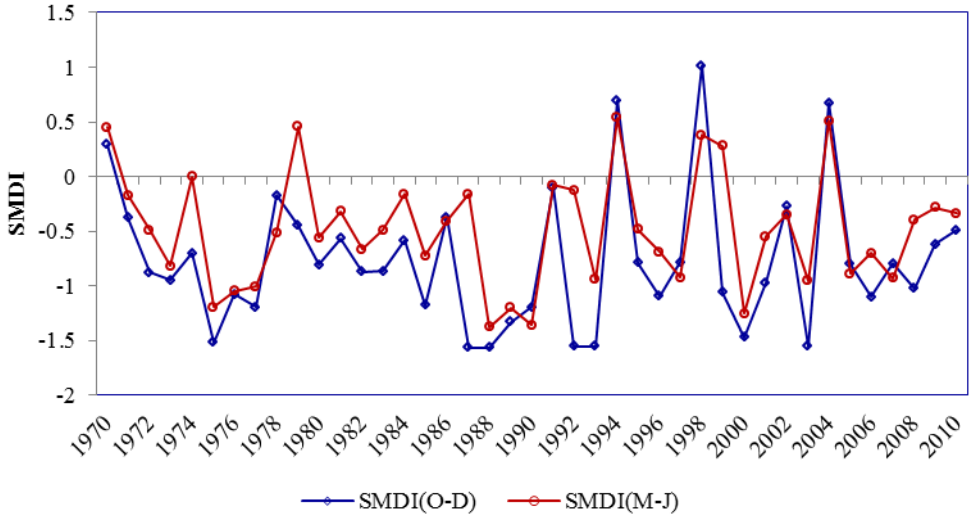


Figure 4. 31: Time series of SMDI for wet season at MIAD meteorological station

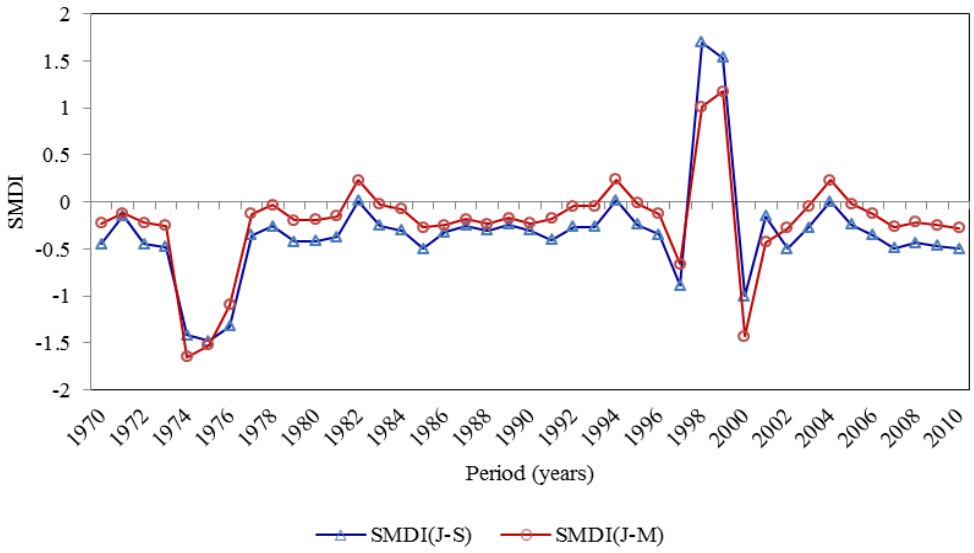


Figure 4. 32: Time series of SMDI for dry season at Naro-moru meteorological station

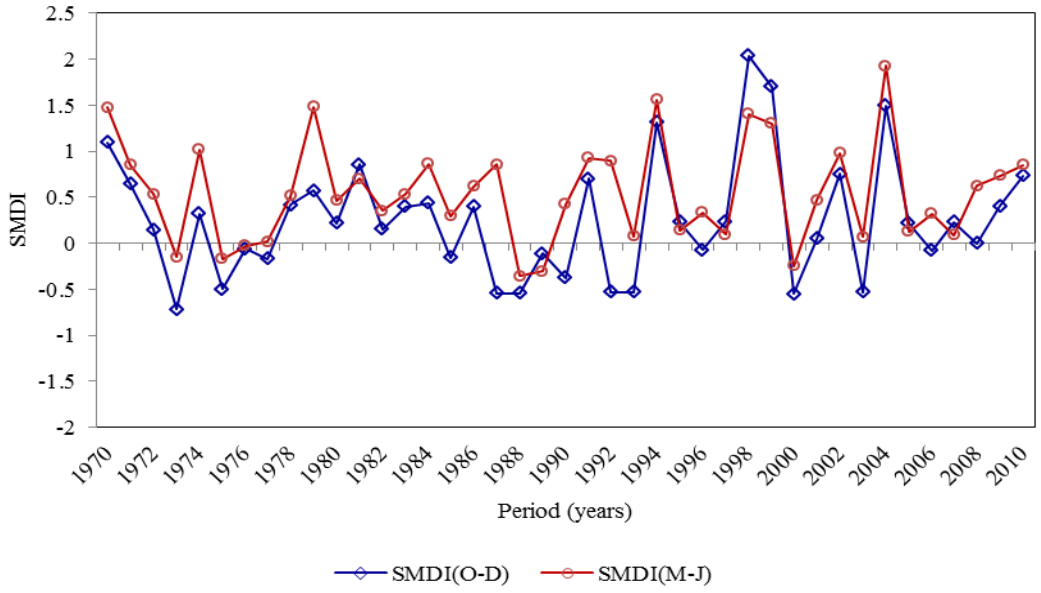


Figure 4. 33: Time series of SMDI for wet season at Naro-moru meteorological station

4.1.11 Spatially distributed drought severity based on SMDI

The results of SMDI spatial distribution of drought as presented in Figure 4.34 show that the minimum and maximum drought severity values in 1970 are from -0.661 to 0.603 and 0.807 to 0.769 respectively. These drought severities are experienced in the north-western and south-eastern areas of the basin as given in Figure 4.34a. From the results given in Figure 4.34c, it is observed that the values increased from -0.715 to -0.658 and -0.886 to -0.829 between 1970 and 2010. This

is an indication that the south-eastern parts of the basin are the most susceptible to droughts as detected by the SMDI while the north-western areas are least prone to the droughts.

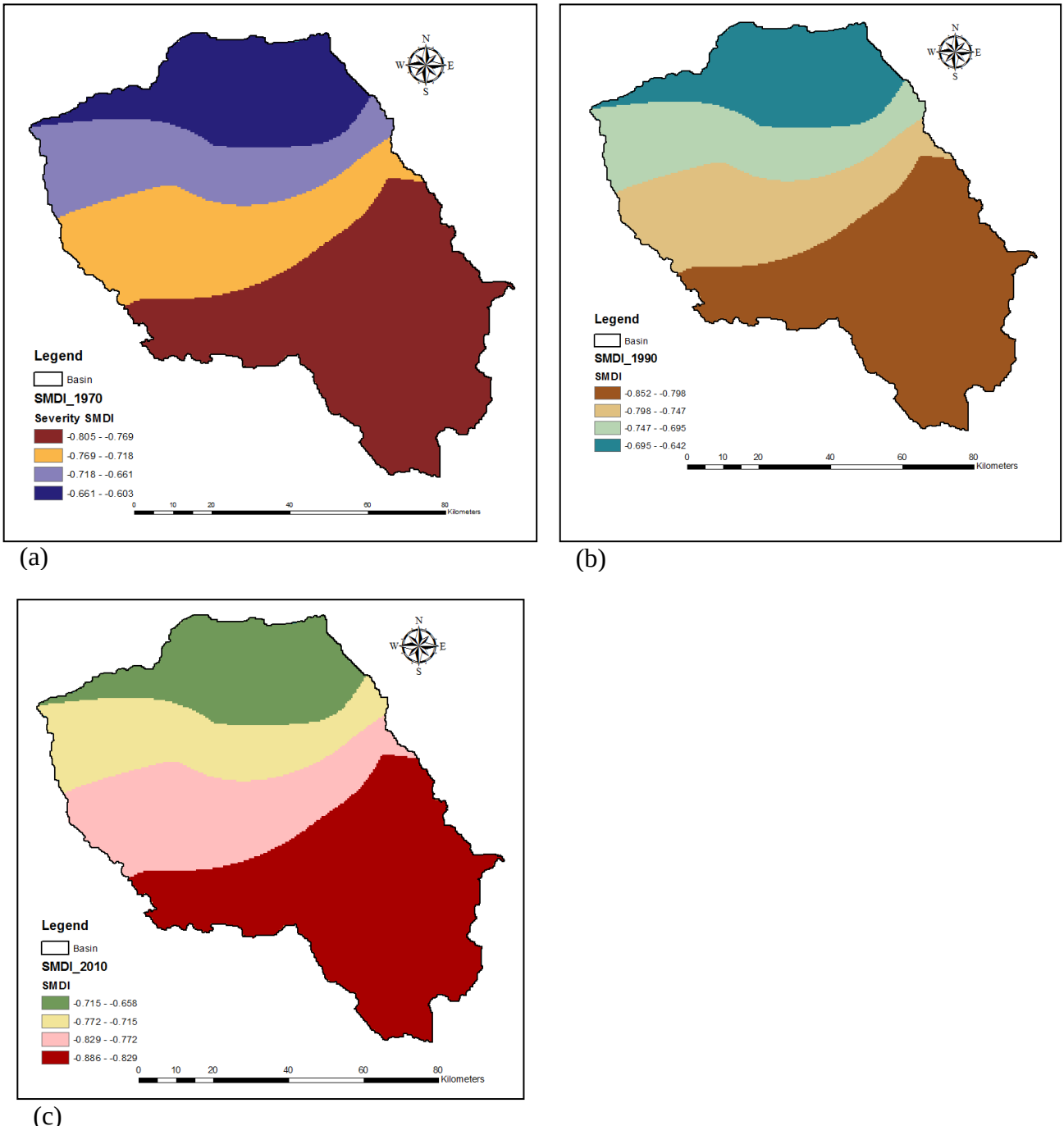


Figure 4. 34(a-c): Spatially distributed drought severity based on SMDI

4.1.12 Time series Palmer Drought Severity Index (PDSI)

From Figures 4.33 to 4.36 it is observed that for dry seasonal PDSI, the values in the months of January to March are constantly higher than the ones for July to September and the trend is

comparable to that of SMDI. However, the range of PDSI values is greater compared to the SMDI. By comparing the time series results of PDSI for dry and wet seasons for MIAD (ID 9037112) meteorological station, it can be seen that the time series plots for the PDSI values for dry season are generally lower than those for the wet seasons (Figures 4.35 and 4.36).

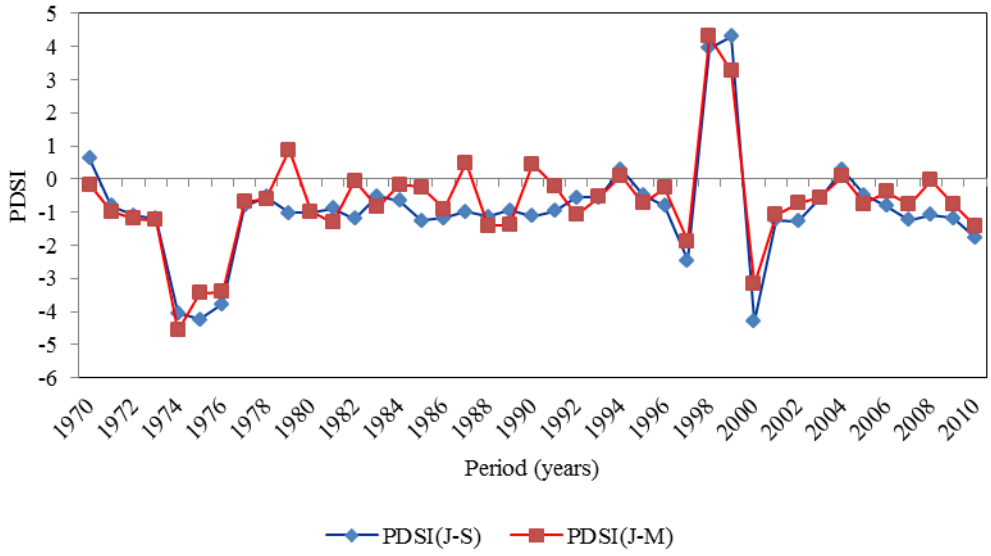


Figure 4. 35: Time series of PDSI for dry seasons of at MIAD meteorological station

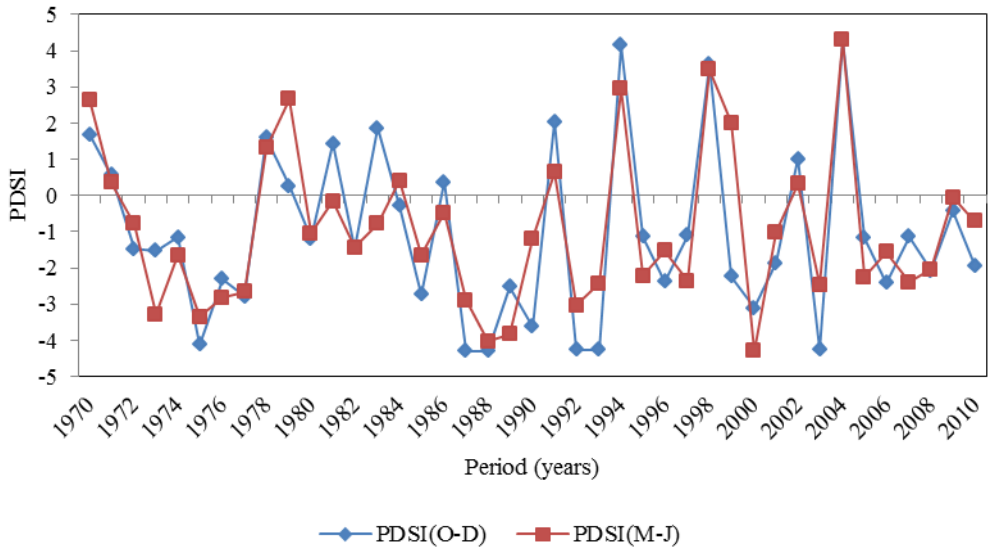


Figure 4. 36: Time series of PDSI for wet seasons at MIAD meteorological station

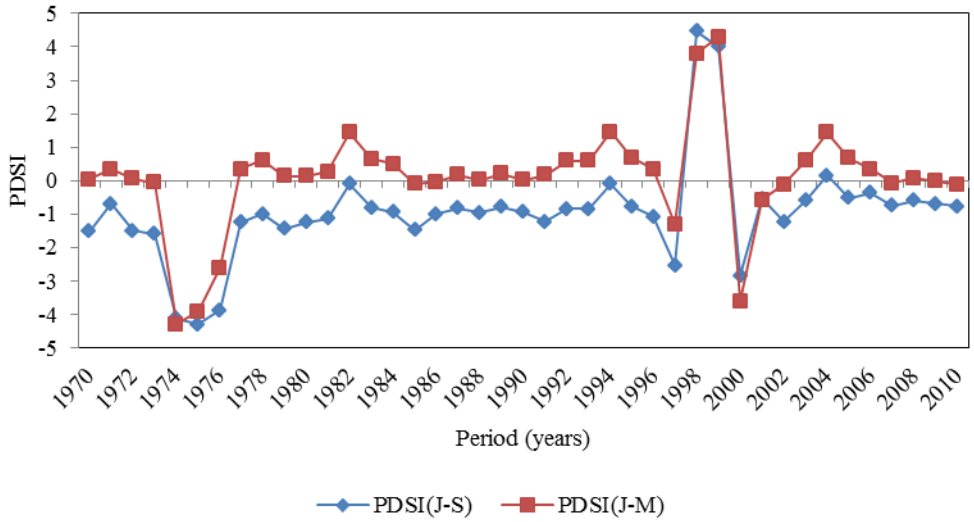


Figure 4. 37: Time series of PDSI for dry seasons at Naro-moru meteorological station

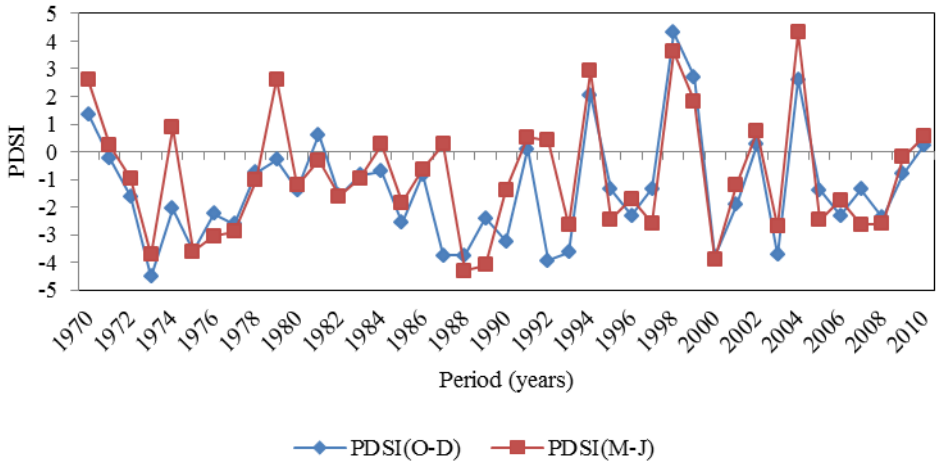
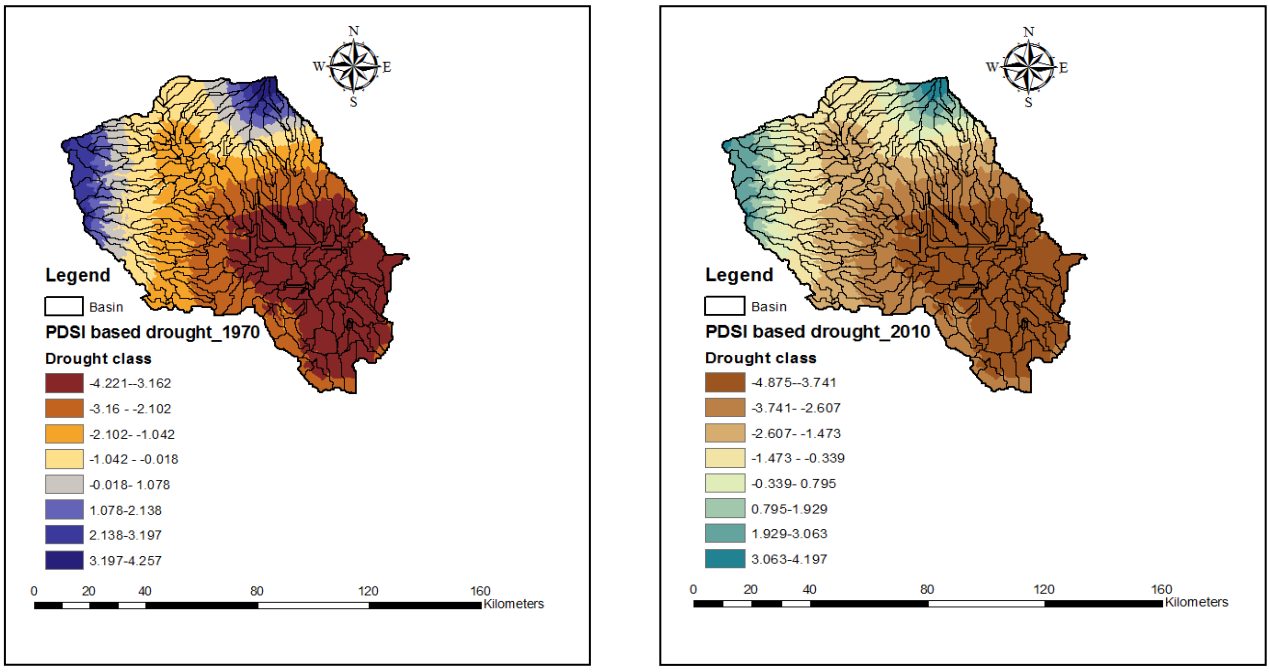


Figure 4. 38: Time series of PDSI for wet seasons at Naro-moru meteorological station

In addition, it can be seen in Figures 4.35 and 4.37 that the PDSI time series values for MIAD meteorological station (ID 9037112) located at the lower elevation of the upper Tana River basin are lower than those for the Naro-moru station (ID 9037064) which is at higher elevation. Thus, similar to the SMDI, the PDSI results indicate that the areas within the lower elevations are more prone to drought risks than those in higher elevations. The average monthly PDSI at Amboni station is presented in Figure 16B, Appendix B whose trend is similar to that of SWSI described in Section 4.1.4.

From the results of spatially distributed drought magnitude, it can be seen that there is a general increase in area under the extreme and severe drought as given by PDSI from 1970 to 2010. For instance, the area under extreme and severe droughts are 3758.01 (21.57%) and 1784.90 (10.25%)

respectively for the year 1970 while the values for 2010 are 4540.36 (26.06%) and 2537.55 (14.57%) respectively (Tables 7A and 8A; Appendix A). The distribution of extreme and severe drought categories dominate in the south-eastern parts of the upper Tana River basin while extreme wet and moderate wet conditions dominate the north-western areas (Figure 4.39). As detected by the other indices, south-eastern parts of the basin have the highest risk of experiencing high drought magnitudes. However, the north-western areas have the lowest drought risks.



(a) (b)

Figure 4. 39(a and b): Spatially distributed magnitude of PDSI-based drought in October

Comparing the results in Figure 4.39 with similar research by Yan *et al*. (2013) in Luanhe River basin, showed that the lowest PDSI values (PDSI< -4.00) are persistently observed in the north-western areas of Luanhe basin. On the other hand, the south-eastern areas of the upper Tana River basin exhibit similar lowest values of PDSI (PDSI< -4.00).

4.1.13 Spatially distributed drought severity based on PDSI

The results of spatially distributed drought severity based on PDSI show that the ranges of maximum and minimum drought severity values in 1970 are -0.868 to -0.804 and -0.675 to -0.610 respectively, and categorized in Figure 23B, Appendix B. These maximum and minimum drought severity values occur respectively in the north-western and south-eastern areas of the basin as summarised in Figure 4.40a. The maximum and minimum values increased from -1.478 to -1.348 and from -1.087 to -.957 in 2010 as presented in Figure 4.40d. This is an indication that the south-

eastern parts of the basin are the most susceptible to droughts while the north-western areas are least prone to the droughts. There is an increase in drought severity over the years of record. The trend in spatial PDSI severity values over time compared closely with the spatial patterns trend explained by Zoljoodi and Didevarasasl (2013). For instance, these authors showed that the PDSI severity values increased from -1.28 (1951-2005) to -7.68 (1999-2002) in Iran. In comparison with the present study, the results show that the PDSI increased from the range -0.675 to -0.610 in 1970 and from -1.087 to 0.957 in 2010 for the north-eastern areas of the upper Tana River basin.

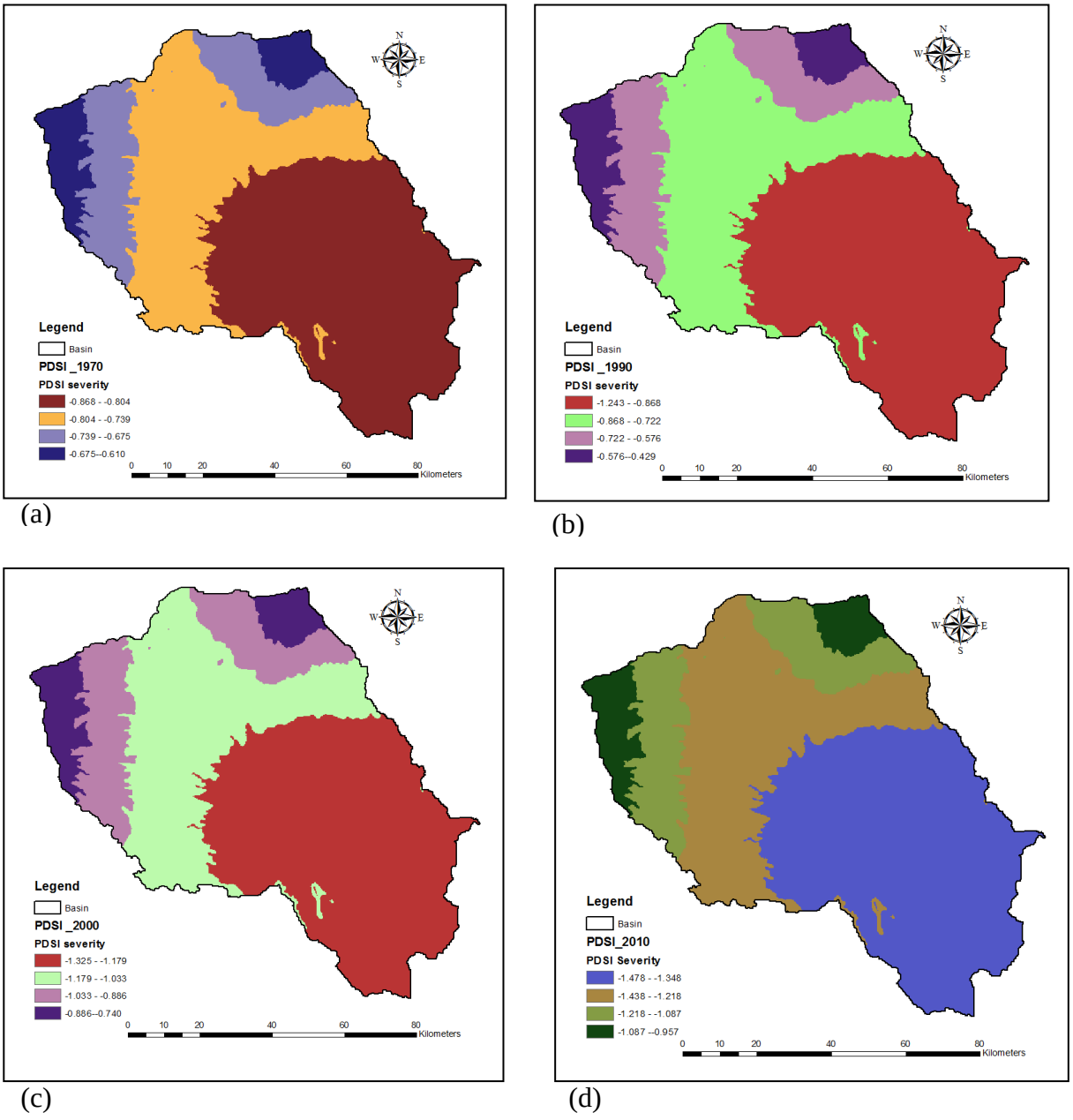


Figure 4. 40(a-d): Spatially distributed PDSI-based drought severity

4.1.4 Characteristics of time series drought conditions

Drought characteristics within the upper Tana River basin were esatablished using characteristic curves that relate drought probability, return period and magnitude. It is possible to determine the probability and frequency of drought with defined severity for different gauging stations. For instance, Figure 4.41 presents the results of drought characteristics for Amboni gauging station (ID 4AB05) where the probability of occurrence of severe (-3.00 to -3.99) and moderate (-1.00 to -1.99) droughts is 0.18 and 0.51 while the return periods for the two drought conditions are 8 years and 1 year respectively. From Figure 4.41, it can also be seen that the probability of drought events increased linearly with increase in magnitude of SWSI while the return period of drought events increased exponentially with decrease in magnitude of SWSI. The associated regression coefficients of the resulting linear and exponential functions of 0.984 and 0.980 respectively indicate that the drought probability and magnitude, and the return period and magnitude have a strong correlation. Thus the functions in Figure 4.41 can effectively be used in determining the probability and return perion for any drought severity in the basin.

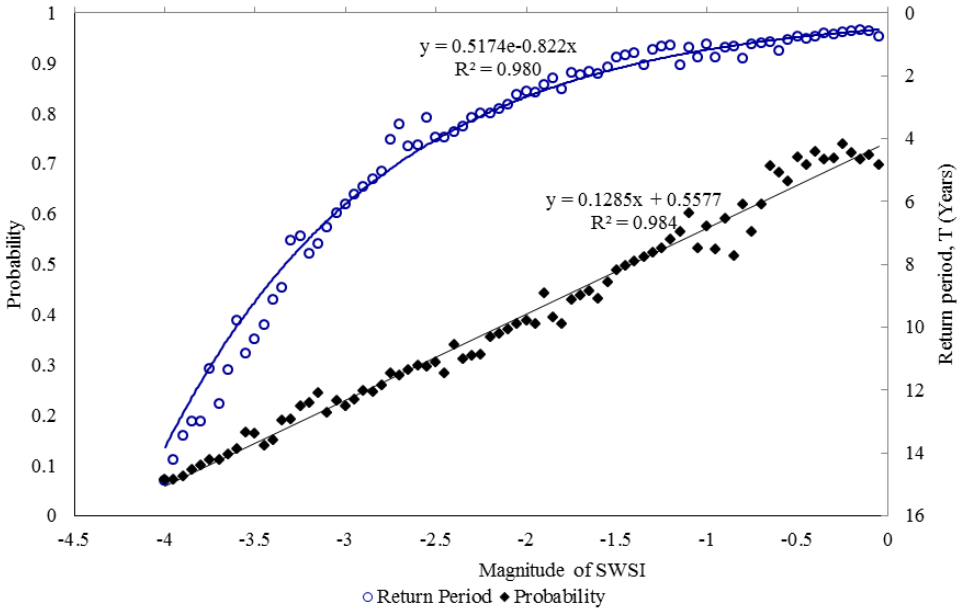


Figure 4. 41: Drought characteristic curves of SWSI for Amboni hydrometric station

In addition, the drought Severity-Duration-Frequency (SDF) curves for the upper Tana River basin were fitted. The results of the SDF curves at Amboni gauge station (ID 4AB05) as presented in Figure 4.42 show that for a 50-year return period, cumulative drought severity of 5.5, 7.2, 9.8, 13 and 17 correspond to 2, 4, 5, 6 and 8-months duration. The drought events of different return

periods exhibit a similar trend. It is thus inferred that the higher the cumulative drought severity for each return period, the higher the duration. The trend of the results presented in Figure 4.42 is similar to the SDF curves developed by Dalezios *et al*. (2000) using cumulative drought severity of PDSI for Volos hydrometric stations in Greece.

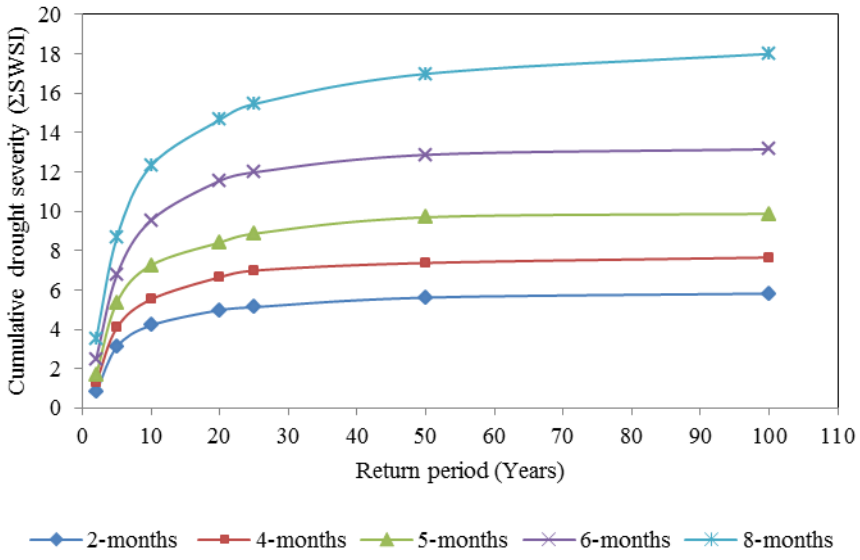
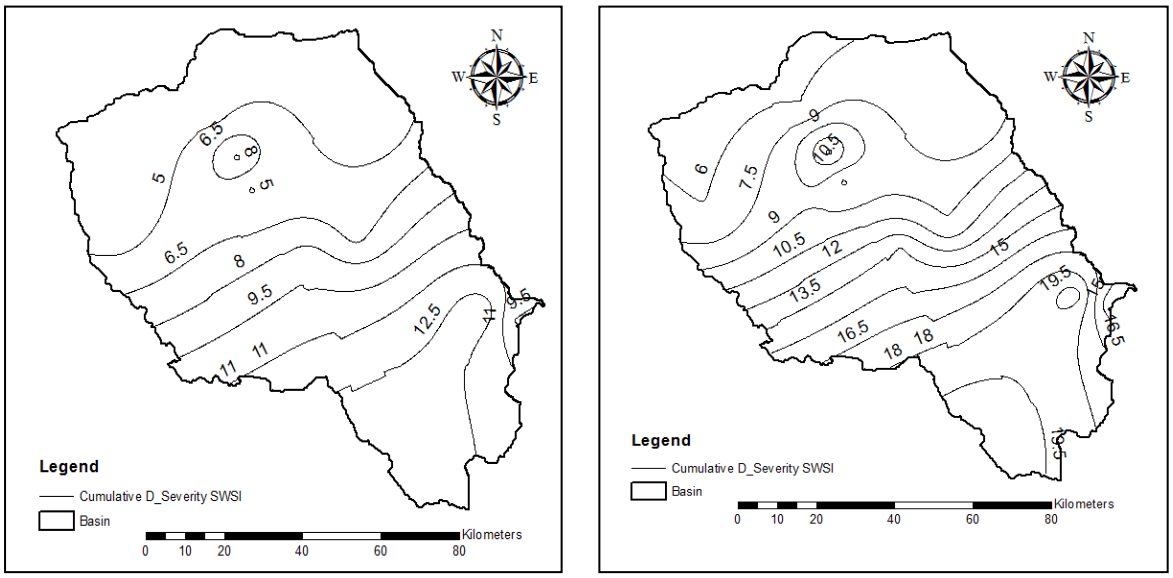


Figure 4.42: Severity-Duration-Frequency (SDF) curves at Amboni based on SWSI

Figure 4.43 shows the maximum cumulative drought magnitude based on SWSI for the duration of 2 months which is expected to be equaled or exceeded once every 10 and 50 years.



(a) (b)

Figure 4. 43: Drought-severity contour map for (a) 10 and (b) 50-year return period

Figure 4.43 present spatial drought for the upper Tana River basin. When such maps are developed using precipitation, as it is in the case of intensity-duration-frequency curves, they are referred to as isopluvial maps. However, in this study on drought, the maps are called iso-drought maps. The lines in Figure 4.43 define points with same drought magnitude for different combinations of one return period for a defined duration. From Figure 4.43, it can be deduced that the expected drought magnitude for a 50 year return period is higher than that of 10 year return period. The south-eastern parts of the basin exhibit higher values than the north-western and north-eastern areas. This means that the south-eastern areas are at the highest risk of drought than the other parts of the basin.

The mean drought frequency was determined and the results are presented in Figure 4.44 and Table 9A in Appendix A.

.

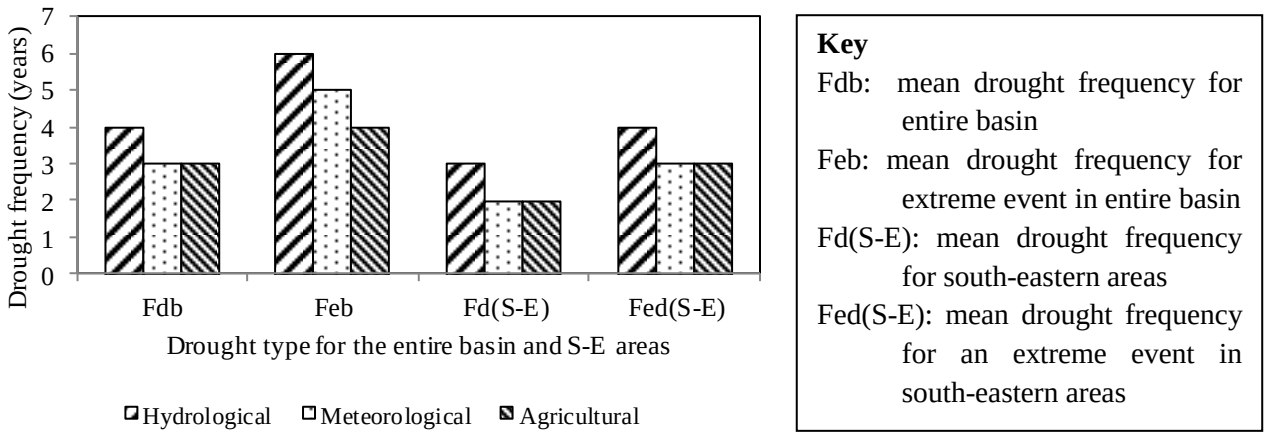


Figure 4. 44: Mean drought frequency for entire basin and south-eastern areas

The results given in Figure 4.44 show that the upper Tana River basin experiences at least a hydrological, meteorological and agricultural droughts on average every 4, 3 and 3 years respectively. Results indicate that an extreme agricultural, meteorological and hydrological drought event is experienced every 4, 5 and 6 years respectively. The results show that there is more frequent drought occurrence in the south-eastern areas of the basin. For instance, the results show that the hydrological drought occurs every 3 years while its extreme event occurs every 4 years.

The Stream Flow Drought Index (SDI) was also fitted for different return periods and then plotted together with the streamflow. Results of the fitted curves show that the absolute SDI increases with the increase in return period as given in Figure 4.45 for gauging stations; Amboni (4AB05), Tana Sagana (4BC02), Sagana (4AC03) and Gura (4AD01) respectively. Results show that the

hydrological droughts represented by absolute *SDI* of 0.667and 1.265 are equaled or exceeded on average once every 2 and 1000 years respectively. The same applies to the other hydrological droughts of defined absolute *SDI*. The results of the absolute SDI were tied to their respective stream flow magnitudes as shown in Figure 4.45. This is critical in water resources management since water managers can match specific absolute drought index with corresponding quantity of stream flow.

Generally the results show that, the minimum and maximum drought episodes occurred within the areas where the gauging stations 4AB05 and 4CC03 are located in highest and lowest elevations, with absolute *SDI* ranging from 0.667 to 1.265 and 1.213 to 2.42 for 2 and 1000-year return period respectively. Using SDI criterion (Table 3.4) and results from Figure 4.45, the critical points are identified. The critical point is the level of hydrological drought beyond which the water facilities are significantly affected by drought. For instance, the critical point for gauging station 4AB05, as shown using the dotted line, coincides with the return period of 28 years with absolute SDI of 0.92 and stream flow of 3.6 m^3/s (Figure 4.45a), while that of 4BC02 is 20 years with absolute SDI of 1.2 and stream flow of 45 m^3/s (Figure 4.45b). In this case, if a water resource system is to be designed for example at the Tana Sagana gauging station (ID. 4BC02) and Amboni (ID. 4AB05), the water storage systems should be designed to cope with drought corresponding to absolute severity 1.2 and of 0.92, with return periods of less or equal to 20 and 28 years respectively.

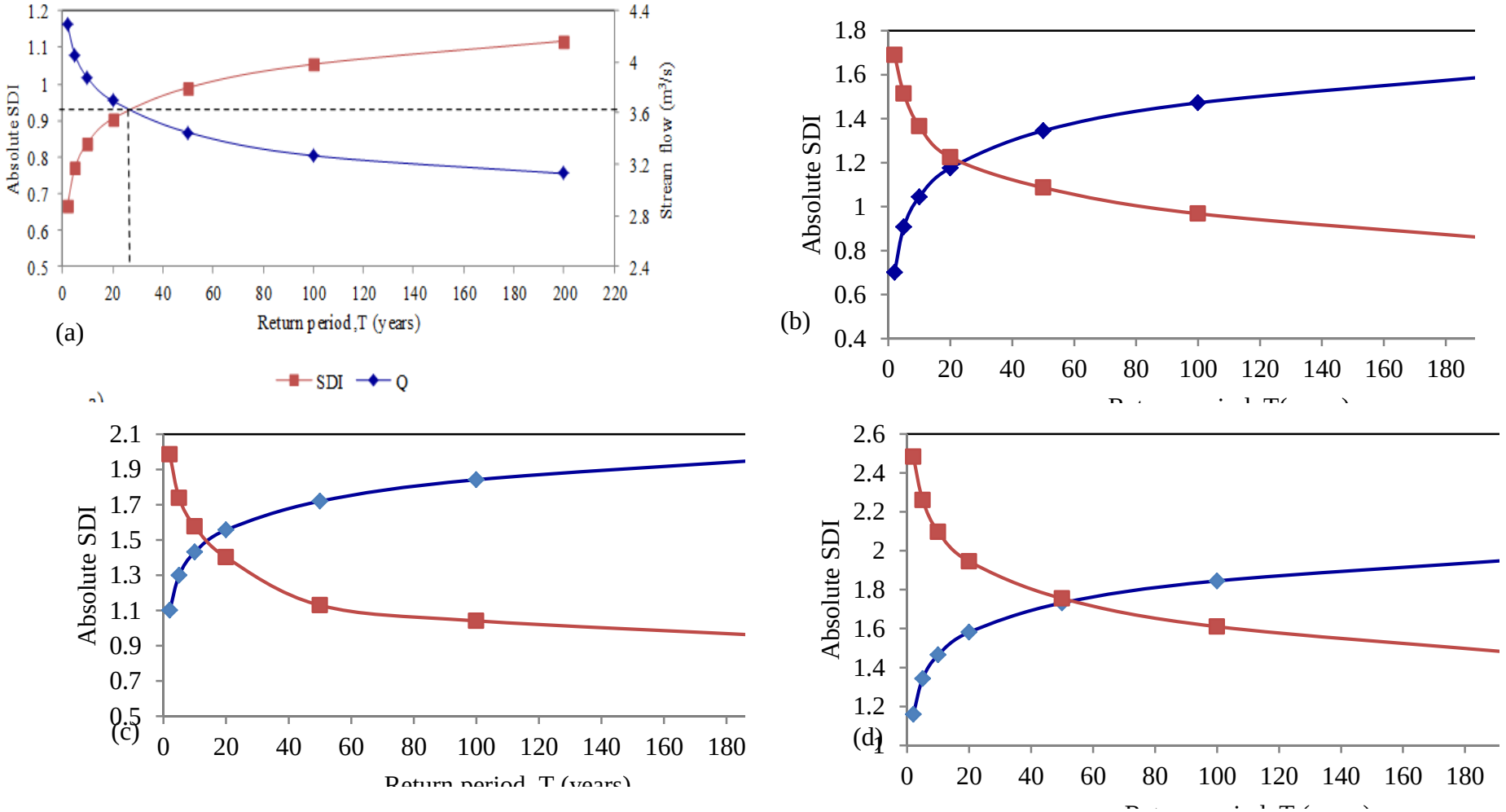


Figure 4. 45: Relation between the Q_m, *SDI* and *T* for different gauging stations

From Figure 4.46, the results show that the ratio of $Q_T/\overline{Q}$ represented by Y (Equation 3.56) increases with C_v for different return periods; 5, 10, 20, 50 and 100-year return periods represented as T5, T10, T20, T50 and T100. This confirms that the Gumbel method is also applicable in drought frequency estimation as supported by Al-Mashindani *et al*. (1978) on development of a simple function of flood estimation in numerous river systems such as Tigris, Euphrates, Adhaim, Diyala and Yamuna.

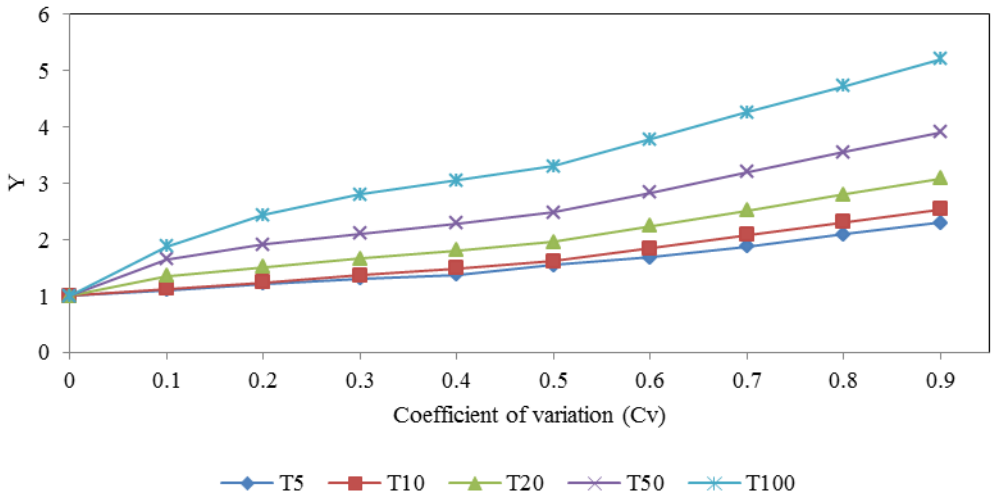


Figure 4. 46: Plot of Y versus C_v at Kamburu station in the upper Tana River basin

4.2 Forecasted drought using DIs and ANNs

The results of DIs under the hydrological, meteorological and agricultural drought types were investigated. As shown in Tables 4.2 to 4.10, the performance of different artificial neural network models in forecasting of different drought indices at different gauging stations vary with lead-time. In this study, the lead times were categorized into short (1 to 3 months), medium (4 to 9 months) and long (10 to 24) lead time as presented in Tables 4.2 - 4.10.

4.2.1 Hydrological drought forecasts

The SWSI and SDI were the two indices used for hydrological drought forecasting at the gauge stations. For instance, at Yatta furrow gauging station (ID 4CC03), the SWSI values show that the best forecasting models give correlation coefficient (R) values of 0.857, 0.764, 0.726, 0.694, 0.674, 0.503, and 0.382 for 1, 3, 6, 9, 12, 18 and 24-month lead times respectively. The corresponding R values for the same lead times based on SDI are slightly lower than R values for SWSI. Generally, the performance of the models in forecasting drought in the upper Tana River basin, declined with increase in lead time. Such a trend conforms to that presented by Mishra and Desai (2006) where

these authors showed a significant decrease of ANN models performance in drought forecasting from 1 to 6-months lead time in Kansabati River Basin, India.

To identify the most appropriate model, different network models were formulated and tested for forecasting each drought index using the performance measures. The formulated models comprise input combination that are in conformity with those illustrated by Morid *et al*. (2007) and Mishra and Nagarajan (2012). From the results, the best forecasting model was adopted whose results are presented for each station in this study. For example, the best network model for forecasting SWSI (Table 10A, Appendix A) is expressed as:

$$SWSI_{(t+n)} = f\left(SWSI_t, SWSI_{t-1}, SWSI_{t-2}\right),\left(P_t, P_{t-1}, P_{t-2}\right) \quad (4.3)$$

Where;

$SWSI_{t+n}$ = the forecasted SWSI for lead time *n* months

$SWSI_t$ = the SWSI at time *t*

t, *t-1* and *t-2* = present, one month before and two months before the current month

For one-month lead time, the best ANNs architecture for forecasting SWSI are 6-5-1, 6-3-1, 6-5-1, 6-9-1, 6-10-1, 6-3-1, 6-7-1 and 6-10-1 for gauging stations 4AB05, 4AC03, 4AD01, 4BC02, 4CC03, 4DA10, 4DC03 and 4EB01 respectively. The ANNs acrchitecture defines the number of neurons in the structure. For instance, the architeture of 6-10-1 means 6, 10 and 1 neurons in the input, hidden and output of the ANN. The ANNs architecture for all the other stations and different lead times are shown in the Tables 4.2-4.4. The best network model for forecasting SDI is also shown in Table 10A, Appendix A. These findings show that the forecasting ability of all the network models decrease with increase in lead-time at all stations as indicated by different performance criteria; R, RMSE, MAE, d_1 and NSE that were used in this study. In addition, the SWSI performed better than SDI in hydrological drought forecasting across all lead times. The SWSI and SDI for instance, depicted R values of 0.752 and 0.732 for station 4AB05 for one-month lead time.

Table 4. 2: B**est ANNs for short-term drought forecasting of SWSI and SDI**

Station ID[a]	SWSI						SDI					
	ANN Arch[b]	R	RMSE	MAE	d1	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
1-month lead time												
4AB05	6-5-1	0.752	0.367	0.244	0.805	0.698	5-4-1	0.732	0.372	0.253	0.745	0.654
4AC03	6-3-1	0.721	0.373	0.249	0.773	0.668	5-3-1	0.703	0.376	0.256	0.719	0.623
4AD01	6-5-1	0.748	0.365	0.242	0.799	0.695	5-4-1	0.728	0.372	0.250	0.745	0.646
4BC02	6-9-1	0.685	0.378	0.255	0.735	0.632	5-5-1	0.666	0.381	0.261	0.683	0.594
4CC03	6-10-1	0.857	0.298	0.175	0.924	0.822	5-6-1	0.839	0.293	0.176	0.852	0.759
4DA10	6-3-1	0.787	0.324	0.209	0.833	0.734	5-4-1	0.762	0.331	0.209	0.771	0.681
4DC03	6-7-1	0.762	0.363	0.239	0.814	0.709	5-5-1	0.740	0.365	0.243	0.751	0.662
4EB01	6-10-1	0.854	0.312	0.189	0.907	0.802	5-6-1	0.821	0.308	0.188	0.822	0.745
3-months lead time												
4AB05	6-5-1	0.662	0.376	0.259	0.717	0.609	5-4-1	0.640	0.382	0.269	0.666	0.582
4AC03	6-9-1	0.643	0.395	0.277	0.696	0.591	5-5-1	0.623	0.415	0.292	0.651	0.555
4AD01	6-5-1	0.658	0.384	0.266	0.713	0.606	5-4-1	0.636	0.391	0.283	0.668	0.575
4BC02	6-5-1	0.612	0.399	0.282	0.667	0.559	5-3-1	0.594	0.408	0.296	0.622	0.528
4CC03	6-10-1	0.764	0.312	0.194	0.816	0.712	5-6-1	0.734	0.319	0.201	0.764	0.676
4DA10	6-3-1	0.688	0.343	0.225	0.745	0.636	5-6-1	0.664	0.368	0.230	0.691	0.606
4DC03	6-10-1	0.669	0.383	0.265	0.721	0.619	5-11-1	0.647	0.390	0.276	0.667	0.589
4EB01	6-10-1	0.775	0.329	0.211	0.786	0.692	5-6-1	0.716	0.331	0.214	0.729	0.664

[a]The station IDs defined as 4AD01, 4AC03, 4AD01, 4BC02, 4CC03, 4DA10, 4DC03 and 4EB01refer to Amboni,Sagana, Gura,Tana sagana, Yatta furrow, Nyamindi, Rupingazi and Kamburu gauge stations respectively

[b]The short form of 'Architecture'

Table 4. 3: B**est ANNs for medium-term drought forecasting of SWSI and SDI**

Station ID[a]	SWSI						SDI					
	ANN Arch[b]	**R**	**RMSE**	**MAE**	d1	**NSE**	ANN Arch[b]	R	RMSE	MAE	d1	NSE
6-months lead time												
4AB05	6-5-1	0.632	0.396	0.274	0.684	0.579	5-4-1	0.631	0.401	0.285	0.637	0.538
4AC03	6-10-1	0.613	0.418	0.296	0.665	0.561	5-9-1	0.613	0.425	0.310	0.622	0.522
4AD01	6-5-1	0.626	0.406	0.284	0.674	0.573	5-4-1	0.625	0.412	0.297	0.633	0.533
4BC02	6-9-1	0.589	0.422	0.299	0.638	0.545	5-5-1	0.591	0.429	0.314	0.597	0.508
4CC03	6-10-1	0.726	0.324	0.202	0.776	0.674	5-6-1	0.722	0.321	0.203	0.719	0.622
4DA10	6-4-1	0.658	0.363	0.241	0.708	0.605	5-6-1	0.655	0.364	0.250	0.657	0.561
4DC03	6-7-1	0.642	0.405	0.283	0.693	0.589	5-4-1	0.638	0.411	0.296	0.642	0.547
4EB01	6-10-1	0.713	0.348	0.227	0.763	0.663	5-6-1	0.705	0.347	0.232	0.691	0.614
9-months lead time												
4AB05	6-5-1	0.611	0.435	0.272	0.665	0.558	5-4-1	0.599	0.442	0.286	0.621	0.521
4AC03	6-12-1	0.582	0.459	0.296	0.632	0.529	5-5-1	0.574	0.468	0.314	0.594	0.495
4AD01	6-5-1	0.598	0.449	0.286	0.654	0.545	5-4-1	0.587	0.457	0.302	0.615	0.509
4BC02	6-9-1	0.563	0.465	0.302	0.612	0.509	5-12-1	0.555	0.475	0.321	0.572	0.477
4CC03	6-10-1	0.694	0.356	0.193	0.758	0.642	5-6-1	0.675	0.368	0.196	0.703	0.595
4DA10	6-10-1	0.629	0.397	0.234	0.685	0.575	5-8-1	0.615	0.399	0.246	0.637	0.536
4DC03	6-7-1	0.615	0.446	0.284	0.669	0.563	5-3-1	0.602	0.453	0.301	0.621	0.525
4EB01	6-10-1	0.681	0.382	0.219	0.737	0.628	5-6-1	0.663	0.382	0.226	0.667	0.584

[a]The station IDs defined as 4AD01, 4AC03, 4AD01, 4BC02, 4CC03, 4DA10, 4DC03 and 4EB01refer to Amboni,Sagana, Gura,Tana sagana, Yatta furrow, Nyamindi, Rupingazi and Kamburu gauge stations respectively

[b]The short form of 'Architecture'

Table 4. 4: **Best ANNs for long-term drought forecasting of SWSI and SDI**

Station ID[a]	SWSI						SDI					
	ANN Arch[b]	R	RMSE	MAE	**d1**	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
12-months lead time												
4AB05	6-5-1	0.589	0.487	0.356	0.647	0.537	5-4-1	0.566	0.502	0.381	0.603	0.505
4AC03	6-4-1	0.562	0.508	0.376	0.619	0.508	5-9-1	0.538	0.523	0.404	0.581	0.477
4AD01	6-5-1	0.578	0.498	0.367	0.636	0.525	5-4-1	0.553	0.511	0.394	0.599	0.492
4BC02	6-9-1	0.544	0.517	0.385	0.601	0.489	5-4-1	0.518	0.532	0.414	0.562	0.461
4CC03	6-10-1	0.673	0.395	0.267	0.731	0.618	5-6-1	0.640	0.397	0.280	0.679	0.575
4DA10	6-7-1	0.613	0.439	0.317	0.668	0.559	5-5-1	0.587	0.446	0.324	0.621	0.522
4DC03	6-5-1	0.594	0.496	0.364	0.652	0.541	5-7-1	0.567	0.509	0.390	0.605	0.507
4EB01	6-10-1	0.657	0.427	0.295	0.715	0.604	5-6-1	0.627	0.432	0.316	0.648	0.564
18-months lead time												
4AB05	6-5-1	0.483	0.576	0.456	0.534	0.482	5-4-1	0.472	0.599	0.493	0.502	0.455
4AC03	6-9-1	0.465	0.643	0.523	0.516	0.453	5-7-1	0.455	0.673	0.593	0.488	0.431
4AD01	6-5-1	0.475	0.612	0.492	0.526	0.472	5-4-1	0.464	0.639	0.535	0.501	0.446
4BC02	6-12-1	0.448	0.608	0.486	0.498	0.437	5-8-1	0.439	0.634	0.528	0.470	0.416
4CC03	6-10-1	0.503	0.487	0.365	0.637	0.564	5-6-1	0.490	0.501	0.395	0.594	0.528
4DA10	6-9-1	0.504	0.543	0.421	0.542	0.505	5-5-1	0.491	0.562	0.442	0.508	0.476
4DC03	6-10-1	0.498	0.587	0.463	0.545	0.487	5-5-1	0.485	0.611	0.501	0.509	0.460
4EB01	6-10-1	0.506	0.502	0.398	0.612	0.549	5-6-1	0.493	0.517	0.427	0.555	0.516

[a]The station IDs defined as 4AD01, 4AC03, 4AD01, 4BC02, 4CC03, 4DA10, 4DC03 and 4EB01refer to Amboni,Sagana, Gura,Tana sagana, Yatta furrow, Nyamindi, Rupingazi and Kamburu gauge stations respectively

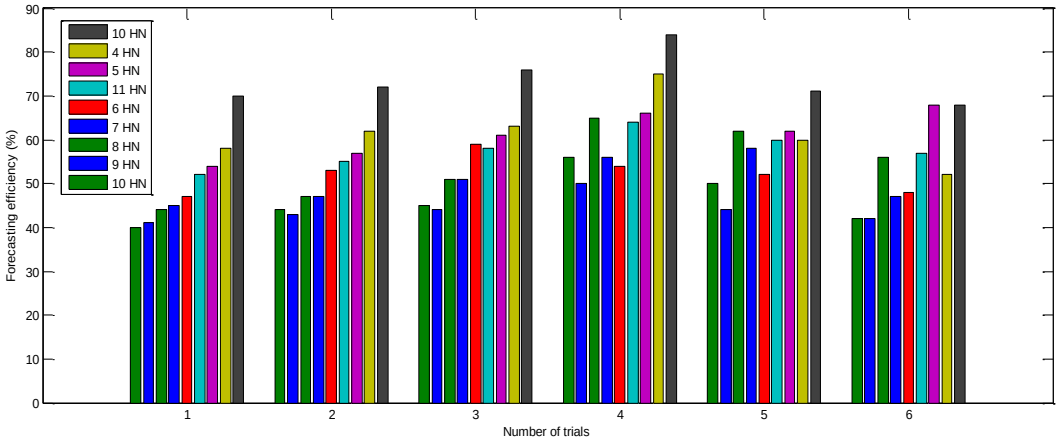
[b]The short form of 'Architecture'

24-months lead time												
4AB05	6-5-1	0.368	0.667	0.547	0.426	0.317	5-4-1	0.348	0.703	0.601	0.368	0.307
4AC03	6-5-1	0.356	0.756	0.636	0.412	0.303	5-7-1	0.337	0.802	0.706	0.395	0.295
4AD01	6-5-1	0.364	0.713	0.586	0.418	0.311	5-4-1	0.345	0.754	0.644	0.403	0.302
4BC02	6-4-1	0.343	0.723	0.595	0.395	0.287	5-3-1	0.327	0.765	0.655	0.377	0.281
4CC03	6-10-1	0.382	0.576	0.458	0.473	0.379	5-6-1	0.359	0.602	0.492	0.478	0.362
4DA10	6-5-1	0.386	0.685	0.541	0.442	0.346	5-4-1	0.363	0.723	0.557	0.418	0.333
4DC03	6-12-1	0.381	0.686	0.561	0.436	0.328	5-9-1	0.359	0.724	0.616	0.411	0.317
4EB01	6-10-1	0.388	0.605	0.492	0.456	0.364	5-6-1	0.365	0.637	0.544	0.442	0.349

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

Figure 4.47 describes the best ANN forecasting performance efficiency with different number of hidden neurons (HN) for SWSI for 1-month lead time at hydrometric station 4CC03.



HN=Number of hidden neuron for inatance 10 HN means ten hidden neurons in the ANN archtecture

Figure 4. 47: Forecasting efficiency verses trials at Yatta furrow gauge station

In addition, Figure 4.47 shows that the highest forecasting efficiency is achieved when the number of Hidden Neurons (HN) in the hidden layer is 10 and which was realised in the fourth trial. However, beyond the fourth trial, the efficiency declined with increase in the number of trials.

The results presented in Figure 4.48 show that the optimum ANN for the Yatta furrow (ID 4CC03) gauging station gave correlation coefficient R values of 0.878, 0.858 and 0.864 for training, validation and testing respectively. The correlation coefficient values were used as a measure of forecasting ability of the model. The results from Figure 4.48 were for a one-month lead time forecasting. This means that 87.8% of the output value of SWSI is explained by the input variables. The dashed line in Figure 4.48b is a perfect fit curve where outputs and targets are equal to each other, while the coloured solid line represents the best fit between outputs (forecasted) and targets (observed) values. From the validation results in Figure 4.48b, the close correlation value of 0.85789 is realized. This means that the model can be used to make accurate predictions for the given conditions.

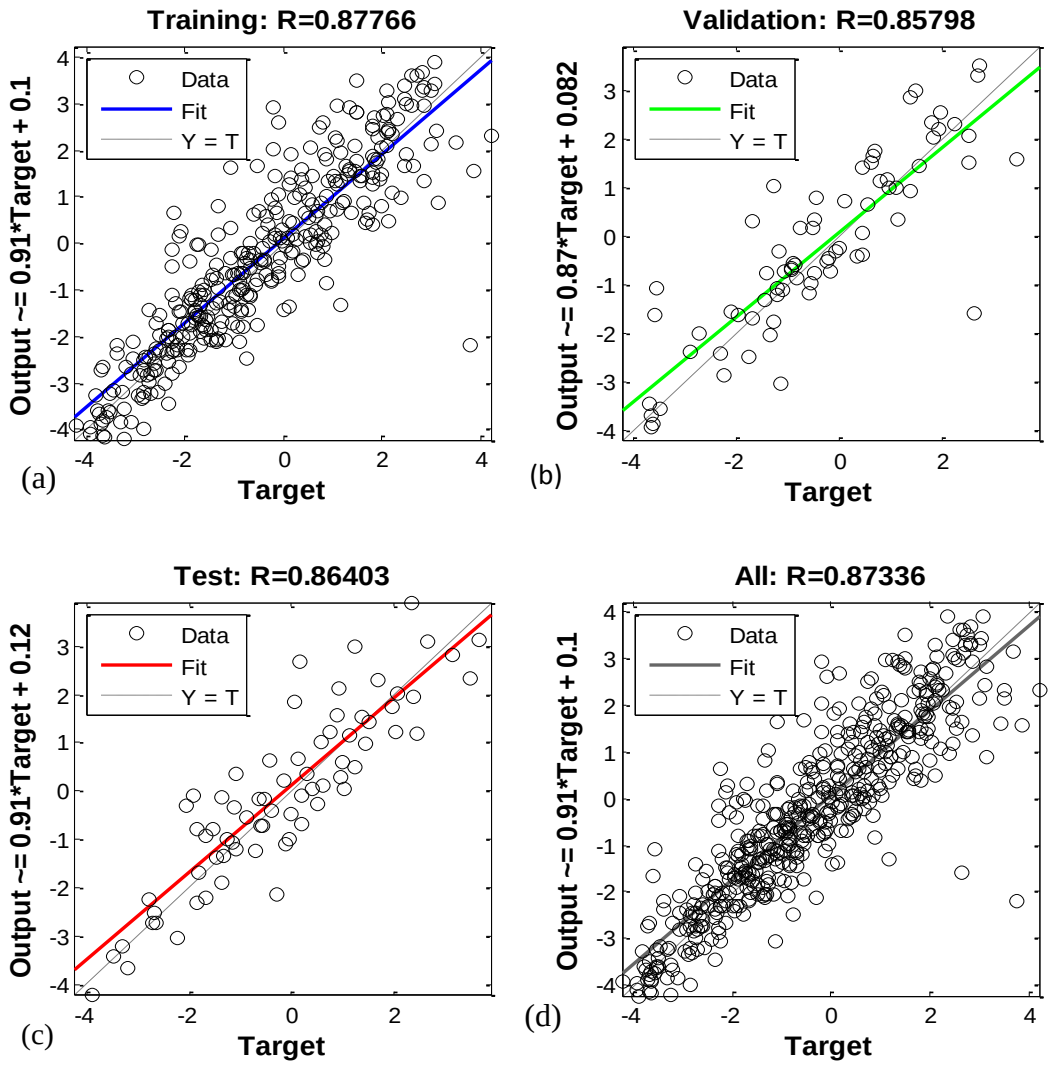


Figure 4. 48: Regression of the best ANNs of SWSI at Yatta furrow gauging station

Figure 4.48 depicts the training, validation and testing mean square error values for Levenberg-Marquardt algorithm with 6-10-1 model architecture. In this architecture, the minimum MSE values for the best model for training, validation and testing are 0.875, 0.926 and 0.863 respectively (Figure 4.49) at epoch 4.

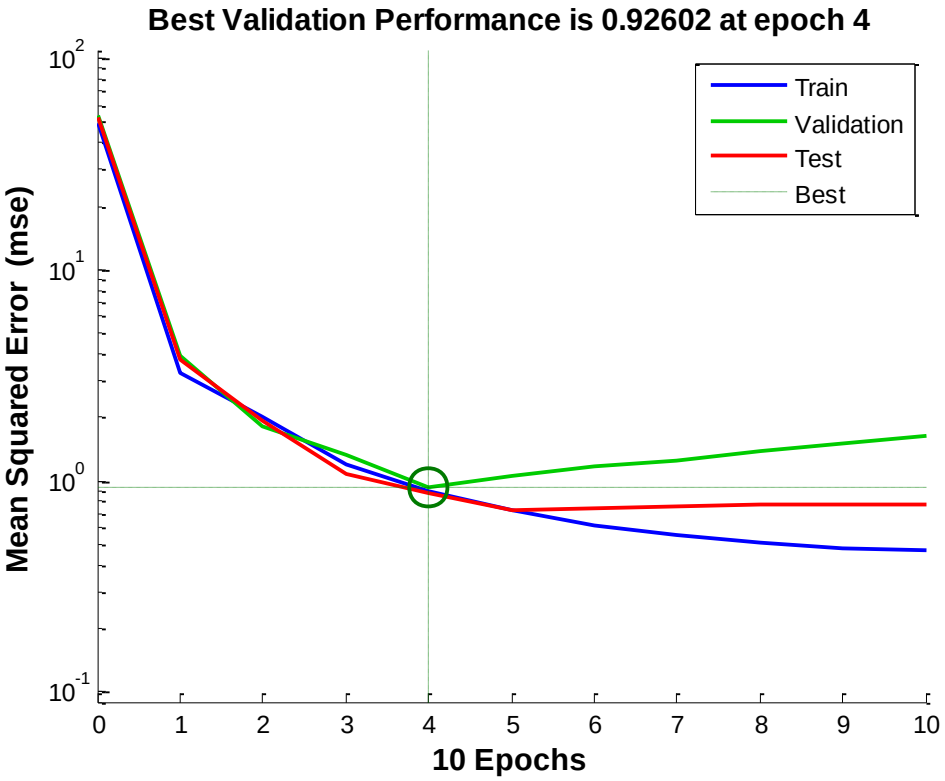
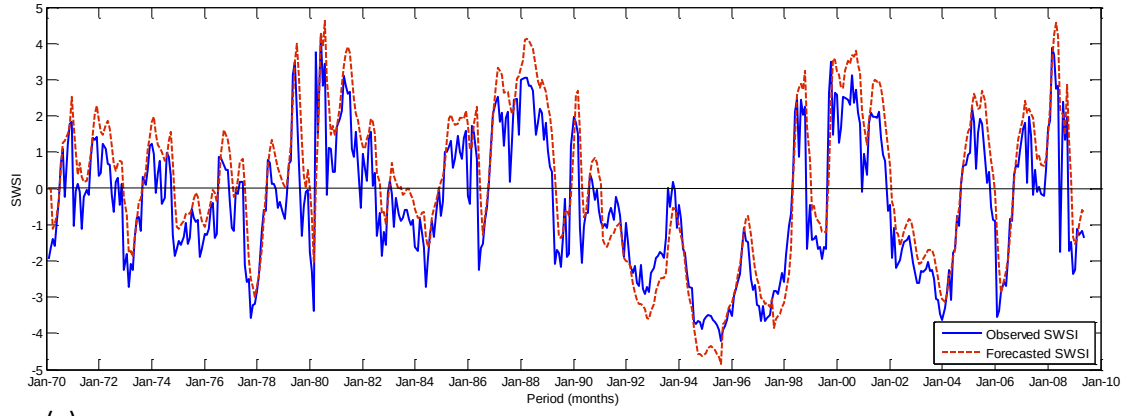
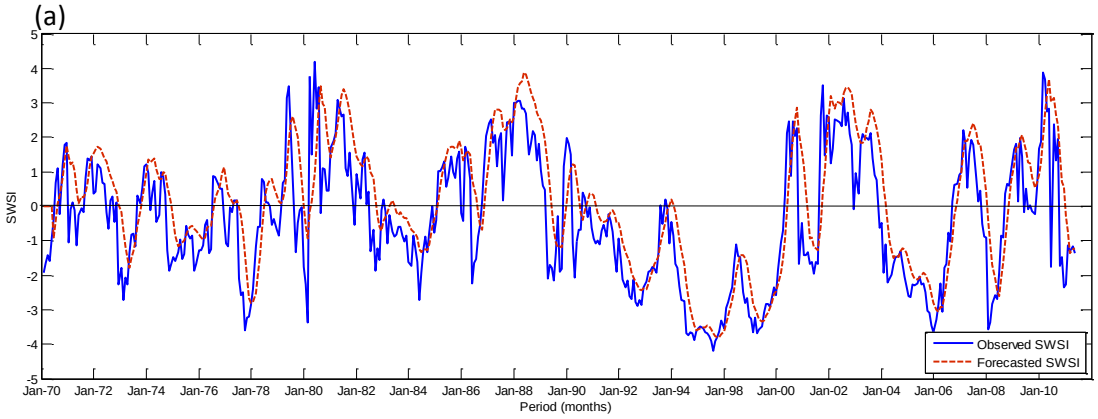


Figure 4. 49: MSE results for L-M algorithm at Yatta furrow station

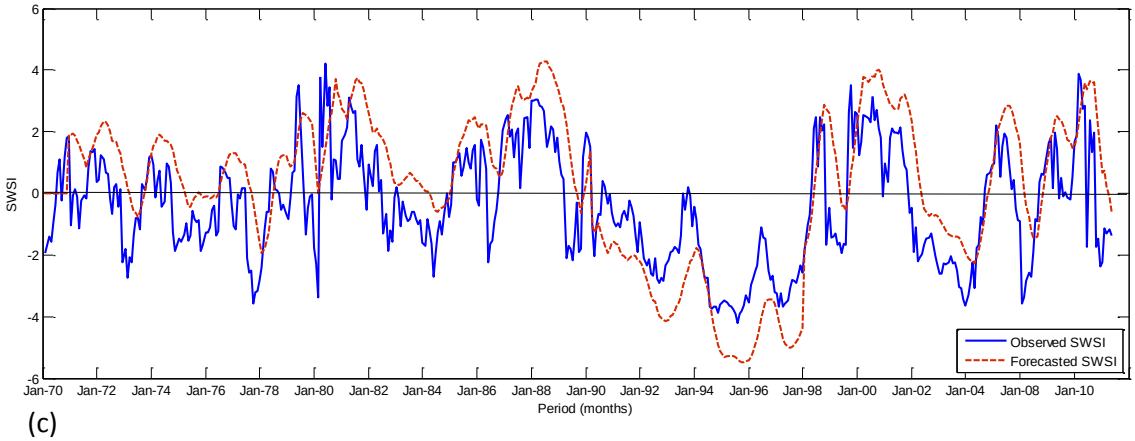
Figure 4.50 shows the results of observed time series of the SWSI values against the finest forecasted values with lead times of 3, 6, 12 and 24 months respectively at gauging station 4AB05. The coefficient of correlation R values for these lead times are 0.662, 0.632, 0.589 and 0.368 rspectively (Tables 4.2 to 4.4). The results show that the forecasted time series SWSI values accurately reflect that of the observed SWSI values at short and medium-term forecasting. The forecasting accuracy of the model declines with increase in lead time. The results of the other gauging stations follow a similar trend but indicate different performance levels of which the gauging station 4AB05 is used to illustrate the variation of time series data. As far as a practical application is concerned, the accurate forecasting can be used to inform water resources managers, agricultural systems and hydropower generation of the expected severities of certain droughts. Such information is useful for timely formulation of mitigation and/or coping mechanisms.



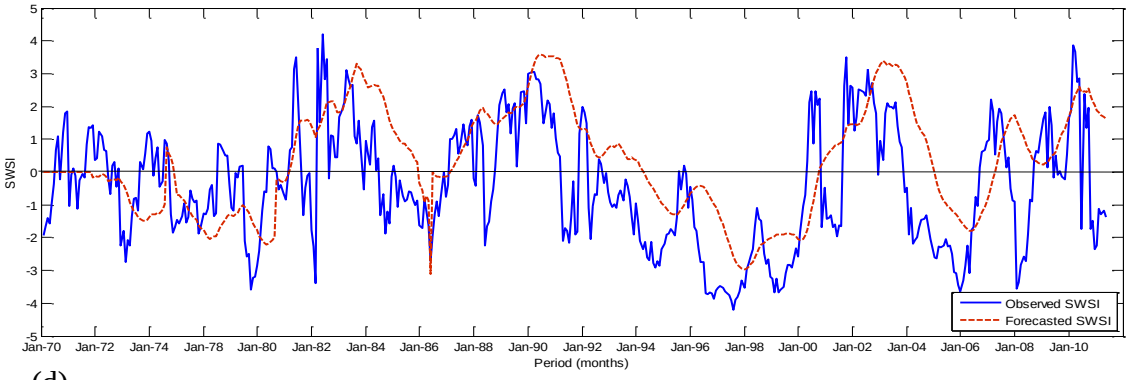
(a)



(b)



(c)



(d)

Figure 4. 50: Observed SWSI and best ANNs forecasts at Amboni gauge station

For the purpose of comparing the performance of the two hydrological drought indices at low and high elevations, Figures 4.51 and 4.52 were plotted. The plots represent the performance of forecasting SWSI and SDI for the Amnoni (4AB05) and Yatta (4CC03) gauging stations and which represent high and low elevations of the basin.

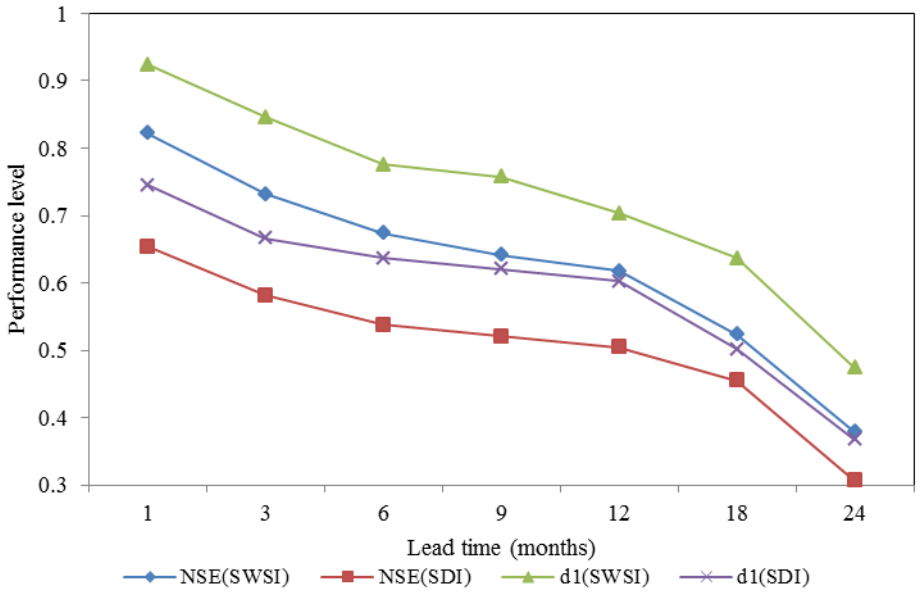


Figure 4. 51: Comparison of SWSI and SDI forecasts at Amboni gauge station

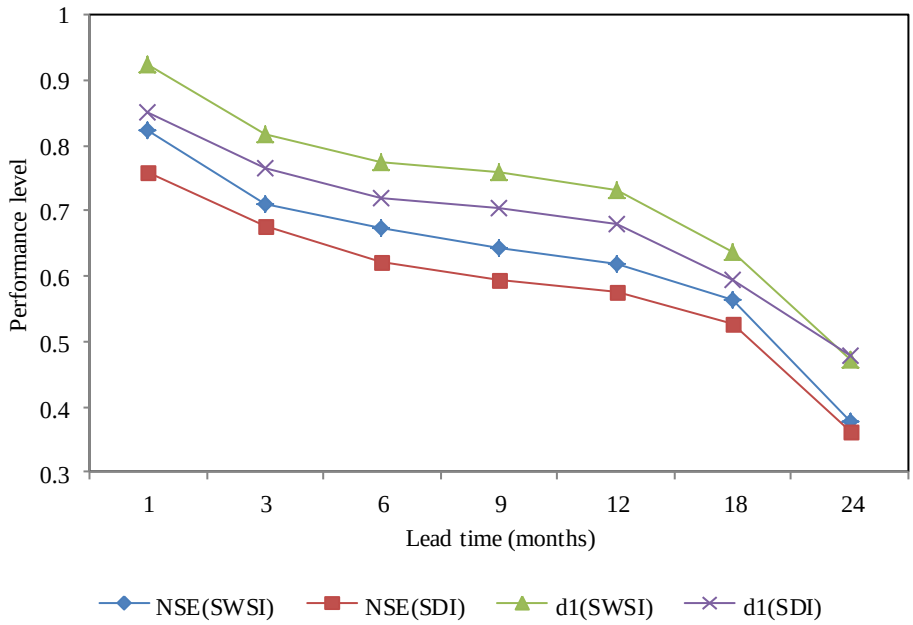


Figure 4. 52: Comparison of SWSI and SDI forecasts at Yatta furrow gauge station

From Figures 4.51 and 4.52, the NSE and modified index of agreement d_1 values in forecasting of the SWSI are higher than the the ones for SDI. These forecasting performance measures at the validation step are therefore better for the SWSI compared to SDI. This is attributed to the fact that stream flow is the only input to the SDI while SWSI has four input variables which when integrated increases the accuracy of the forecasting. The results show that forecasting performance declines with increase in lead time and reduces more rapidly at gauging station 4AB05 than at station 4CC03. This means that it is possible to forecast drought at 4CC03 more accurately at longer lead times than at 4AB05. These stations are located in Arid and Semi-arid Lands (ASALs), and humid areas respectively.

4.2.2 Meteorological drought forecasts

The best network models for meteorological drought forecasts are illustrated in Table 10A in Appendix A. The SPI and EDI network models formulated in this study for forecasting SPI and EDI were found to be in conformity with those developed by Morid *et al*. (2007) in a similar study. Different ANNs architecture were tested for the meteorological drought indices. For the meteorological drought forcasting, the SPI and EDI were used to formulate the network models. Although some of the ANNs artchitecture for SPI such as 5-6-1 for meteorological staion 9037212 are consistent with Morid *et al*. (2007), the other best ANNs exhibit different and unique arcthitecture as shown in Tables 4.5 to 4.7.

Table 4. 5: B**est ANNs for short-term drought forecasting of SPI and EDI**

Station ID[a]	SPI						EDI					
	ANN Arch[b]	R	RMSE	MAE	d1	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
1-months lead time												
9037112	5-4-1	0.865	0.326	0.288	0.791	0.739	5-8-1	0.887	0.299	0.251	0.841	0.765
9037096	5-11-1	0.832	0.296	0.268	0.761	0.711	5-4-1	0.849	0.266	0.235	0.809	0.735
9037202	5-9-1	0.761	0.284	0.248	0.728	0.681	5-8-1	0.764	0.255	0.217	0.774	0.704
9037031	5-7-1	0.768	0.276	0.241	0.735	0.687	5-4-1	0.774	0.248	0.211	0.782	0.711
9036288	5-4-1	0.726	0.269	0.235	0.695	0.649	5-5-1	0.747	0.241	0.205	0.739	0.672
9137123	5-9-1	0.732	0.274	0.24	0.677	0.633	5-5-1	0.755	0.246	0.199	0.72	0.654
9037212	5-6-1	0.703	0.255	0.223	0.653	0.608	5-11-1	0.718	0.229	0.181	0.691	0.629
9037064	5-3-1	0.684	0.23	0.203	0.634	0.591	5-10-1	0.694	0.226	0.167	0.673	0.612
3-months lead time												
9037112	5-6-1	0.826	0.335	0.299	0.774	0.713	5-10-1	0.838	0.326	0.261	0.809	0.733
9037096	5-4-1	0.764	0.313	0.288	0.754	0.686	5-5-1	0.813	0.306	0.251	0.779	0.725
9037202	5-4-1	0.733	0.302	0.264	0.693	0.657	5-4-1	0.729	0.294	0.231	0.715	0.665
9037031	5-8-1	0.724	0.292	0.258	0.707	0.663	5-8-1	0.737	0.286	0.226	0.742	0.681
9036288	5-5-1	0.695	0.284	0.248	0.682	0.64	5-5-1	0.706	0.278	0.217	0.707	0.646
9137123	5-6-1	0.704	0.29	0.254	0.664	0.623	5-8-1	0.695	0.284	0.223	0.698	0.628
9037212	5-6-1	0.666	0.256	0.224	0.638	0.599	5-6-1	0.669	0.251	0.195	0.661	0.603
9037064	5-7-1	0.649	0.233	0.214	0.626	0.582	5-4-1	0.658	0.228	0.187	0.649	0.599

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

Table 4. 6: B**est ANNs for medium-term drought forecasting of SPI and EDI**

Station ID[a]	SPI						EDI					
	ANN Arch[b]	R	RMSE	MAE	**d1**	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
6-months lead time												
9037112	5-6-1	0.767	0.380	0.328	0.735	0.698	5-4-1	0.812	0.362	0.296	0.777	0.723
9037096	5-4-1	0.745	0.348	0.305	0.703	0.672	5-12-1	0.775	0.341	0.266	0.729	0.673
9037202	5-5-1	0.695	0.334	0.292	0.673	0.643	5-7-1	0.697	0.327	0.255	0.698	0.647
9037031	5-10-1	0.681	0.325	0.284	0.679	0.646	5-9-1	0.708	0.318	0.248	0.704	0.653
9036288	5-8-1	0.647	0.315	0.276	0.669	0.639	5-3-1	0.682	0.309	0.241	0.695	0.628
9137123	5-4-1	0.667	0.323	0.283	0.652	0.623	5-7-1	0.678	0.316	0.247	0.676	0.622
9037212	5-6-1	0.626	0.299	0.252	0.614	0.586	5-8-1	0.638	0.293	0.230	0.636	0.578
9037064	5-8-1	0.616	0.282	0.237	0.597	0.571	5-6-1	0.628	0.257	0.197	0.629	0.562
9-months lead time												
9037112	5-8-1	0.715	0.389	0.348	0.692	0.683	5-7-1	0.742	0.391	0.304	0.715	0.689
9037096	5-9-1	0.674	0.367	0.321	0.652	0.643	5-10-1	0.681	0.359	0.283	0.698	0.673
9037202	5-6-1	0.659	0.352	0.308	0.643	0.615	5-7-1	0.667	0.344	0.269	0.668	0.636
9037031	5-5-1	0.656	0.342	0.299	0.649	0.621	5-4-1	0.663	0.335	0.266	0.673	0.648
9036288	5-4-1	0.618	0.332	0.291	0.647	0.619	5-7-1	0.625	0.325	0.254	0.67	0.616
9137123	5-6-1	0.635	0.340	0.298	0.633	0.603	5-6-1	0.641	0.333	0.260	0.653	0.598
9037212	5-7-1	0.589	0.315	0.276	0.614	0.568	5-4-1	0.596	0.308	0.241	0.635	0.575
9037064	5-7-1	0.567	0.277	0.252	0.589	0.559	5-8-1	0.573	0.271	0.221	0.614	0.561

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

Table 4. 7: B**est ANNs for long-term drought forecasting of SPI and EDI**

Station ID[a]	SPI						EDI					
	ANN Arch[b]	R	RMSE	MAE	d1	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
12-months lead time												
9037112	5-5-1	0.681	0.421	0.389	0.676	0.646	5-9-1	0.695	0.432	0.338	0.684	0.669
9037096	5-8-1	0.658	0.407	0.366	0.634	0.606	5-11-1	0.673	0.398	0.334	0.657	0.645
9037202	5-11-1	0.635	0.390	0.349	0.602	0.576	5-7-1	0.648	0.382	0.305	0.624	0.618
9037031	5-71	0.622	0.380	0.333	0.607	0.581	5-5-1	0.635	0.372	0.290	0.629	0.624
9036288	5-9-1	0.588	0.369	0.323	0.607	0.579	5-8-1	0.605	0.361	0.282	0.627	0.591
9137123	5-6-1	0.599	0.377	0.33	0.596	0.567	5-10-1	0.616	0.369	0.288	0.615	0.576
9037212	5-6-1	0.567	0.354	0.316	0.576	0.551	5-4-1	0.578	0.342	0.27	0.597	0.554
9037064	5-7-1	0.548	0.329	0.288	0.517	0.535	5-8-1	0.559	0.322	0.262	0.573	0.534
18-months lead time												
9037112	5-10-1	0.657	0.523	0.479	0.593	0.563	5-3-1	0.675	0.521	0.438	0.615	0.620
9037096	5-7-1	0.625	0.487	0.456	0.577	0.548	5-7-1	0.638	0.476	0.398	0.598	0.596
9037202	5-5-1	0.596	0.467	0.429	0.551	0.523	5-8-1	0.609	0.457	0.374	0.571	0.571
9037031	5-5-1	0.586	0.454	0.397	0.541	0.514	5-9-1	0.598	0.445	0.347	0.563	0.576
9036288	5-4-1	0.556	0.441	0.386	0.528	0.502	5-8-1	0.568	0.432	0.337	0.547	0.544
9137123	5-8-1	0.569	0.451	0.395	0.506	0.481	5-6-1	0.581	0.441	0.345	0.535	0.532
9037212	5-4-1	0.537	0.418	0.366	0.495	0.472	5-7-1	0.548	0.409	0.319	0.515	0.509
9037064	5-11-1	0.516	0.394	0.345	0.474	0.451	5-7-1	0.527	0.386	0.301	0.491	0.504

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

24-months lead time												
9037112	5-8-1	0.606	0.702	0.615	0.514	0.502	5-4-1	0.618	0.688	0.537	0.532	0.584
9037096	5-9-1	0.585	0.680	0.596	0.511	0.481	5-8-1	0.598	0.667	0.520	0.512	0.558
9037202	5-6-1	0.562	0.653	0.572	0.497	0.467	5-9-1	0.573	0.638	0.499	0.515	0.514
9037031	5-11-1	0.548	0.634	0.555	0.486	0.458	5-4-1	0.558	0.621	0.485	0.504	0.539
9036288	5-8-1	0.531	0.616	0.529	0.474	0.446	5-7-1	0.542	0.605	0.462	0.492	0.510
9137123	5-9-1	0.542	0.630	0.552	0.453	0.426	5-6-1	0.554	0.601	0.482	0.469	0.497
9037212	5-4-1	0.509	0.584	0.501	0.454	0.425	5-8-1	0.514	0.572	0.438	0.468	0.477
9037064	5-10-1	0.474	0.542	0.455	0.437	0.411	5-7-1	0.484	0.514	0.419	0.434	0.446

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

To compare the performace of the SPI and EDI in drought forecasting, the NSE and d_1 for the meteorological stations MIAD (7ID 037112) and Nro-moru (ID 9037064) are plotted and results given in Figures 4.53 and 4.54.

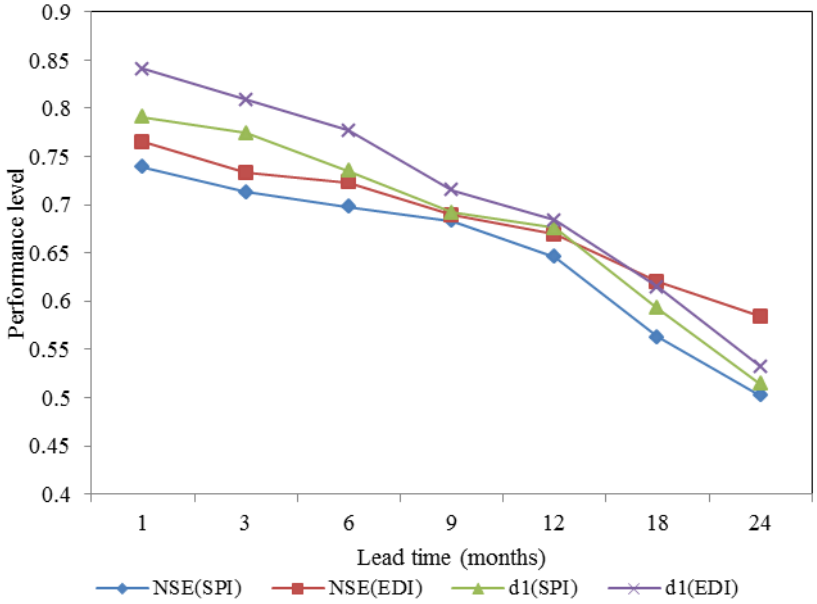


Figure 4. 53: **Comparison of SPI and EDI forecasts at MIAD meteorological station**

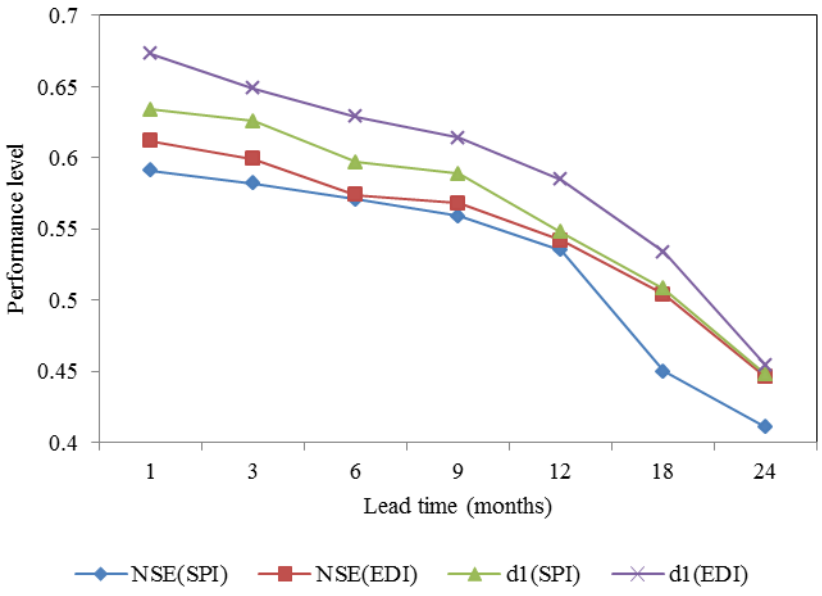


Figure 4. 54: Comparison of SPI and EDI forecasts at Naro-moru meteorological station

The NSE and d_1 for EDI are higher than those of the SPI for both stations. Similar trend is depicted in the other staions. This means that EDI performs better than the SPI in meteorological drought forecasting for the upper Tana River basin.

4.2.3 Agricultural drought forecasts

Agricultural drought forecasting was done using SMDI and the PDSI. The SMDI is a monthly average indiactor of soil moisture stress on land (Tessema, 2007). The results of the best network model for forecasting both the SMDI and PDSI are as indicated in Table 10A of Appendix A. These results indicate that the best network model for forecasting SMDI declined in accuracy as shown by decrease in coefficient of correlation from 0.689 to 0.324 at MIAD meteorological station (ID 9037112) for 1 and 24-months lead time respectively (Tables 4.8 to 4.10). The decline in network forecasting ability for the SMDI with increase in lead time is a trend that closely matches the findings presented by Morid *et al*. (2007) and Belayneh and Adamowski (2013). In their research, the drought forecasting was impelemented using the SPI and EDI.

Table 4. 8: B**est ANNs for short-term drought forecasting of SMDI and PDSI**

Station ID[a]	SMDI						PDSI					
	ANN Arch[b]	R	RMSE	MAE	**d1**	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
					1-months lead time							
9037112	5-4-1	0.689	0.389	0.277	0.659	0.580	5-10-1	0.553	0.472	0.347	0.507	0.464
9037096	5-7-1	0.662	0.394	0.280	0.636	0.552	5-8-1	0.532	0.478	0.351	0.489	0.413
9037202	5-8-1	0.686	0.389	0.273	0.659	0.573	5-5-1	0.551	0.472	0.342	0.507	0.429
9037031	5-7-1	0.627	0.399	0.285	0.604	0.527	5-6-1	0.504	0.484	0.357	0.464	0.395
9036288	5-6-1	0.790	0.307	0.192	0.753	0.673	5-7-1	0.635	0.372	0.241	0.579	0.504
9137123	5-6-1	0.718	0.347	0.229	0.682	0.604	5-9-1	0.577	0.421	0.287	0.524	0.452
9037212	5-7-1	0.697	0.382	0.266	0.664	0.587	5-4-1	0.560	0.463	0.334	0.511	0.440
9037064	5-3-1	0.773	0.322	0.206	0.727	0.661	5-6-1	0.621	0.390	0.258	0.559	0.545
					3-months lead time							
9037112	5-3-1	0.597	0.400	0.297	0.624	0.538	5-8-1	0.574	0.436	0.334	0.487	0.438
9037096	5-6-1	0.581	0.434	0.323	0.575	0.494	5-8-1	0.552	0.442	0.338	0.509	0.455
9037202	5-6-1	0.593	0.409	0.313	0.591	0.512	5-6-1	0.572	0.436	0.329	0.527	0.473
9037031	5-7-1	0.554	0.427	0.327	0.550	0.470	5-10-1	0.522	0.447	0.344	0.483	0.435
9036288	5-8-1	0.684	0.334	0.222	0.675	0.602	5-7-1	0.658	0.344	0.231	0.602	0.555
9137123	5-4-1	0.619	0.385	0.254	0.611	0.540	5-6-1	0.598	0.389	0.276	0.545	0.498
9037212	5-4-1	0.603	0.408	0.305	0.590	0.524	5-10-1	0.581	0.428	0.321	0.531	0.484
9037064	5-7-1	0.667	0.347	0.236	0.644	0.591	5-8-1	0.644	0.361	0.248	0.531	0.499

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

Table 4. 9: B**est ANNs for medium-term drought forecasting of SMDI and PDSI**

Station ID[a]	SMDI						PDSI					
	ANN Arch[b]	R	RMSE	MAE	**d1**	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
					6-months lead time							
9037112	5-5-1	0.595	0.423	0.274	0.575	0.494	5-7-1	0.582	0.429	0.315	0.464	0.399
9037096	5-4-1	0.571	0.434	0.276	0.561	0.460	5-6-1	0.566	0.450	0.343	0.641	0.570
9037202	5-8-1	0.592	0.428	0.271	0.571	0.470	5-10-1	0.577	0.436	0.328	0.664	0.592
9037031	5-2-1	0.543	0.440	0.281	0.539	0.448	5-8-1	0.545	0.454	0.347	0.609	0.545
9036288	5-3-1	0.682	0.338	0.190	0.649	0.548	5-6-1	0.666	0.340	0.224	0.759	0.696
9137123	5-6-1	0.620	0.383	0.226	0.593	0.494	5-10-1	0.604	0.385	0.276	0.687	0.624
9037212	5-7-1	0.601	0.421	0.263	0.580	0.482	5-7-1	0.589	0.435	0.327	0.669	0.607
9037064	5-5-1	0.650	0.354	0.203	0.624	0.541	5-6-1	0.667	0.367	0.256	0.512	0.423
					9-months lead time							
9037112	5-4-1	0.547	0.468	0.313	0.566	0.479	5-8-1	0.680	0.380	0.221	0.453	0.395
9037096	5-7-1	0.524	0.495	0.343	0.542	0.436	5-8-1	0.653	0.385	0.224	0.705	0.681
9037202	5-6-1	0.536	0.484	0.330	0.561	0.448	5-6-1	0.677	0.380	0.218	0.730	0.707
9037031	5-3-1	0.507	0.503	0.351	0.522	0.420	5-10-1	0.618	0.390	0.228	0.669	0.650
9036288	5-8-1	0.616	0.389	0.214	0.641	0.524	5-7-1	0.779	0.300	0.153	0.834	0.830
9137123	5-6-1	0.562	0.422	0.269	0.581	0.472	5-8-1	0.708	0.339	0.183	0.756	0.745
9037212	5-3-1	0.550	0.479	0.329	0.566	0.463	5-8-1	0.687	0.373	0.212	0.736	0.724
9037064	5-5-1	0.605	0.404	0.247	0.586	0.515	5-6-1	0.762	0.314	0.165	0.482	0.41

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

Table 4. 10: B**est ANNs for long-term drought forecasting of SMDI and PDSI**

Station ID[a]	SMDI						PDSI					
	ANN Arch[b]	R	RMSE	MAE	**d1**	NSE	ANN Arch[b]	R	RMSE	MAE	d1	NSE
					12-months lead time							
9037112	5-5-1	0.533	0.537	0.417	0.525	0.447	5-7-1	0.542	0.427	0.310	0.433	0.387
9037096	5-3-1	0.506	0.559	0.442	0.522	0.422	5-8-1	0.521	0.432	0.313	0.461	0.387
9037202	5-8-1	0.520	0.546	0.431	0.545	0.435	5-8-1	0.539	0.427	0.305	0.460	0.376
9037031	5-2-1	0.487	0.569	0.453	0.508	0.408	5-8-1	0.493	0.438	0.319	0.438	0.374
9036288	5-8-1	0.602	0.424	0.306	0.614	0.509	5-8-1	0.621	0.337	0.215	0.523	0.482
9137123	5-6-1	0.552	0.477	0.354	0.562	0.462	5-6-1	0.565	0.381	0.256	0.487	0.441
9037212	5-5-1	0.534	0.544	0.426	0.547	0.449	5-7-1	0.548	0.419	0.297	0.408	0.439
9037064	5-4-1	0.590	0.462	0.346	0.566	0.499	5-8-1	0.608	0.353	0.230	0.474	0.347
					18-months lead time							
9037112	5-3-1	0.435	0.647	0.539	0.445	0.414	5-8-1	0.506	0.509	0.366	0.388	0.361
9037096	5-4-1	0.486	0.419	0.370	0.559	0.486	5-8-1	0.727	0.516	0.648	0.432	0.377
9037202	5-8-1	0.504	0.428	0.361	0.580	0.504	5-8-1	0.690	0.509	0.585	0.444	0.391
9037031	5-8-1	0.461	0.405	0.377	0.531	0.464	5-6-1	0.685	0.523	0.577	0.416	0.364
9036288	5-7-1	0.581	0.452	0.254	0.662	0.592	5-6-1	0.541	0.402	0.432	0.526	0.462
9137123	5-6-1	0.528	0.453	0.303	0.600	0.532	5-7-1	0.607	0.454	0.483	0.450	0.417
9037212	5-4-1	0.512	0.447	0.352	0.584	0.517	5-8-1	0.660	0.500	0.548	0.451	0.403
9037064	5-3-1	0.454	0.559	0.467	0.492	0.452	5-10-1	0.568	0.422	0.272	0.434	0.321

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

24-months lead time												
9037112	5-4-1	0.324	0.763	0.657	0.357	0.322	5-8-1	0.430	0.583	0.449	0.316	0.284
9037096	5-7-1	0.314	0.591	0.454	0.445	0.374	5-8-1	0.413	0.870	0.762	0.352	0.262
9037202	5-4-1	0.321	0.583	0.443	0.460	0.391	5-8-1	0.428	0.818	0.624	0.358	0.267
9037031	5-9-1	0.305	0.598	0.462	0.423	0.358	5-8-1	0.391	0.830	0.716	0.336	0.249
9036288	5-7-1	0.334	0.460	0.311	0.526	0.459	5-6-1	0.493	0.653	0.538	0.424	0.323
9137123	5-6-1	0.338	0.520	0.371	0.477	0.411	5-10-1	0.448	0.754	0.609	0.372	0.295
9037212	5-7-1	0.334	0.573	0.431	0.464	0.399	5-7-1	0.435	0.765	0.664	0.365	0.282
9037064	5-2-1	0.340	0.691	0.595	0.392	0.359	5-9-1	0.482	0.483	0.334	0.326	0.276

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

The performance of SMDI and the PDSI in forecasting drought at diffferent lead times was illustrated by ploting the performance level verses lead time for NSE and the index of agreement (d_1) as shown in Figures 4.55 and 4.56.

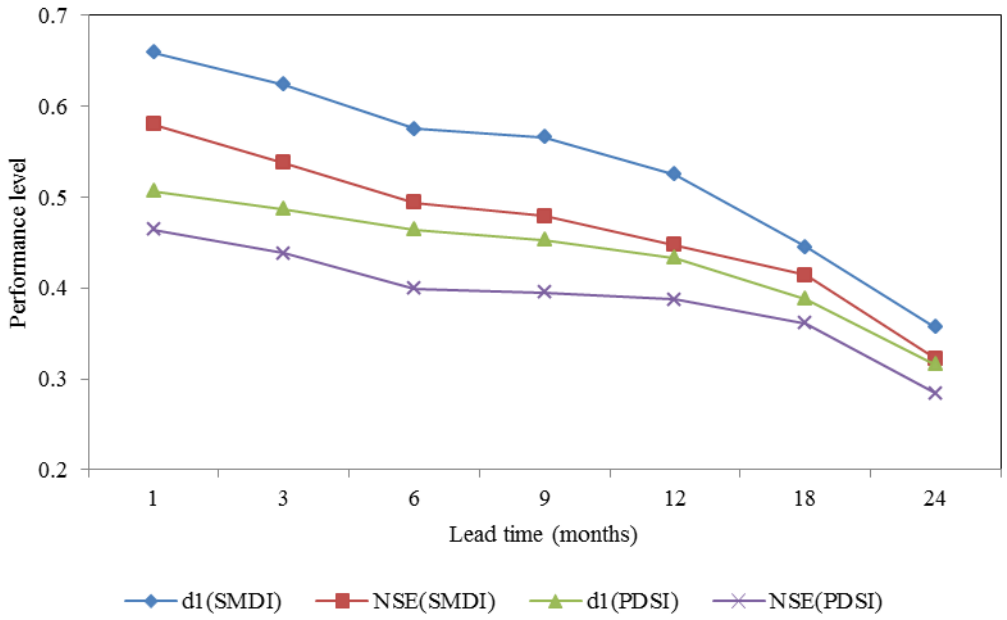


Figure 4. 55: Comparison of SMDI and PDSI forecasts for drought at MIAD station

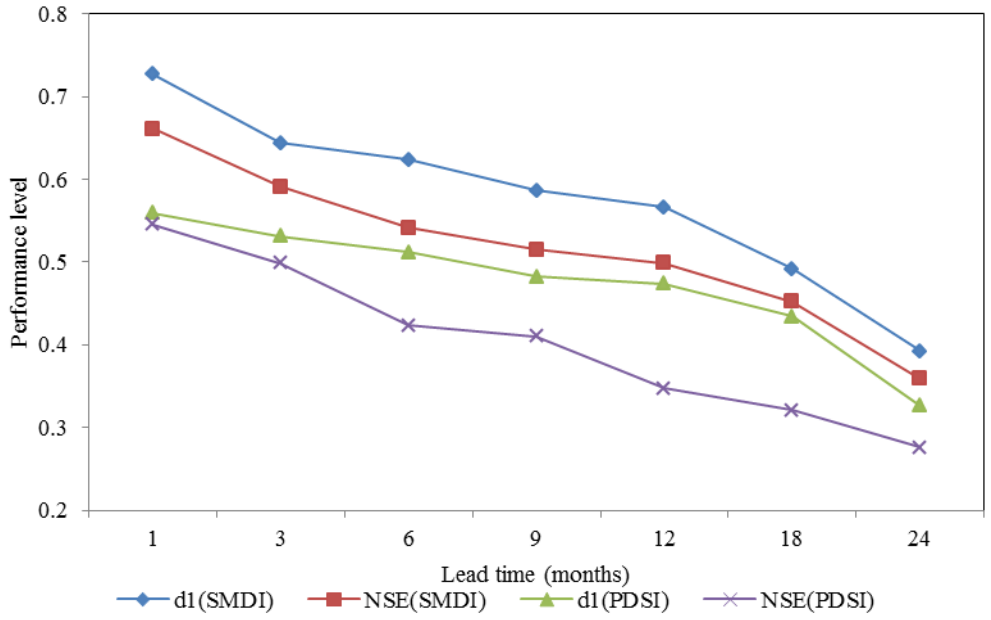


Figure 4. 56: Comparison of SMDI and PDSI forecasts for drought at Naro-moru station

4.3 Formulated NDI for the upper Tana River basin

The (41x4) matrix of observations for the month of January, which was designated as H_{Jan} is as shown in Figure 26B (Appendix B). The matrix H_{Jan} was used in the principal component analysis module of MATLAB to generate a (41x1) matrix which is defined as Q_{Jan} of transformed variables, and a (4x1) matrix of eigenvectors called E_{Jan}. The (4x1) matrix was linked to the first Principal Component (PC1) used in this study. The two matrices Q_{Jan} and E_{Jan} computed using Principal Component Analysis (PCA) are in agreement with the matrix generated by Keyantash and Dracup (2004) and Barua (2010) in development of Agrregated Drought Index (ADI), and Non-linear Aggregated Drought Index (NADI) respectively. These indices are as presented in Figure 27B (Appendix B). The (41x1) matrix of PC1 designated as Y_{Jan} was computed for the month of January based on Equation 3.50. The results are as shown in Figure 28B (Appendix B). The NDI values for the month of Janaury for each of the 41 years were calculated using Equation 4.4. The NDI results are shown in Figure 28B where the standard deviation of PC1 for January for all the years is 2.857 and the resuting NDI is presented in Table 4.11. The NDI values for all the other months were derived using the same procedure and then the data for all months for different years organized in a chronological order and plotted.

$$NDI = \frac{Y_{Jan}}{\sigma_{Jan}} \tag{4.4}$$

For the month of January the standard deviation was determined as 2.857 and thus Equation 4.4 reduces to:

$$NDI = \frac{Y_{Jan}}{2.857} \tag{4.5}$$

Based on Equation 4.5, the NDI for January across the study period was computed and the monthly NDI time series

Table 4. 11: Computed NDI for January at Sagana FCF hydrometric station

Year (January)	NDI
1970	-0.5964
1971	-0.5785
1972	-0.1773
1973	-0.3406
1974	-0.4022
1975	-0.1293
1976	0.0453
1977	0.0095
1978	0.9991
1979	-0.3368
1980	-2.2193
1981	-1.2231
1982	1.9215
1983	-0.2722
1984	-0.2836
1985	0.4942
1986	-0.5588
1987	1.7361
1988	-0.6682
1989	-1.9123
1990	2.6676
1991	0.1884
1992	-0.4358
1993	0.5454
1994	0.5912
1995	-0.0134
1996	0.6487
1997	1.1426
1998	1.3675
1999	-0.0562
2000	-1.1461
2001	-0.9235
2002	-0.9679
2003	0.1722
2004	0.5311
2005	0.1092
2006	0.3776
2007	0.7366
2008	1.2192
2009	-0.8266
2010	-1.4348

4.3.1 Sensitivity of NDI to the input parameters

To test the sensitivity of the NDI to its inputs, the absolute sensitivity was computed based on percent increase and decrease of the variables. Results of the sensitivity analysis are as shown in Figure 4.60.

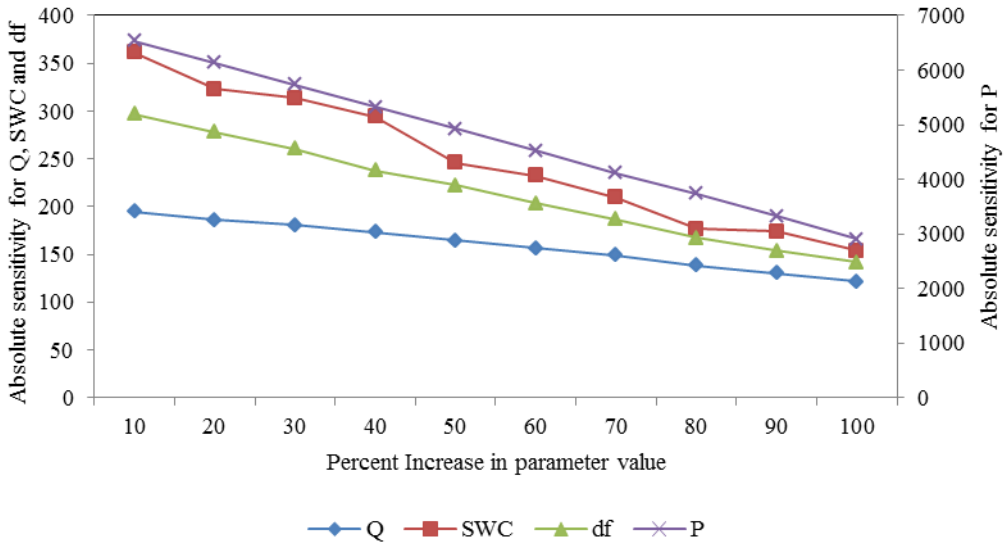


Figure 4. 57: Absoulute sensitivity for NDI at MIAD meteorological station

Results of the sensitivity analysis show that the NDI is most sensitive to changes in precipitation (P), followed by soil water content (SWC), dam levels (df) and Stream flow (Q) respectively. Generally, the sensitivity of NDI declined with percent increase in input variables (Figure 4.60). However, its sensitivity increased with decrease in input variables as shown in (Figure 10B; Appendix B).

4.4 Forecasts of NDI values using ANNs

The results of the best network model for forecasting both the NDI is as shown in the Table 10A of Appendix A. This model was used in forecasting time series NDI at different lead times. As shown in Table 4.11, the value of R decreased from 0.716 to 0.337 at meteorological station 9037112 for 1 and 24 months lead times respectively. The best NDI forecasting for 1 month and 24 months lead times are 0.804 and 0.353 respectively at meteorological station 9037064. From the results, the forecasting ability of NDI model using ANNs decreases with the increase in lead time. Such a trend is consistent with the findings by Morid *et al*. (2007) where it was found that the performance in prediction ability of drought indices decreased with increase in lead time. In their study, for instance, the coefficient of determination (R^2) for forecasting EDI decreased from 0.85 to 0.44 for 1

and 12 months lead time respectively. The same R^2 decreased from 0.78 to 0.42 in forecasting SPI for the same period.

Table 4. 12: Best ANNs for different months lead time forecasts of NDI

Station ID[a]	ANN Arch[b]	R	RMSE	MAE	d1	NSE
1 month lead time						
9037112	6-5-1	0.716	0.348	0.244	0.705	0.631
9037096	6-3-1	0.688	0.351	0.247	0.680	0.601
9037202	6-4-1	0.713	0.348	0.241	0.705	0.623
9037031	6-5-1	0.652	0.356	0.252	0.646	0.573
9036288	6-6-1	0.801	0.274	0.170	0.806	0.732
9137123	6-4-1	0.746	0.309	0.202	0.729	0.657
9037212	6-7-1	0.724	0.341	0.235	0.710	0.639
9037064	6-4-1	0.804	0.288	0.182	0.778	0.719
3 months lead time						
9037112	6-5-1	0.620	0.357	0.262	0.630	0.564
9037096	6-3-1	0.604	0.388	0.285	0.616	0.538
9037202	6-4-1	0.616	0.365	0.276	0.632	0.557
9037031	6-6-1	0.575	0.381	0.289	0.588	0.512
9036288	6-5-1	0.711	0.298	0.196	0.723	0.655
9137123	6-4-1	0.643	0.344	0.224	0.654	0.587
9037212	6-7-1	0.627	0.365	0.269	0.631	0.571
9037064	6-4-1	0.694	0.309	0.209	0.690	0.643
6 months lead time						
9037112	6-5-1	0.605	0.379	0.278	0.615	0.516
9037096	6-3-1	0.588	0.401	0.302	0.601	0.500
9037202	6-4-1	0.599	0.389	0.290	0.611	0.511
9037031	6-4-1	0.567	0.405	0.306	0.577	0.487
9036288	6-5-1	0.692	0.303	0.198	0.694	0.596
9137123	6-4-1	0.628	0.344	0.244	0.635	0.538
9037212	6-7-1	0.612	0.388	0.289	0.620	0.524
9037064	6-4-1	0.676	0.328	0.226	0.667	0.589
9 months lead time						
9037112	6-5-1	0.568	0.418	0.276	0.606	0.500
9037096	6-3-1	0.545	0.442	0.303	0.580	0.475
9037202	6-4-1	0.557	0.432	0.292	0.600	0.488
9037031	6-5-1	0.527	0.449	0.310	0.558	0.457
9036288	6-6-1	0.640	0.348	0.189	0.686	0.570
9137123	6-4-1	0.583	0.377	0.238	0.622	0.514
9037212	6-7-1	0.571	0.428	0.291	0.606	0.503
9037064	6-4-1	0.629	0.361	0.218	0.651	0.560
12 months lead time						
9037112	6-5-1	0.553	0.479	0.368	0.583	0.486
9037096	6-3-1	0.526	0.499	0.390	0.562	0.459
9037202	6-4-1	0.541	0.488	0.380	0.580	0.474

9037031	6-6-1	0.506	0.508	0.400	0.544	0.444
9036288	6-5-1	0.626	0.379	0.270	0.657	0.554
9137123	6-6-1	0.574	0.426	0.313	0.601	0.503
9037212	6-7-1	0.554	0.486	0.377	0.585	0.488
9037064	6-4-1	0.613	0.412	0.305	0.627	0.543
18 months lead time						
9037112	6-5-1	0.452	0.578	0.476	0.476	0.434
9037096	6-3-1	0.436	0.649	0.573	0.462	0.411
9037202	6-4-1	0.444	0.616	0.517	0.475	0.425
9037031	6-5-1	0.420	0.612	0.510	0.445	0.396
9036288	6-4-1	0.469	0.483	0.381	0.563	0.503
9137123	6-4-1	0.470	0.542	0.427	0.481	0.454
9037212	6-7-1	0.465	0.589	0.484	0.482	0.438
9037064	6-4-1	0.472	0.499	0.412	0.526	0.492
24 months lead time						
9037112	6-5-1	0.337	0.681	0.580	0.349	0.296
9037096	6-3-1	0.326	0.777	0.682	0.375	0.284
9037202	6-4-1	0.334	0.730	0.622	0.383	0.291
9037031	6-7-1	0.316	0.741	0.632	0.358	0.271
9036288	6-5-1	0.347	0.583	0.475	0.454	0.349
9137123	6-4-1	0.351	0.700	0.538	0.397	0.321
9037212	6-7-1	0.347	0.701	0.595	0.390	0.305
9037064	6-4-1	0.353	0.617	0.525	0.420	0.336

[a]The station IDs defined as 9037112, 9037096, 9037202, 9037031, 9036288, 9137123, 9037212 and 9037064 refer to MIAD,Sagana FCF, Embu, Kerugoya DWO, Nyeri, Mangu HS, Maragua GEF and Naro-moru FGP weather stations respectively

[b]The short form of 'Architecture'

The ability of the NDI forecasting was also found to differ with elevation. Generally the performance was better for higher elevation than lower elevations. This is attributed to the high sensitivity of the NDI to precipitation as illustrated by absolute sensitivity analysis (Figure 4.57) and that high precipitation amounts occur in higher elevations. For the purpose of illustrating this, the NSE and d_1 for Mwea Irrigation and Agricultural Centre (MIAD) and Naro-moru meteorological stations were plotted. From the results given in Figures 4.61 and 4.62, it can be seen that the NDI can forecast the drought accurately (NSE<0.5) upto 9 and 15-months lead time at MIAD (ID 9037112) and Naro-moru F.G.P. (ID 9037064) stations respectively.

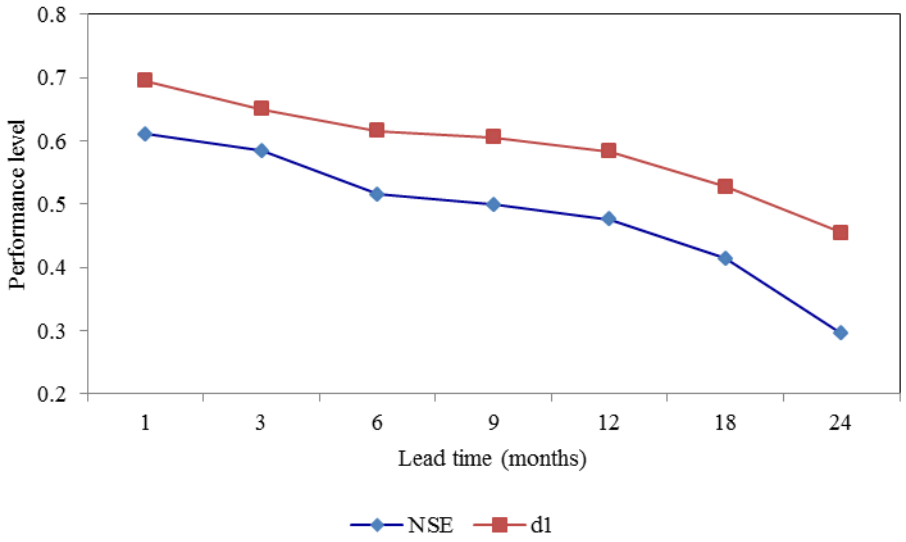


Figure 4. 58: Performance of NDI forecasts at MIAD meteorological stations

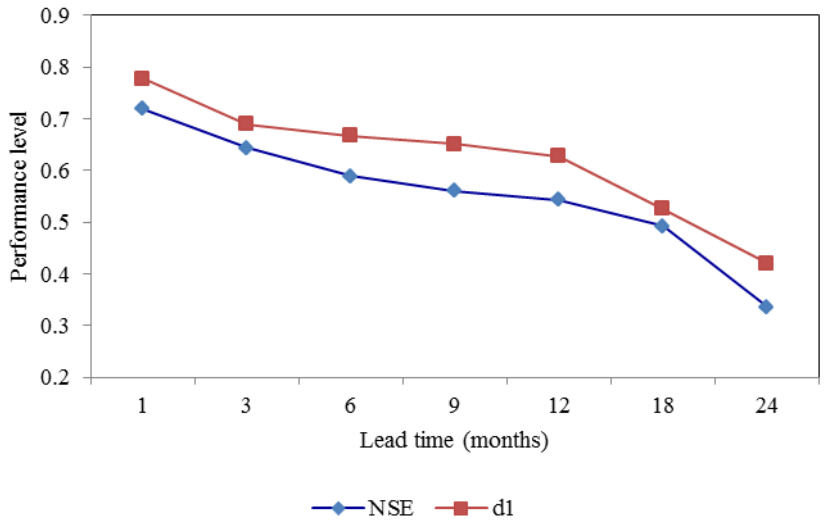
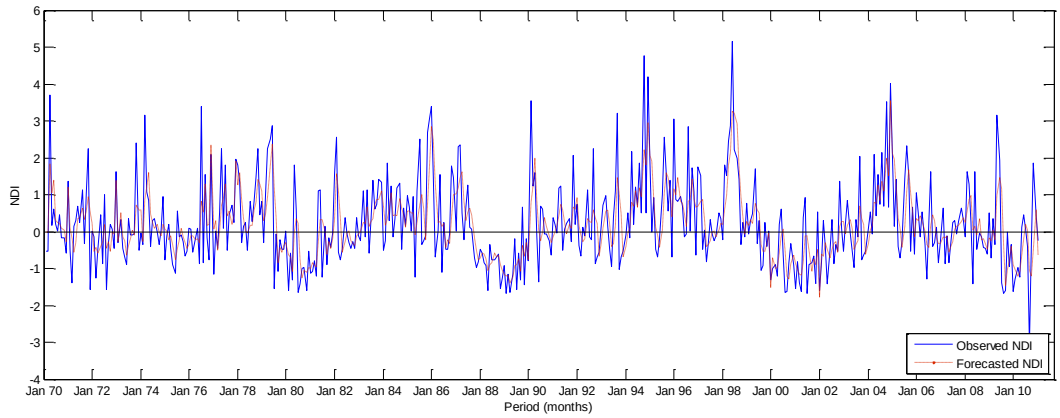
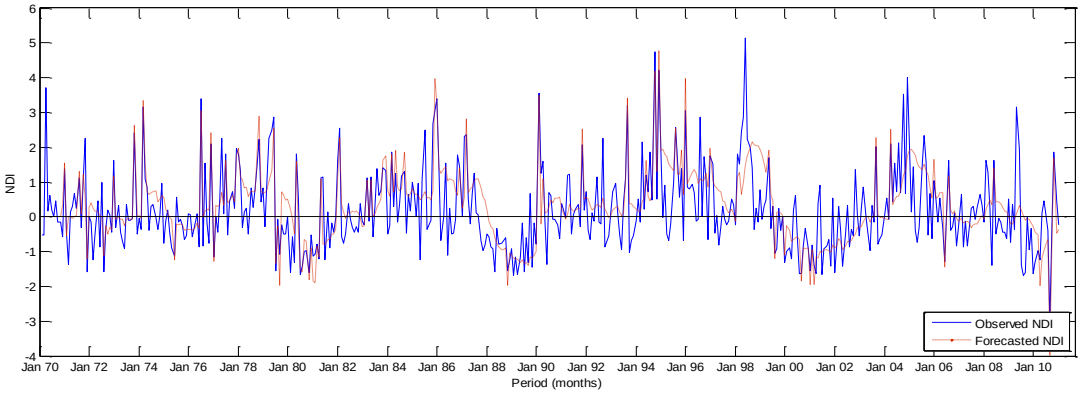


Figure 4. 59: Performance of NDI forecasts at Naro-moru meteorological stations

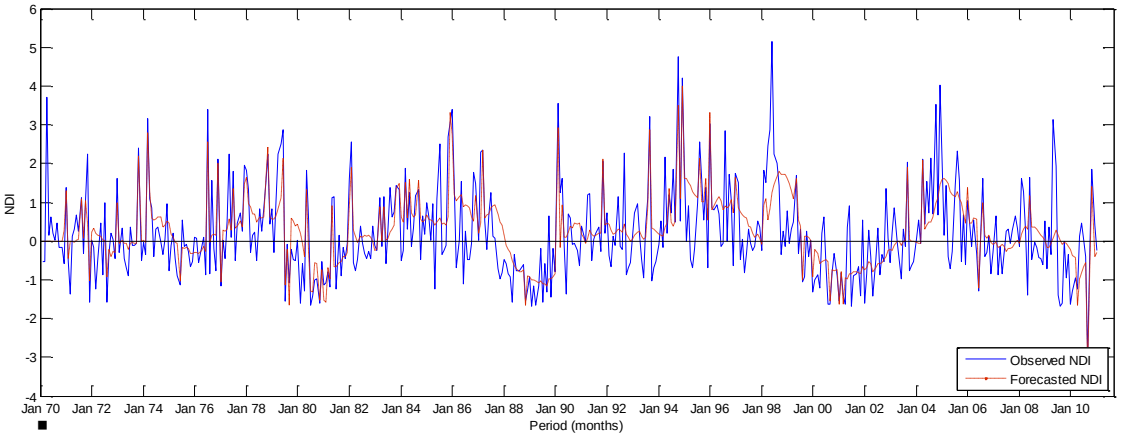
To illustrate the pattern of NDI forecasting, the time series plots of both curves of the observed and forecasted values for the best ANN models is presented in Figure 4.63. The forecast results are for 3, 12 and 24 months lead time respectively. The observed NDI time series curves are for Sgana F. C. F and Naro- moru hydrometric stations are presented in Figures 17B and 18B, Appendix B.



(a)



(b)



(c)

Figure 4. 60(a-c): Observed NDI and best ANNs forecast results at Sagana FCF station

From the time series plots at the at Sagana FCF metorological station (ID 9037096), and using the NDI drought of magnitude -1.7, -1.8, -1.8, -1.5, -2.0, and -2.2 for the years 1973, 1981, 1983, 1992, 2000 and 2010 was detected. According to the NDI crietrion (Table 3.12) the values -1.7, -1.8, -2.0 and -2.2 represent severe drought while -1.5 represent moderate drought. The occurrence of the

drought episodes for these years is consistent with the historical droughts in the area as given by UNDP (2012) report (Table 4.13). The NDI values for Kamburu hydrometric station (ID 4ED01) were used to detect and classify drought for the basin. It was found that the droughts years were effectively detected and severity of drought classified as shown in Table 4.13.

Table 4. 13: Classification of the reported drought at Kamburu hydrometric station

Reported drought (Year)	Remarks on the reported drought	Classification of drought based on the developed NDI
1972	Water shortage livestock deaths	Extreme drought
1973-1974	Human and livestock deaths	Severe drought
1974-1976	Heavy livestock losses, food and water shortages	Severe drought
1980	Low crop production, water shortages	Severe drought
1981	Famine and water shortage	Severe drought
1983	Water shortages, human and livestock migration	Severe drought
1984	Huge food shortages	Extreme drought
1987	Severe food shortages	Severe drought
1992-1994	Moderate food shortages	Extreme drought
1999-2000	Deficit food supply, interuption of electricity supply, water scarcity	Severe drought
2010	Food and water deficit	Severe drought

4.4.1 Drought projections based on NDI and RMSNN

The most notable drought episodes expected for the upper Tana river basin is as presented in Figure 4.61 at Kamburu hydrometric station (ID 4ED01). Kamburu station is selecetd because it coincides with the outlet of the basin. Thus an output representative of drought for whole basin; the projected extreme droughts (Values of NDI less than -2.27) include the years 2019 to 2020, 2028 to 2030, 2040 to 2041, 2050, 2053 to 2054, 2060 to 2063, 2070 to 2073 and 2081 to 2083. In addition the projected severe drought (NDI values $\geq 2.27\ to\ -1.64$) events are in the years 2017, 2026, 2033, 2046, 2056 to 2057, 2066 to 2067, 2074 to 2075, 2079 and 2087 to 2088 as shown in Figure 4.61. The drought projection for SMDI and SWSI are as shown in the Figure 19B and 21B Appendix respectively. B.

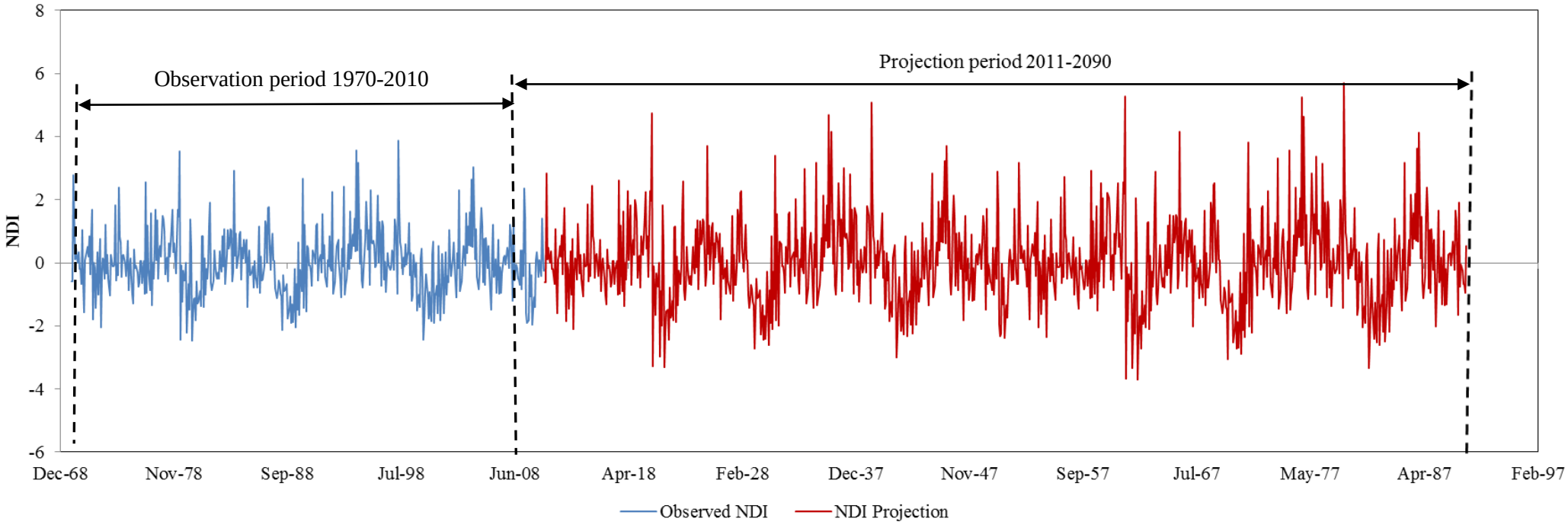
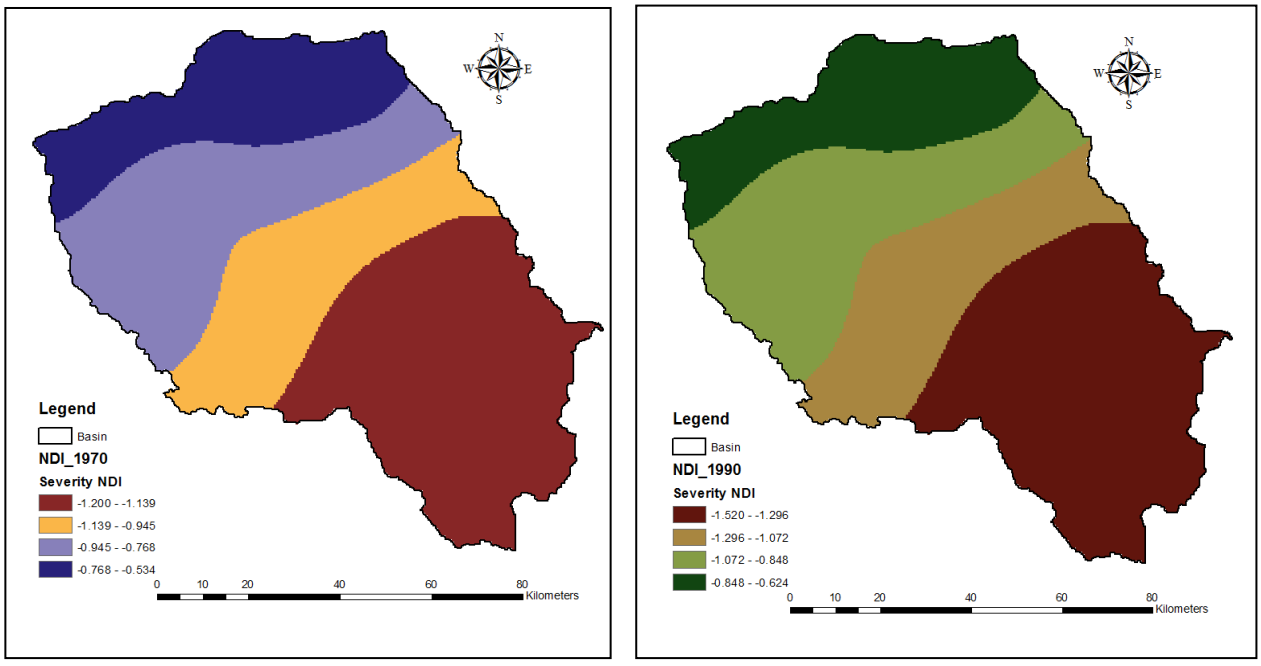


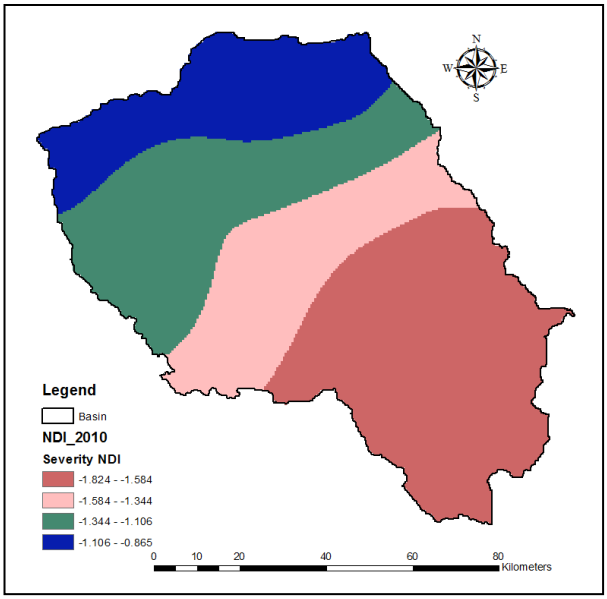
Figure 4. 61: Observed and projected drought at Kamburu hydrometric station

4.4.2 Spatially distributed drought severity based on NDI

Different parts of the basin exhibit different drought severity magnitudes as detected by the NDI. The magnitude of drought severity decreased from the areas located in Low Elevations (LE), Lower Middle Elevation (LME), Middle Elevation (ME) and High Elevation (HE) in that order for different years. The highest severity range -1.200 to 1.390 was for 1970, while the lowest is –0.768 to -0.534 for the Lower and Higher elevations respectively (Figure 4.62). A similar trend is reflected in the spatially distributed severities for the years 1990 and 2010.



(a) (b)



(c)

Figure 4. 62: Spatially distributed drought severity for the years based on NDI

CHAPTER FIVE
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This research assessed and forecasted drought conditions using indices and artificial neural networks. Hydro-meteorological data from sixteen hydrometric stations in the upper Tana River basin was used. The focus was to formulate the most appropriate models for assessment and forecasting of drought using Indices and Artificial Neural Networks (ANNs). In addition drought characteristics that can be used for timely drought mitigation measures and/or coping mechanisms, guide in decision making for water resources planning and management were determined. The specific conclusions drawn from the research are:

i) The temporal variation of drought showed that the DIs were able to effectively detect severe or moderate droughts in the upper Tana River basin during the study period. The findings indicate that the hydrological, meteorological and agricultural droughts presented an extreme drought event with a frequency corresponding to 6, 5 and 4 years respectively for the entire basin. On further analysis using Man-Kendall trend test trend test it was observed that there was an increasing trend of severe drought in south-eastern areas and no trend in north-western and north-eastern parts of the basin. In spatial context, south-eastern areas located in the lower elevations of the basin are more prone to severe drought risks than the areas which are at higher elevations. The least drought risk areas are in the north-western parts of the basin.

ii) From this research, the best models for drought forecasting at 1, 3, 6, 9, 12, 18 and 24-months lead time were developed for each drought type. Based on this study the best drought indices are presented as SWSI, EDI and SMDI for forecasting short, medium and long-term hydrological, meteorological and agricultural droughts respectively.

iii) A Non-linear Integrated Drought Index (NDI) was formulated for the upper Tana River basin. This was based on an aggregation process called Principal Component analysis (PCA) of four hydro-meteorological variables; precipitation, stream flow, soil-moisture and dam levels. The NDI was found to effectively detect droughts in conformity with a similar and previously developed index called NADI. Since the NDI is developed by aggregating input variables that represent different types of droughts; it is thus presented as a single tool which can be adopted for a synchronized assessment and forecasting of all the three

operational drought types (hydrological, metorological and agricultural droughts) in the basin.

iv) Using the formulated NDI, the drought was forecasted at different lead times of 1, 3, 6, 9, 12, 18 and 24. Based on the performance measures (R, RMSE, d1, NSE) of drought forecasting models, it is inferred that the formulated NDI effectively forecasts drought in high and low elevations at a range of 1 to 12, and 1 to 10-months lead time respectively with R value greater than 0.5. Beyond 10 and 12-months lead times, the forecast accuracy significantly decline giving R and NSE values below 0.5 and thus cannot be relied upon.

5.2 Recommendations

From these research findings, more research work can be done:

i) Temporal and spatial drought is influenced by both climate and land use/cover changes. Thus, there is need for further research on relative effect of climate change and land use/cover change on drought based on DIs in the upper Tana River basin. This research used monthly and 90-m temporal and spatial resolutions in assessing drought using the DIs. Further research involving finer resolutions is recomemded to compare the effectiveness of of DIs in drought assessment at finer resolutions.

ii) Further research is required involving a comparative study to compare multiple data driven and remote sensed drought indices in forecasting of drought conditions at different lead-times in the upper Tana River basin. In addition, drought forecasting based on climate projections and scenarios to formulate risk assessment tools should be conducted in the upper Tana River basin.

iii) A further research to formulate a Nonlinear-Integrated Drought Index (NDI) using principle component analysis based on additional basin hydro-meteorological variables such as temperature, evaporation, humidity, sunshine should be conducted.

iv) Further research to calibrate, validate and assess the performance of formulated NDI in drought forecasting in other river basins is required. It is recommended that the formulated DIs and be adopted as decision support tool in managing drought and water resources systems. A network of automatic hydrometric stations that convey real- time data to the algorithms of DIs for managing impending droughts should be incorporated (Figure 11B of Appendix B).

REFERENCES

Abad, M. B. J., Zade, A. H., Rohina, A., Delbalkish, H. and Mohagher, S. S. (2013). The effect of climate change on flow regime in Basher river using two meteorological and hydrological standards, *International Journal of agricultural Sciences*, 5: 2852-2857.

Adindu, R., Igboekwe, M., Isu, J. O. and Igbokwe, K. (2013). Estimation of groundwater recharge at Amudike watershed, Abia State Nigeria, *European Journal of applied engineering and scientific research*, 2(3): 38-43.

Agwata, J. F. (2014b). Spatial characteristics of drought duration and severity in the upper tana basin, Kenya, *International research journal of environmnent sciences*, 3(4): 18-26.

Agwata, J. F., Wamicha, W. N. and Ondieki, C. M. (2014a). Modelling hydrological drought events in the upper Tana river basin of Kenya, *IOSR Journal of mechanical and civil engineering*, 11(1) 41-48.

Agwata, J. F., Wamicha, W. N. and Ondieki, C. M. (2015). Analysis of hydrological drought events in the upper Tana river basin of Kenya, *Journal of environment and earth science* 5(2) 22-31.

Aksoy, H. (2000). Use of Gamma distribution in hydrological analysis, *Turk Journal of Engineering Sciences*, 24: 419-428.

Alaa, B. S. (2014). Water use optimization based on the concept of partial root zone drying, *Ain Shams Engineering Journal*, 5: 55-62.

Al-Mashindani, G. Lal, P. B. B. and Mujda, M. F. (1978). A simple version of Gumbel's method for flood estimation, *Hydrological sciences journal*, 23 (3): 373-379.

Arnell, N. W., (2004). Climate change and global water resources: SRES emissions and socio-economic scenarios, *Global Environmental Change*, 14: 31–52.

Asad, M. A., Ahmeduzzaman, K., Kar, S., Khan, M. A., Rahman, M. N. and Islam, R. (2013). Flood frequency modelling using Gumbel's and Powell's method in Dudhkumar River, 2(2): 25-28.

Asgari, H., Mohsenipour, M., Shahid, S., Hadipour, S., Shafieifar, M. and Roushenas, P. (2014). Spatio-temporal charactersitics of droughts and drought trends in Qazvin province of Iran, *Research Journal of Applied sciences, Engineering and Technology*, 8(11): 1299-1311.

Awass, A. A. (2009). Hydrological drought analysis-occurrence, severity and risks, the case of Wabi Shebele River Basin, Ethiopia PhD thesis Universität Siegen.

Awass, A. A. and Foerch, G. (2006a). Assessing the threshold levels used to detect hydrological and environmental drought in semi-arid areas. In proceedings of man and river systems II December 4-6th Paris, France.

Ayyad, M. A. (2003). Case studies in the conservation of biodiversity: degradation and threats, *Journal of Arid Environments*, 54(1): 165-185.

Bacanli, U. G., Firat, M. and Dikbas, F. (2008). Adaptive neuro-fuzzy inference system for drought forecasting, *Journal of stochastic environmental research and risk assessment*, 23(8): 1143-1154.

Balint, Z, Mutua, F, Muchiri, P. and Omuto, C. T. (2013). Monitoring drought with the combined drought index in Kenya, *Developments in earth surface processes*, 16: 341-356.

Barua, S. (2010). Drought assessment and forecasting using a non-linear aggregated drought index, PhD thesis, Victoria University, Australia.

Bassey, P. C. and Akinkunmi, B. O. (2013). Introducing the spatial qualification problem and its qualitative model, *African journal of computing and ICT*, 6(1): 191-197.

Beale, M. H., Hagan, M. and Demuth, H. B. (2014). Neural network toolbox^{TM}, User's Guide, Mathsworks Inc.

Belayneh, A. and Adamowski, J. (2012). Standard precipitation index drought forecasting using neural networks, wavelet neural networks and support vector regression, *Journal of applied computational intelligence and soft computing*, 2012: 1-13.

Belayneh, A. and Adamowski, J. (2013). Drought forecasting using new machine learning methods, *Journal of Water and Land Development*, 18 (I-IV): 3-12.

Biamah, E. K., Sharma, T. C. and Stroosnijder, L. (2002). Simulation of watershed peak runoff using the Nash model, *Journal of Environment and Agricultural Engineering* (JEAE), 2(1): 49-56.

Birkel, C. (2005). Temporal and spatial variation of drought indices in Costa Rica, MSc thesis, Istitut für Hydrologie, der Albert-Ludwigs-Universität Freiburg i. Br.

Bishop, C. M. (1995). Neural networks for pattern recognition NY, Oxford University press.

Bonacci, O. (1993). Hydrological identification of drought, hydrological processes, 7: 249-262.

Brown, J. F. Wardlow, B. D., Tadesse, T., Hayes, M. J. and Reed, B. C. (2008). The vegetation drought response index (VegDRI). A new integrated approach for monitoring drought stress in vegetation, *geosciences and remote sensing*, 45(1): 16-46.

Bryant, S. and Arnell, N. W. and Law, F. M. (1992). The long-term context for the current hydrological drought, institute of water and environmental management conference on management of scarce water resources October 1992 18^{th}-24^{th}.

Bulu, A. and Aksoy, H. (1998). Low flow and drought studies in turkey, proc. low flows expert meeting 10-12 June 1998, University of Belgrade.

Byurn, H. R. and Wilhite, D. A. (1999). Objective quantification of drought severity and duration, *Journal of climatology*, 12(9): 2747-2756.

Casa, A., Ovando, G., Bressanini, L. and Martinez, J. (2013). Aquacrop model calibration in potato and its use to estimate yield variability under field conditions, *Atmospheric and Climatic sciences*, 3: 397-407.

Cassiamani, C., Morgillo, A., Marchesi, S. and Pavan, V. (2007). Monitoring and forecasting drought on a regional scale: Emilia Romagna region, *Water science and technology*, 62(1): 29-48.

Castano, A. (2012). Monitoring drought at river basin and regional scale: application in Sicily, PhD Dessertation in Hydraulic Engineering, University of Catania, Italy.

Ceppi, A, Ravazzani, G., Corbari, C., Mencci, R., Saleno, S., Meuccr, and Mancini, M. (2014). Real time drought forecasting system for irrigation management, *Journal of Hydrology and Earth Systems Science*, 18: 3353-3366.

Chai, S. S. (2010). An Artificial Neural Network approach for soil moisture retrieval using passive microwave data, PhD thesis, Curtin University of Technology.

Chatfield, C. (2003). The analysis time series, an introduction, Chapman and Hall press, USA.

Chen, S. and Rao, P. (2008). Land degradation monitoring using multi-temporal Landsat TM data in transition zone between grassland and crop land of northeast China, *International Journal of Remote Sensing*, 29(7): 2055-2073.

Chen, W. Xiano, Q. and Sheng, Y. (1994). Application of anormally vegetation index to monitoring heavy drought remote sensing *Journal Environmental Science* 9(2): 106-112.

Chiang, Y. M., Chang, L. C. and Chang, F. J. (2004). Comparison of static feed forward and dynamic feedback neural networks for rainfall runoff modelling, *Journal of Hydrology*, 290(3-4): 297-311.

Clevers, J. G. P. W. (1988). The derivative of simplified reflectance model for the estimation of leaf area index, *Remote Sensing and Environment*, 25(1): 53-69.

Crippen, R. E. (1990). Calculating the vegetation index faster, *Remote Sensing of Environment*, 34(1): 71-73.

Cutore, P., Mauro, G. D. and Cancelliere, A. (2009). Forecasting Palmer index using artificial neural networks and climate indices, *Journal of Hydrologic Engineering*, 14(6): 588-595.

Dai, A. (2011). Drought under global warming: A review, *WIREs climate change global circulation*, 2: 45-65.

Dalezios, N., Loukas, A., Vasiliades, L. and Liakopoulos, E. (2000). Severity-duration-frequency analysis of droughts and wet periods in Greece, *Hydrological sciences journal*, 45(5): 751-769.

Demuth, H.B., Beale, M. H. and Hagan, M. (2009). Neural network toolbox 6, users guide, Nattick M. A, Mathsworks Inc.

Dracup, J. A., Lee, K. S. and Paulson, E. G. (1980a). On the statistical characteristics of drought events, *Journal of water resources research*, 16(2): 289-296.

Droogers P, Kauffman JH, Dijkshoorn JA, Immerzeel W and Huting JRM 2006. Green Water Credits: Basin identification. Green Water Credits Report 1, ISRIC - World Soil Information, Wageningen.

Du, W., He, X., Shamaila, Z., Hu, Z., Zeng, A. and Muller, J. (2011). Yield and biomass prediction testing of AquaCrop model for winter wheat, *Transactions of the Chinese society of agricultural machinery*, 42: 174-178.

Eden, U. (2012). Drought assessment by evapotranspiration mapping in Twente, The Netherlands, MSc Thesis, Enchede, the Netherlands.

El-Jabi, N., Turkkan, N. and Caissie, D. (2013). Regional climate index for floods and droughts using Candian climate model, *American journal of climate variability*, 2: 106-115.

FAO. (2013). Crop evapo-transpiration guidelines for computing crop water requirements; cooperate document repository report.

FAO. (2013). Report on land and water, www.fao.org/nrl/aboutnr/nrl, accessed on 5th February, 2014.

FAOSTAT. (2000). AQUASTAT Land use database, FAO, Rome.

Firat, M. and Gungor, M. (2008). Hydrological time-series modelling using an adaptive neuro-fuzzy inference system, *Journal of hydrological processes*, 22 (13): 2122-2132.

Fiseha, B. M., Setagn, S. G., Melesse, A. M., Volpi, E. and Fiori, A. (2013). Hydrological analysis of the upper Tiber river basin, central Italy: A watershed modelling approach, *Hydrological processes*, 27:2339-2357.

Food and Agriculture Organization of the United Nations (FAO). (2011). Kenya country profile, http://www.fao.org./countries/55528/en/ken, accessed on may 26th, 2014.

Gao, B. C. (1996). NDWI, A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote sensing and environment*, 58(3): 257-266.

Ghulam, A., Li, Z. L., Qin, Q. and Tong, Q. (2007a). Exploration of the spectral space based on vegetation index and albedo for surface drought estimation, *Journal of Applied remote sensing*, 1(13529): 1-12.

Gibbs, W. J. and Meher, J. V. (1967). Rainfall Deciles as drought indicators bureau of Meteorology, Bulletin No. 4, commonwealth of Austarlia, Melbbourne Austria.

GoK. (2012).Tana river delta strategic environmental assessment scoping report, ministry of lands, in collaboration with Ministry of Prime minister, Ministry of state for planning national development and vison 2030.

GoK. (2007). The Kenya Vision 2030, report by the Ministry of planning, Republic of Kenya.

GoK. (2014). The outlook for March-April-May 2014 long rains season and review of weather during Oct-Dec, 2013 short rains season, Jan-Feb 2014 period, *press release*. Ref No. KMD/FCST/5-2014/SO/01. www.meteo.go.ke/ranet/Wx/seasonal.pdf, accessed on 8th March, 2014.

GoK. (2012). Upper Tana natural resources management project; A strategic environmental assessment draft report.

Gommes, R. A. and Petrassi, F. (1994). Rainfall variability and drought in sub-saharan Africa since 1960, FAO Agromet report series WP9, Rome Italy.

Gonzalez, J. and Valdes, J. (2006). New drought frequency index, definitions and evaporative performance analysis, *Water resources Res,* 42(11): 333-349.

Government of Kenya (GoK). (2009). Changes in forest cover in Kenya's five water towers 2000-2007.

Gu, Y., Jesslyn, F., Brown, J., Verdin, P. and Wardlow, B. (2007). A five year analysis of MODIS NDVI for grass drought assessment over central Great Plains of the U.S. *Journal of geophysical research letters*, 34: 1-6.

Hayes, M. J. (2003). Drought Indices, national drought mitigation centre, university of Nebraska-Lincoln, http://www.drought.unl.edu/whatis/indices.htmspiaccessed on February, 9th 2014.

Haykin, S. (1994). Neural networks: A comprehensive foundations, Prentice Hall, Upper Saddle River New Jersey, U.S.A.

Huete, A Jackson, R. and Post, D. (1985). Spatial response of a plant canopy with different soil backgrounds, remote sensing *Environment Sci*ence, 17(1): 37-53.

Hundecha, Y. and Bardossy, A. (2004). Modelling of the effect of land use changes on the runoff generation of river basin through regionalization of watershed model, *Journal of hydrology*, 292(1-4): 281-295.

IFAD. (2012). Upper Tana catchment natural resource management project report, east and southern Africa division, project management department.

IGAD. (2007). Annual report and strategic plan for 2004-2008.

IPCC. (1995). Climate variability. The second IPCC Scientific Assessment. Intergovernmental Panel on Climate variability, World Meteorological Organization/United nations Environment Programme, Houghton, J.T., Meira Filho, L.G., Callendar, B.A. (Eds), Cambridge University press.

IPCC. (2001). Third Assessment Report of the Intergovernmental Panel on Climate variability. Parts 1, 2 and 3, Synthesis Report and Policy Makers Summaries. Cambridge University Press, Cambridge, UK.

IPCC. (2007). A report of Working Group I of the IPCC Summary for policy makers.

IPCC. (2007). Climate variability 2007-synthesis Report. Contribution of working Groups, I, II and III The 4th assessment report of the intergovernmental panel on climate variability, Pachauri, R. K. and Resinger, A. Eds, IPCC Secretariat, Geneva.

IPCC. (2013). Climate change 2013, fifth assessment report of IPCC, the physical science basis, summary for policy makers.

IPCC. (2014). Climate Change 2014, fifth assessment report on impacts, adaptation and vulnerability, WGII AR5, Summary for policy makers.

Iqbal, M. A., Shen, Y., Stricevic, R., Pei, H., Sun, H., Amini, E., Penas, A. and Rios, S. (2014). Evaluation of the FAO AquaCrop model for winter wheat on the north China plain under deficit irrigation from field experiments to regional yield simulation, *Agricultural Water Management Journal*, 135(2014): 61-72.

Jacobs, J. Angerer, J., Vitale, J., Srinivasan, R., Kaitho, J. and Stuth, J. (2004). Exploring the Potential Impact of Restoration on Hydrology of the Upper Tana RiverCatchment and Masinga Dam, Kenya, a Draft Report, Texas A & M University.

Jahangir, A. T. M., Sayedur, R. M. and Saadat, A. H. M. (2013). Monitoring meteorological and agricultural drought dynamics in Barind region, Bangladesh using SPI and Markov chain model, *International Journal of Geomatics and Geosciences*, 3 (3): 511-524.

Jain, S., Keshri, R., Goswami, A. and Saker, A. (2010). Application of meteorological and vegetative indices for evaluation of drought impact; a case study of Rajasthan, *India, national hazards,* 540: 643-656.

Jones, J. A. A. (2010). Soil piping and catchment response. *Journal Hydrological Processes*, 24: 1548–1566.

Jorgensen, S. E. and Bendoricchio, G. (2001). Fundamentals of ecological modelling, 3rd edition, Elsevier, Amsterdam.

Kalter, A. M. (2007). Rainfall-runoff Modelling using Artificial Neural Networks (ANNs), PhD thesis, Department of Water Resources Engineering, Lund University.

Kamga, F. M. (2001). Impact of greenhouse gas induced climate variability on the runoff of the upper Benue River, Cameroon, *Journal of hydrology,* 252: 145-156.

Karamouz, M. Rasouli, K. and Nazi, S. (2009). Development of a hybrid index for drought prediction: case study, *Journal of Hydrologic Engineering*, 14(6): 617-627.

Karamouz, M., Szidarovszky, F. and Zaharaie, B. (2003). Water resources systems analysis, Lewis Publishers, Florida, U.S.A

Karl, T. R. and Knight, R. W. (1985). Atlas of monthly palmer hydrological drought indices for the continuous United States, Asheville, N.C USA national climatic data centre, climatology series (3-7) report.

Kaufman, Y. J. and Tanne, D. (1992). Atmospherically resistant vegetation index for EOS-MODDIS IEEE trans. *Geoscence, remote sensing*, 30(2): 261-270.

Keyantash, J. A. and Dracup J. A. (2004). An aggregate drought index: assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage, *Journal of Water resources research*, 40(9): 1-14.

Keetch, J. J. and Byuram, C. M. (1968). A drought index for forest fire control, Res. Pap, SE-38, Asheville NC, US Department of agriculture, forest service, south eastern forest experimental station, Sheville N. C.

Kenya Food Security Steering Group (KFSSG). (2009). The 2009 long rains season assessment report, government of Kenya.

Kenya Food Security Steering Group (KFSSG). (2011). The 2009 long rains season assessment report, government of Kenya.

Kim, T. W. and Valdes, J. B. (2003). Non-linear model for drought forecasting based on a conjunction of wavelet transforms and neural networks, *Journal of Hydrologic Engineering*, 8 (6): 319-328.

Kjeldsen, T. R., Lundorf, A. and Dan, R. (2000). Use of two component exponential distributions in partial duration modelling of hydrological drought in Zimbabean river, *Journal of Hydrological Sci*ences, 45: 285-298.

Kogan, F. N. (1990). Remote sensing of weather impacts of vegetation in non-homogeneous areas. *International Journal of remote sensing*, 11(8): 1405-1419.

Kogan, F. N. (1995). Application of vegetative index and brightness temperature for drought detection, *Advances in Space Research*, 15(11): 91-100.

Krause, P., Boyle, P. P. and Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment, *Journal of advances in Geosciences*, 5: 89-97.

Kwon, H. J. and Kim, S. J. (2010). Assessment of distributed hydrological drought based on hydrological unit map using SWSI drought index in south Korea, *Journal of Civil Engineering*, 14(6): 923-929.

Lillesand, T. and Kieffer, R. (1987). Remote sensing and image interpretation, John Willey new Conc: 721.

Liu, X, Wang, S., Zhou, Y, Wang, F., Li, W. and Liu, W. (2015). Regionalization and spatiotemporal variation of drought in China based on standardized precipitation evapotranspiration index (1961-2013). *Advances in meteorology*, 2015: 1-18.

Loucks, D. P. and van-Beek, E. (2005). Water resources systems planning and management, an introduction to methods, models and applications, studies and reports in hydrology, UNESCO publishing Paris.

Luk, K. C., Ball, J. E. and Sharma, A. (2000). A study of optimal model lag and spatial inputs to artificial neural networks for rainfall forecasting, *Journal of hydrology,* 227(1-4): 56-65.

Luke, J., Wersonger, J., Jianshen, S. and Kopalle, D. (2001). Remote sensing applications from the Albama Space Program, Report.

Ma, X., Xu, J., Luo, Y., Aggarwal, S. P. and Li, J. (2009). Response of hydrological processes to land-cover and climate variability in Kejie watershed, south-west China, *Journal of Hydrological processes,* 23: 1179-1191.

Mahajan, D. R. and Dodamani, B. M. (2015). Trend analysis of drought events over upper Krishna basin in the Maharashtra, *Journal of Aquatic procedia,* 4: 1250-1257.

Maier, A. R., Jain, A., Dandy, G. C. and Sudheer, K. P. (2010). Methods used for development of neural networks for the prediction of water resource variables in river systems: current status and future directions, *Journal of Environmental modelling,* 25(8): 891-909.

Manikandan, M. and Tamilmani, D. (2013). Development of drought vulnerability maps in the Parambikulam-Alyar Basin, India, *Academic Journals*, 8(20): 778-790.

Markovic, R. D. (1965). Probability functions of the best fit to distributions of annual precipitation and runoff hydrology paper No. 8, Colorado State University.

Matera, A., Fontana, G. and Marleto, V. (2007). Use of a new agricultural drought index within a regional drought observatory. In methods and tools for drought analysis and management Edited by G. Rossi, Springer the Netherlands: 103-124.

Mckee, T. B. and Edwards, D. C. (1997). Characteristics of 20th century droughts in the United States at multiple time scales, *Journal of atmospheric science,* 634: 97-92.

Mckee, T. B., Doesken, N. J. and Kleist, J. (1993). The relationship of drought frequency and duration to time scales in proceedings of 8th conference on applied climatology, Anaheim, California, U.S.A: 179-184.

Meigh, J. R., Tate, E. L. and McCartney, M. P. (2002). Methods for identifying and monitoring river flow drought in Southern Africa. In proc. 4th Int. Conf. on FRIEND, March, 2002, Cape Town, SA IAHS Pub. No. 274.

Mendicino, G., Senatore, A. and Versace, P. (2008). Aground water resource index for drought monitoring and forecasting in mediterean climate, *Journal of hydrology*, 357 (3-4): 282-302.

Mishra, A. K. and Desai, V. R. (2006). Drought forecasting using feed-forward recursive models, *Journal of Ecological Modelling*, 198: 127-138.

Mishra, A. K. and Singh, V. P. (2010). A Review of Drought Concepts, *Journal of Hydrology*, 391 (1-2): 202-216.

Mishra, A. K. and Singh, V. P. (2011). Drought modelling-A Review, *Journal of Hydrology*, 403(2011): 157-175.

Mishra, A. K., Desai, V. R. and Singh, V. P. (2007). Drought forecasting using a hybrid stochastic and neural net-work models, *Journal of Hydrological Engineering*, 12(6): 626-638.

Mishra, S. S. and Nagarajan, R. (2011). Spatio-temporal drought assessment in Tel river basin using standardized precipitation index (SPI) and GIS, *Journal of Geomatics, natural hazards and risk*, 2(1): 79-93.

Mishra, S. S. and Nagarajan, R. (2012). Forecasting drought in the Tel River basin using feed forward recursive neural networks, international conference on environmental, biomechanical and biotechnology, 41 IPCBEE, IACSIT Press, Singapore.

Modarres, R. (2007). Stream flow drought time series forecasting, *Journal of Stochastic Environment Research and Risk Assessment*, 21: 223-233.

Mondal, A. and Mujumndar, P. P. (2015). Regional hydrological impacts of climate change implications and for water management in India, hydrological sciences and water security, past, present and future, *Proceedings of the 11th Kovacs Colloquium, Paris France, June 2014 IAHS Pub.366(2015),* doi:10.5194 piahs-366-34-2015

Moreira, E. E., Coello, C. A., Paulo, A. A., Pererira, L. S. and Mexia, J. T. (2008).SPI-based drought category prediction using log-linear models, *Journal of hydrology*, 352 (1-4): 116-130.

Morid, S. Smakhtin, V. and Bagherzadeh, K. (2007). Drought forecasting using artificial neural networks and time series of drought indices *International Journal of climatology*, 27 (15): 2103-2111.

Morid, S., Smakhtin, V. and Moghaddasi, M. (2006). Comparison of seven meteorological indices for drought in Iran, *International Journal of Climatology*, 26: 971-985.

Mulla, D. (2013). Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps, *Journal of Biosystems Engineering*, 114 (4): 358-371.

Musaningabe, C. G. (2012). Assessing vegetative drought from multi-temporal NDVI images, PhD thesis, university of Twente, the Netherlands.

Mustafa, M. R., Isa, M. H. and Rezaur, R. B. (2012). Artificial neural networks in Water Resources Engineering; Infrastructure and applications, *Journal of World Academy of Science, Engineering and Technology*, 6: 2-24.

Mutua, B. M. and Klik, A. (2007). Predicting daily stream flow in un-gauged rural catchments; the case of upper Tana River basin, Kenya, *Journal of spatial hydrology*, 5(2): 64-80.

Mwangi, E., Watterhall, F., Dutra, E., Giuseppe, F. D. and Pappenbenger, F. (2013). Forecasting Droughts in East Africa, *Journal of Hydrology and Earth Sciences*, 10: 10209-10230.

Nalbantis, I. (2008). Evaluation of hydrological drought index, *European water journal*, 23(24): 67-77.

Narasimhan, B. and Srinivasan, R. (2005). Development and evaluation of soil moisture deficit (SMDI) and evapo-traspiration deficit index (ETDI) for agricultural drought monitoring, *Journal of agricultural and forest meteorology*, 133: 69-88.

Nash, J. E. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models. *Journal of hydrology*, 10: 282-290.

NEMA. (2004). Kenya state of environment report: Chapter 7, fresh water, coastal and marine resources, Nairobi, Government printer.

Nikbakht, J., Tabari, H. and Talace, P. H. (2013). Streamflow drought severity analysis by percent of normal index (PNI) in northwest Iran, *Journal of Theoretical and Applied Climatology*, 112: 565-573.

Nohara, D., Kitoh, A., Hosaka, M. and Oki T. (2006). Impact of climate change on river discharge projected by multi-model ensemble, *Journal of Hydrometeorology*, 7: 1076-1089.

Ntale, H. K. and Gan, T. Y. (2003). Drought indices and their application to East Africa, *International journal of climatology*, 23: 1335-1357.

Ochoa-Rivera, J. C. (2008). Prospecting droughts with stochastic artificial neural networks, *Journal of Hydrology*, 352(1-2): 174-180.

Okoro, B. C., Uzoukwu, R. A. and Chiomezie, N. M. (2014). River basins of Imo State for sustainable water resources management, *Journal of Civil and Environmental Engineering*, 4(1): 1-8.

Oludhe, C. (2012). Impacts of climate variability on power generation within the 7-forks dams in Tana River basin, *Journal of meteorological and related sciences*, KMS 10th conference, 13-20.

Onyango, O. A. (2014). Analysis of meteorological drought in th north eatsren province of Kenya, *Journal of earth sciences and climate change*, 5(8): 1-6.

Otieno, F. A. O. and Maingi, S. M. (2000). Sedimentation problems of Masinga reservoir. In land and water management in Kenya. Eds. Gichuki F. N., Mungai, D. N., Gachere, C. K. and Thomas, D. B., published by Soil and Water Conservation branch of Ministry of Agriculture and Rural development, Nairobi, Kenya.

Palmer, W. C. (1965). Meteorological drought research paper 45, weather Bureau, Washington D.C, U.S.A.

Paulo, A. A. and Pereira, L. S. (2008). Stochastic Prediction of drought class transitions, *Journal of water resources management*, 22(9): 1277-1296.

Paulo, A. A. and Pereira, L. S. (2007). Prediction of SPI drought class transitions using Markov chains, *Journal of water resources management*, 21(10): 1813-1827.

Paulo, A. A., Ferreira, E., Coelho, C. and Pereira, L. S. (2005). Drought class transition analysis through markov and log linear model; an approach to early warning system, *Journal of agricultural water management*, 77(1-3): 59-81,

Pearson, R. C. and Miller, L. D. (1972). Remote mapping of crop biomass for estimation of productivity of short grass prairie, *Remote sensing environment*, 8(1): 1355-1367.

Peng, S., Liu, W., Wang, W. Shao, Q.,Jiao, X., Yu, Z., Xing, W., Xu, J., Shang, Z. and Luo, Y. (2013). Estimating the effect of climatic variability and human activities on stream flow in the Hutuo River Basin, China, *Journal of Hydrologic Engineering*, 18(4): 422-430.

Peters, E., Torfs, P., van-Lanen, H. and Bier, G. (2003). Propagation of drought through ground water-a new approach using linear reservoir theory, *Hydrological processes*, 17(15): 1-16.

Raes D., Steduto, P., Hsiano, T. C. and Fereres, E. (2009). AquaCrop-the FAO crop model to simulate yield response on water, *Agronomic Journal*, 101: 438-477.

Rajput, P., Sinha, M. K., Verma, M. K. and Ahmad, I. (2014). Drought hazard assessment and mapping in upper seonath sub-basin using GIS, *International Journal of emerging technology and advanced Engineering*, 4(10): 210-218.

Ramazani, E. H., Liaghat, A., Parsinezhad, M. and Ramazani, E. M. (2012). The study of agriculturl drought on soil moisture in Qazvin station, *Journal of water in agriculture*, 2(1): 84-93.

Raude, J. M. (2013). Hydraulic performance of horizontal constructed sub-surface flow constructed wetland system in treatment of grey water; the case of peri-urban areas in Nakuru Municipality, Kenya, PhD theis, Egerton University.

Richardson, A. J. and Wiegand, C. L. (1977). Distinguishing vegetation from soil background information, *photogram Engineering Remote sensing*, 43: 1541-1552.

Roudier, P. and Mahe, G. (2010). Study of water stress and droughts with indicators using daily data on Bani river, Niger basin, Mali, *International Journal of climatology*, 30:1689-1705.

Saenyi, W. W. (2002). Sediment management in Masinga reservoir, Kenya, PhD thesis (Published), University of Agricultural Sciences (BOKU), Vienna Austria.

Sayanjali, M. and Nabdel, O. (2013). Remote sensing satellite design using model based systems Engineering, *Journal of Science and Engineering*, 1(1): 43-54.

Scheffran, J., Marmer, E. and Show, P. (2012). Migration as a contribution to resilience and innovation in climate adaptation: social networks and co-development in North West Africa, *Applied geography,* 33: 119-127.

Shafer, B. A. and Desman, L. E. (1982). Development of a Surface Water Supply Index (SWSI) to assess drought conditions in snowpack Runoff Areas, proceedings of the Western snow conference Reno, Nevada, U.S.A.: 164-175.

Shamar, T. C. and Panu, U. S. (2013). A semi-empirical method for predicting the hydrological drought magnitude in Canadian Prairies, *Hydrological sciences journal*, 58(3): 549-569.

Sharda, V., Srivasta, P., Kalin, L., Ingram K. and Chelliah, M. (2012). Development of community water deficit index (CWDI) -Drought forecasting tool for small to mid-size communities of south eastern united state, *Journal of Hydrologic Engineering*, 18(7): 846-858.

Shatanawi, K., Rahbeh, M. and Shatanawi, M. (2013). Characterizing, monitoring and forecasting of drought in Jordan river basin, *Journal of water resources and protection*, 5: 1192-1202.

Shen, H., Yuan, F., Ren, L., Ma, M., Kong, H. and Tong, R. (2015). Regional drought assessment using a distributed hydrological model coupled with standardized runoff, *Journal of remote sensing and GIS for hydrology and water resources*, 368: 397-402.

Sivaprakasam, S., Murugappan, A. and Mohan, S. (2011). Modified Hargreaves equation for estimation of ETo in a hot and humid location in Tamilnadu state, India, *International Journal of engineering science and technology*, 3(1): 592-600.

Smakhtin, V. U. and Hughes, D. A. (2007). Automated estimation and analysis of meteorological drought charactersitsics from monthly rainfall data, *Journal Environmental modelling and software*, 22(6): 880-890.

Stahl, K. (2001). Hydrological drought, a study across Europe, PhD thesis, Institut für Hydrologie der Universität, Freiburg i. Br.

Sung, J. H., Chung, E. S. and Lee, K. S. (2013). Development of stream flow drought severity and magnitude duration-frequency curves using threshold level method, *Journal of Hydrology and Earth Sci*ences, 10: 14675-14704.

Tadesse, T. and Wardlow, B. (2007). The vegetation outlook, a new tool for providing outlooks of general vegetation data mining tech, In IEEE Int. Conference on data mining workshops :667-672.

Tessema, R. S. (2007). Agricultural drought assessment for upper Blue Nile basin Ethiopia using SWAT, MSc. Thesis, UNESCO-IHE Institute of water education

Thenkabail, P. S., Ward, A. D. and Lyon, J. G. (1994). Landsat 5 thematic mapper models for soybean and corn characteristics. *International Journal of remote Sens*ing, 15(1): 49-61.

Timm, L. C., Gomes, D. T., Barbosa, E. D., Reichardt, K., Souza, M. D. and Dynia, J. F. (2006). Neural Network and state space models for studying relationship among soil properties, *Journal Scientia Agricola*, 63: 386-395.

Tong, S. T. Y. and Chen, W. (2002). Modelling relationship between land use and surface water quality, *Journal of Environmental Management*, 66(44): 377-393.

Tran, H. D., Perera, B. J. C. and Ng, A. W. M. (2009). Comparison of structural deterioration models for storm water drainage pipes, *computer aided civil and infrastructure engineering journal*, 24(2): 145-156.

Tsakiris, G. and Vangelis, H. (2005). Establishing a drought index incoproprating evapo-transpiration, *European water Journal*, 9(10): 3-11.

Tsakiris, G. and Nalbantis, I. (2009). Assessment of hydrological drought revised, *water resources management*, 23: 881-897.

Toushmalani, R. (2013). Comparison results of inversion of gravity data a fault by particle swarm optimization and Levernberg-Marquardt methods, *springerplus journal*, 2: 1-6

Tychon, B., Pecheur, C. and Ozer, P. (2007). The NDWI as a drought index applied to Belgium and Heilongjians in Begian, *Chinese crop growth monitoring systems, comparison, adaptation and improvement*, I: 111-120.

UN. (2007). African review report on drought and desertification, fifth meeting on African committee on sustainable development, Addis Ababa, October 22nd-25th.

UN. (2008). Trends in sustainable development, agriculture, rural development, land desertification and drought. Department of economic and social affairs, United Nations, New York.

UN. (2016). Sustaibale development goals; 17 goals to transform our world, www.un.org/sustainable development/sustainable-development-goals/, accessed on 16th August 2016.

UNEP. (2015). Promoting ecosystems for disaster risk and climate change adaptation, opportunities for intergration, Discussion paper.

UNDP. (2012). Kenya: adapting to climate variability in Arid and Semi-Arid Lands (KACCAL), project report Wilby, R. L., Orr, H. G., Hedger, M, Forrow D. and Blakmore, M. (2006a). Risks posed by climate variability to delivery of water framework directive objectives. *Environ. Int.*, in press.

Van-loon, A. F. (2015). Hydrological drought explained, *WIREs, water journal*, 2:359-392.

Van-loon, A. F. and Laaha, G. (2015). Hydrological drought severity explained by climate and catchment characteristics, *Journal of hydrology*, 526 (2015): 3-14.

Van-rooy, M. P. (1965). A Rainfall Anormally Index (RAI), independent of the time and space, *Notos*, 14: 43-48.

Vasiliades, L. Loukas, A. and Liberis, N. (2011). A water balance derived drought index for Pincos River basin, Greece, *Journal of Water resources management*, 25(4): 1087-1101.

Vicente-Serano, S. M., Begneria, S. and Lopez-Moreno, J. I. (2010). A multi-scalar drought index sensitive to global warming, the standardized precipitation evapotranspiration index, *Journal of climatology*, 23(7): 1696-1711.

Wang, L. and Qu, J. J. (2007). NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote Sensing, *Journal of Geophysics Reserach Letters*, 34(20): 1-5.

Wang, W. (2010). Drought analysis under climate change by application of drought indices and copulas, MSc thesis in Civil and Environmental Engineering, Portland State University.

Weghorst, K. (1996). The reclamation drought index, guidelines and practical applications, ASCE, Denver Colorado.

Wilby, R. L. and Harris. (2006). A framework for assessing uncertainities in climate change impacts, low flow scenarios for the river thames, UK, water resources research, 42: 1-10.

Wilhite, D. A. (2004). Drought as a natural hazard, in international perspectives on natural disasters ;occurrence, mitigation, and consequences, edited by J. P. Stollman, J. Lidson and L. M. Dechano, Kluwer academic publishers, the Netherlands: 147-162.

Wilhite, P. A. and Clantz, M. (1955). Understanding the drought phenomena, the role of definitions, *Water International*, 10(3): 111-120.

Wilmot, C. J. (1984). On the evaluation of model performance in physical Geography in spatial statistic and models, edited by Graile G. L. and Wilmot C. J, Deidel D: 443-460.

Wilschut, L. I. (2010). Land use in the upper Tana: Technical report of a remote sensing based land use map. In green water credits report 9 edited by Mcmillan B., Kauffmann, S. and De Jon, R. Wageningen, ISRIC-world soil information.

World Bank. (2014). Report on climate variability and water resources degradation in Kenya; improving water resources development and management, Report No. 69, http://wdronline.worldbank.org/worldbank/p/developmentdatabase, accessed on 15th March, 2014.

World Bank. (2013). A world bank report No. 8 on Kenya economic update, time to shift gears; accelerating growth and poverty reduction in the new Kenya.

World food programme (WFP). (2011). Drought and famine in Horn of Africa.

World Health Organization (WHO). (2010). Water requirement, Geneva, Switzerland.

World meteorological organization (WMO). (2015). A report by commission for climatology, http://public.wmo,int, accessed on 15th February 2016.

World Resources Institute (WRI). (2011). Kenya GIS data –world resources institute, retrieve from www.wri.org/resources/data-sets/kenya-gis-data on January 15, 2014.

World Wild Life Fund (WWF). (2006). Climate variability impacts on east Africa, a review of scientific literature, Gland, Switzerland.

WRMA. (2010). Physiological survey in the upper Tana catchment, a natural resources management project report, Nairobi.

WSI. (2011). Baseline review of the upper tana, Kenya, green water credits report 8.

Wu, Y., Bake, B., Zhang, J. and Rasulov, H. (2015). Spatio-temporal patterns of drought in north Xinjiang, China 1961-2012 based on meteorological drought index, *Journal of Arid Lands*, 7(4): 527-543.

Yamada, K. (2011). Network parameter setting for reinforcement learning approaches using neural networks, *Journal of advanced computational intelligence and intelligent informatics*, 15 (7): 822-930.

Yan, D. H., Wu, D., Huang, R., Wang, L. R. and Yang, G. Y. (2013). Drought evolution characteristics and precipitation intensity changes during alternating dry-wet changes in Huang-Huai-Hai River basin, *Journal of Hydrology and Earth Systems Sciences*, 10: 2665-2696.

Yang, W. (2010). Drought analysis under climate change by application of drought indices and copulas, MSc thesis, Portland State University, USA.

Yan, D., Shi, X., Yang, Z., Li, Y., Zhao, K. and Yuan, Y. (2013). Modified palmer drought severity index based on distributed hydrological simulation, *Journal of mathematical problem in Engineering*, 2013: 1-8.

Yao, Y. B., Yang, J. H. Yue, P. and Lu, D. (2013). Climatic change of terrestrial surface humid index and its impact factors over the source region of three rivers in recent 50 years, *ecology and environmental science journal*, 20(1): 1585-1593.

Zarco-tejada, P. J. and Ustin, S. L. (2001). Modelling canopy water content for carbon estimates form MODDIS data at land EOS validation sites in *geosciences and RS symbosium*, 1: 342-344.

Zargar, A., Sadiq, R., Naser, B. and Khan, F. I. (2011). A review of drought indices, *Journal of Environmental Research*, 19: 333-349.

Zhou, S. W., Wang, C. H., Wu, P. and Wang, M. (2012). Temporal and spatial distribution of strong precipitation days over the Tibetan plateau, *Arid Land Geography*, 35(1): 23-31.

Zoljoodi, M. and Didevarasl, A. (2013). Evaluation of Spatio-temporal variability of droughts in Iran using Palmer Drought Severity Index and its precipitation factors through (1951-2005), *Atmosphere and Climate Sciences Journal*, 3: 193-207.

APPENDICES

APPENDIX A: TABLES

Table 1A: The major drought impacts for different drought categories

Impact category	Impact sub-category	Hydrological drought	Meteorological drought	Agricultural/ soil moisture drought
	Rain-fed		x	x
Agriculture	Irrigation	X		x
River	Terrestrial	X	x	x
basins/ecosystems	Cooling	X		
Energy and	Hydro-power	X		
industry	Cooling water	X		
Navigation		X		
Drinking water		X		
Recreation		X		

Source: Van-loon and Laaha (2015)

Table 2A: Best ANN structure used in filling of missing stream flow data for each gauge station

Gauge station	Gauge ID	Epoch No	Coordinates		Input function into the neurons	ANN architecture	R Train	R validation	MSE train	MSE validation
			Longitude	Latitude						
Amboni	4AB05	9	36.989	-0.350	$Q_{(t+1)} = f(Q_t, Q_{t-1}, Q_{t-2})$	3-9-1	0.545	0.735	0.214	0.105
Sagana	4AC03	14	37.043	-0.449	$Q_{(t+1)} = f(Q_t, Q_{t-1})$	2-9-1	0.621	0.724	0.312	0.213
Gura	4AD01	20	37.076	-0.517	$Q_{(t+1)} = f(Q_t, Q_{t-1})$	2-6-1	0.583	0.655	0.362	0.351
Tana sagana	4BC02	9	37.207	-0.672	$Q_{(t+1)} = f(Q_t, Q_{t-1})$	2-5-1	0.643	0.563	0.384	0.346
Yatta furrow	4CC03	20	37.361	-1.094	$Q_{(t+1)} = f(Q_t, Q_{t-1}, Q_{t-2})$	3-9-1	0.701	0.732	0.276	0.242
Nyamindi	4DA10	8	37.317	-0.621	$Q_{(t+1)} = f(Q_t, Q_{t-1})$	2-2-1	0.684	0.654	0.332	0.344
Rupingazi	4DC03	9	37.438	-0.533	$Q_{(t+1)} = f(Q_t, Q_{t-1}, Q_{t-2})$	3-6-1	0.595	0.673	0.346	0.318
Kamburu	4ED01	10	37.683	-0.800	$Q_{(t+1)} = f(Q_t, Q_{t-1})$	2-2-1	0.643	0.686	0.278	0.229
Mean values							0.627	0.678	0.313	0.269

Table 3A: Best ANN structure used in filling of missing precipitation data for each gauge station

Gauge station	Gauge ID	Epoch No	Coordinates		Input neurons	ANN architecture	R Train	R validation	MSE train	MSE validation
			Longitude	Latitude						
MIAD	9037112	1000	37.350	-0.700	$P_{(t+1)} = f(P_t, P_{t-1}, P_{t-2})$	3-10-1	0.758	0.897	0.325	0.056
Embu	9037202	80	37.450	-0.500	$P_{(t+1)} = f(P_t, P_{t-1}, P_{t-2})$	3-9-1	0.654	0.752	0.287	0.172
Kerugoya DWO	9037031	100	37.327	-0.382	$P_{(t+1)} = f(P_t, P_{t-1})$	2-5-1	0.664	0.718	0.361	0.123
Sagana FCF	9037096	90	37.054	-0.448	$P_{(t+1)} = f(P_t, P_{t-1}, P_{t-2})$	2-4-1	0.587	0.653	0.294	0.195
Nyeri	9036288	100	36.970	-0.500	$P_{(t+1)} = f(P_t, P_{t-1}, P_{t-2})$	3-10-1	0.625	0.694	0.302	0.237
Muragua G. E. F.	9036212	50	36.850	-0.750	$P_{(t+1)} = f(P_t, P_{t-1})$	2-4-1	0.581	0.635	0.275	0.228
Naro-moru F.G.P.	9037064	80	37.117	-0.183	$P_{(t+1)} = f(P_t, P_{t-1})$	2-3-1	0.596	0.736	0.311	0.279
Mangu HS	9137123	100	37.033	-1.100	$P_{(t+1)} = f(P_t, P_{t-1})$	2-5-1	0.603	0.729	0.273	0.234
Mean values							0.634	0.727	0.304	0.191

Table 4A: Procedure for objectively determining of weighting parameters for SWSI

Steps:

i) Determine the monthly averages of each component

ii) Normalize the average monthly values by using the normalization equation 3.38 (Xmax=0.9; Xmin=0.1) /dividing through by average value for the highest month

$$X_n = X_{\min} + \frac{(X_o - x_{\min})}{(x_{\max} - x_{\min})} \times (X_{\max} - X_{\min}) \quad (3.38)$$

iii) Monthly coefficient for each component is determined by determining/calculating each fractional contribution (to a total of 1 as per the illustration below)

Example coefficient a,b,c,d,e (average for one station)

Month (C_1)	Average value of P (mm)=(C_2)	Xn=C_3	Highest value of the record H =(C_4)	Precipitation $X_P = \frac{C_3}{C_4}$ =(C_5)	Stream flow X_n=C_6	Dam inflow, X_n=C_7	a =C_8	b =C_9	c =C_{10}
Jan	12	0.90	0.90	**1.00**	**0.93**	**0.25**	**0.46**	**0.43**	**0.11**
Feb	10	0.77	0.90	0.85	0.81	0.72	x	x	x
March	0	0.10	0.90	0.11	0.3	0.61	x	x	x
Apr	5	0.43	0.90	0.48	0.52	0.35	x	x	x
May	6	0.50	0.90	0.56	0.71	0.24	x	x	x
Jun	7	0.57	0.90	0.63	0.75	0.37	x	x	x
Jul	8	0.63	0.90	0.70	0.5	0.67	x	x	x
Aug	10	0.77	0.90	0.85	0.41	0.51	x	x	x
Sept	11	0.83	0.90	0.93	0.62	0.5	x	x	x
Oct	2	0.23	0.90	0.26	0.25	0.34	x	x	x
Nov	4	0.37	0.90	0.41	0.48	0.21	x	x	x
Dec	11	0.83	0.90	0.93	0.85	0.57	x	x	x

Consider the row for January

The ratio is of a:b:c is 1.00:0.93:0.25 which must total to 1

Thus the total for this January row is 1.00+0.93+0.25=2.18

The contribution of component a (for rainfall) is thus calculated as

$$a = \frac{NP_{Jan}}{T} \times 1 = \frac{1}{2.18} \times 1 = 0.46$$

$$Similarly \;\; b = \frac{0.93}{2.18} \times 1 = 0.43$$

$$c = \frac{0.25}{2.18} \times 1 = 0.11$$

$$Thus\, checking \;\; 0.46 + 0.43 + 0.11 = 1.00$$

Similarly the procedure is repeated for coefficients for all the other months to fill the x values

Table 5A: Average characteristics of dominant soils for upper Tana River basin

Elevation	Dominant soil type	MC at saturation %	MC at field capacity %	MC at wilting point %	AWC (%)	TAW (mm)
HE	Andosols	60	40	24	16	172
ME	Nitosols	53	31	22	9	98
LME	Cambisols	48	28	14	14	74
LE	Ferrasols	53	26	17	9	82

HE, ME, LME, LE means highest elevation, middle elevation, lower middle elevation and lowest elevation respectively

Source: Hunink *et al*., (2009)

Table 6A: Agro-climatic zones of Kenya

Zone	Clasiification	Mean annual rainfall R (mm)	Evapo-transpiration E_0 (mm)	R/E_0 ratio	Approximate actual area (km^2)	% of the total land area	Potentail for plant growth
I	Humid	1400-2700	1200-2000	>80	17411.1	3	Very high
II	Sub-Humid	1000-1600	1300-2100	65-80	23214.8	4	high
III	Semi-Humid	800-1400	1450-2200	50-65	29018.5	5	High to medium
IV	Medium to Semi-Arid	600-700	1500-2200	40-50	29018.5	5	Medium
V	Semi-Arid	500-600	1650-2300	25-40	87055.5	15	Medium to low
VI	Arid	300-550	1900-2400	15-25	127681.4	22	Low
VII	Very Arid	<300	2100-2500	<15	266970.2	46	Very low

58 % of the land amounting to 394,651.6 km^2 lies within the arid and semi-arid lands (ASALs)

Source: FAO (2006)

Table 7A. Different drought categories distribution as detected using PDSI in October 1970

Drought category	Drought criterion	Area (km^2)	Percent
Extreme drought	-4 or less	3758.01	21.57
Severe drought	-3 to -2.99	1784.90	10.25
Mild drought	-2.00 to -2.99	2062.56	11.84
Slight drought	-1.00 to -1.99	2643.58	15.18
Normal	0.49 to -0.49	1946.48	11.17
Slightly wet	2.00 to 2.99	1782.32	10.23
Moderate wet	3.00 to 3.99	1681.05	9.65
Extremely wet	4.00 or more	1761.10	10.11

Table 8A: Different drought categories distribution as detected using PDSI in October 2010

Drought category	Drought criterion	Area (km^2)	Percent
Extreme drought	-4 or less	4540.36	26.06
Severe drought	-3 to -2.99	2537.551	14.57
Mild drought	-2.00 to -2.99	1675.444	9.62
Slight drought	-1.00 to -1.99	1824.072	10.47
Normal	0.49 to -0.49	1964.556	11.28
Slightly wet	2.00 to 2.99	1893.08	10.87
Moderate wet	3.00 to 3.99	1420.637	8.16
Extremely wet	4.00 or more	1564.297	8.98

Table 9A: Mean frequency of drought in the upper Tana River basin

Frequency for entire basin		
Drought type	Mean drought Frequency (years)	Mean Frequency of extreme drought (years)
Hydrological	4	6
Meteorological	3	5
Agricultural	3	4
Most prone-lower southeastern areas		
Hydrological	3	4
Meteorological	2	3
Agricultural	2	3

Table 10A: List of the best network models for drought forecasting of different Drought Indices

Drought Index	Network input variable	The best net work model
SWSI	*SWSI, P*	$SWSI_{(t+n)} = f(SWSI_t, SWSI_{t-1}, SWSI_{t-2}), (P_t, P_{t-1}, P_{t-2})$
SDI	*P, Q, SDI*	$SDI_{(t+n)} = f(P_t, P_{t-1}), (Q_t, Q_{t-1}), (SDI_t, SDI_{t-1})$
SPI	*P, SPI*	$SPI_{(t+n)} = f(P_t, P_{t-1}, P_{t-2}), (SPI_t, SPI_{t-1})$
SDI	*P, EDI*	$EDI_{(t+n)} = f(P_t, P_{t-1}, P_{t-2}), (EDI_t, EDI_{t-1})$
SMDI	*P, SMDI*	$SMDI_{(t+n)} = f(P_t, P_{t-1}), (SMDI_t, SMDI_{t-1}, SMDI_{t-3})$
NDI	*P, NDI*	$NDI_{(t+n)} = f(NDI_t, NDI_{t-1}, NDI_{t-2}), (P_t, P_{t-1}, P_{t-2})$
PDSI	*PDSI, P*	$PDSI_{(t+n)} = f(PDSI_t, PDSI_{t-1}, PDSI_{t-2}), (P_t, P_{t-1})$

Table 11A: Elevation bands for the upper Tana River basin used in this study

Elevation band	Elevation symbol	Elevation (m amsl)
Lowest elevation	LE	700-1800
Lower middle elevation	LME	1700-2800
Middle elevation	ME	2800-3900
Highest elevation	HE	3900-5000

Table 12A: Occurrence of historical drought in Kenya

Period (years)	Areas significantly affected	Remarks on the drought effects
1883	Coast	Caused worst famine in 30 years
1889-1890	Coast	One year of drought and femine
1894-1895	Coast	Information on distinct effects not available
1896-1900	Countrywide	Three consecutive rain seasons failed causing human deaths
1907-1911	Lake victoria, Machakos, Kitui and Coastal	Minor food shortages
1913-1919	Eastern and coastal areas	Impacts increased by war
1921	Coastal	Dry year recorded at coast
1925	Rift valley, central and coastal	Food shrtage, crop and livestock losses
1938-1939	Norhern, rift valley and central regions	Loss of livestock, deaths occured, Lorian swamp dried up
1947-1950	Central and coastal lands	Very severe droughtespecially in coast region
1952-1955	Eastern, central, coast, nyanza, western and rift valley regions	Food and water shortages
1960-1961	Eastern, rift valley	Caused Cattle deaths
1972	Widespread countrywide drought	Water shortage livestock deaths
1973-1974	Most areas in Kenya	Human and livestock deaths
1974-1976	Eastern, nothern, and central regions	Heavy livestock losses, food and water shortages
1980	Central, eastern, western and coast	Low crop production, water shortages
1981	Eastern	Famine and water shortage
1983	Country wide	Water shortages, human and livestock migration
1984	Central, rift valley, eastern, north eastern	Huge food shortages
1987	Eastern and central	Severe food shortages
1992-1994	Norhern, central and eastern	Moderate food shortages
1999-2000	Countrywide	Deficit food supply, interuption of electricity supply, water scarcity
2010-2011	Eastern, central, coastal,north eastern(Machakos, Kitui,Mwingi and Tharaka)	Food and water deficit

Source: (UNDP, 2012)

Table 13A: Definition of soil types found within different elevations within upper Tana River basin (FAO)

S. No	Soil type	Description	Elevation
1	Andosols	-These are soil types which are highly porous, dark colured soils. the are developed from parent material or volcanic origin , with high aluminium content	High
2	Vertisols	-Clayey soils that form deep and wide cracks when dry	Mid
3	Ferrasols	-Highly weathered soils rich in sesquiside clays and with low cation exchange capacities	Low
4	Nitosols	-Are soils with shiny surfaces on structural faces of the soil	Low

Table 14A: Summary of different drought indices for (Hydrological (H), Meteorological (M), Agricultural (A) and Remote Sensing (RS))

Drought index/Developed by	Type of drought	Key notes ,
Surface Water Supply Iindex (SWSI), (Shafer and Dezman, 1982)	H	Calculates the weighted average of the standardized anomarlies for precipitation,reservoir storage, runoff and snowpack
Palmer Drought Severity Index (PDSI), (Palmer, 1965)	H	Analysies precipitation and temperature in a water balance model. It compares meteorological and hydrological drought on spatial and temporal domain
Aggregated Drought Index (ADI), (Kayantash and Dracup, 2004)	MHA	It is a multivariate, aggregate drought index with five to six hydrological inputs such as precipiattion, streamflow, reservoir storage, evapotranspiration, soil moisture and snow water content. Its fist priciple component(PC1) is normalized by the standard deviation
Standardized Precipitation Index (SPI), (Mckee *et al*., 1993)	M	It is a simple index,has reliable spatial consistency, its probablistic in nature and useful in risk and decision analysis. It is adjustable to user defined time periods and can be used to determine three main dimensions of drought; intensity, duration and spatial extend
Crop Moisture Index (CMI), (Palmer 1968)	A	It analyses precipitation and temperature in a water balance model
Normalized Difference Drought Index (NDVI)	A	It uses visible red and near infrared bands for calculation of vegetation conditions. An adnanced radiometer is used to capture ad reflect the bands.
Regional Stream flow Defiiciency Index (RSDI), (Stahl, 2001)	H	It is used to characterize drought within a homogeneous region. It uses flow duration curve for for discharges that exceed 90% of the time defined as Q90. A RSDI is computed for each homogeneous region using time series stream flow
Soil moisture deficit index (SMDI), (Narasimhan and	A	It considers spatial variability of hydrological parameters of soil type, land cover and meteorological

Srinivasan, 2005)		variables.
Palmer Modified Drought Index (PMDI), (Palmer 1965)	M	It is used to calculate the beginning and end time of drought and wet periods which may otherwise be impossible to detect using the PHDI.
Z-Index, (Palmer 1965)	M	a monthly standrdized anormally of the available moisture; its an intermediate term within the oroginal PDSI and used for short term drought forecasting
Effective Drought Index (EDI), (Byun and Wilhite, 1999)	M	Used to detect drought beginning and ending and accumulated stress while ignoring the effect of runoff and ET
Reclamation Drought Index (RDI), (Weghorst, 1996)	H	Similar to SWSI and incorporates temperaure-variable and duration into basin-wise index computation
Crop Water Stress Index (CWSI), (Idso *et al*., 1981; Jackson *et al*., 1981)	RS	The CWSI is calculated from the expression: $CWSI = 1 - \frac{AET}{PET}$ where aet is the actual ET, PET is the potential ET. The index is mainly used for irrigation scheduling
Vegetation Condition Index (VCI), (Kogan 1990)	RS	It is used to determine the departure of amount of NDVI from long term NDVI as a measure of the health of the vegetation
Crop Specific Drought Index (CSDI) (Meyer *et al*., 1993)	A	It uses the crop and soil phenology and climatological data to estimate soil water availability for different zones and soil types
Keetch-Byram Drought Index (KBDI), (Keetch and Byram, 1968)	M	M for analysinsg precipitation and soil moisture in a water budget model. It is used to monitor forest fires and is for fire control and management
Temperature Condition Index (TCI), (Kogan, 1995)	RS	It uses te brightness temperature to determine the deviation of month from recorded maximum temperature.in general, the highre the temperature the higher the drought. It is computed from: $TCI = \left(\frac{BT_{\max} - BT_j}{BT_{\max} - BT_{\min}} \right) \times 100$ Btmax=maximum brightness temperature,Btmin=minimum brightness temperaturebtj=the jth month brightness temperature
Normalized Difference Water Index (NDWI), (Gao, 1996)	RS	Used to determine volumetric water content based on physical principles
Vetetaion Health Index (VHI), (Kogan, 1995)	RS	It combines vci and tci with some weight factor a for the contribution of vci and tci. The factos amay be set to 0.5 where information is lacking . the following relation applies: $VHI = a \times VCI + TCI \times (1 - a)$
Evapo-transpiration Drought Index (ETDI) (Narasimhan and Srinivasan, 2005)	A	It combines the hydrological parameters of soil and land use material in a spatial domain. The ETDI considers water stress ratio as expressed in the form: $ETDI = \frac{PET - AET}{PET}$ weekly values reflect short-term droughts
Standardized precipitation index (SPI), (Vicente-Serrano *et al*.,	M	Used to fatcor in the impact of cliamte change on drought characteristics which could not be detected by

2010)		the original spi. It uses both the precipitation and the temperature data sets. It considers water balance and evapotranspiration
Hybrid Drought Index (HDI), (Karamourz *et al*., 2009)	M,H,A Integrated	It combines the SPI, PDSI, and SWSI to represent various drought impacts
Vegetation Drought Response Index (VegDRI) (Brown *et al*., 2008)	M.A,RS	Used to characterize specific droughts and combines indices SPI,PDSI and NDVI
Modified Perpendicular Drought Index (MPDI), (Ghulam *et al*., 2007a)	RS	It is best used to accurately differentiate between the surfaces that are bare soils and densely vegetated agricultural or forest areas. A vegetation fraction that considers soil moisture and vegetation growth is used for undulating topography and variable soil types, the mpdi performs better than PDI.
Recconnaissance Drought Index (RDI), (Tsakiris and Vangelis, 2005)	M	It uses the precipitation and potential evapotranspiration. It is a more comprehensive DI than SPI. It is a physically based and calculates the deficit between the evaporative demand of atmosphere and precipitation. It is a simple index in computation of drought.
Normalized Multi-band Drought Index (NMDI), (Wang and Qu, 2007)	RS	It is an index used to improve sensitivity od ndwi and ndii by using inrformation from nir and swi bands. It extracts both vegetation and soil water content.
Relative Soil Moisture (RSM) Index, (Thornthwaite and Mather, 1955)	A	It is a water component calculated from water balance. It considers climate, soil physical properties, potentail ET ,precipitation, and crop charcetristics and managemnet pracices variables
Agricultural Drought Index, (Matera *et al*., 2007)	A	Uses water balance model and crop transpiration to calculate an integrated transpiration deficit
Rainfall Anormally Index (RAI), (Van-rooy 1965)	M	An average precipitation based on weekly, monthly and annual time periods is used to characterize relative drought. The resulting drought is then ranked with respect to fisrt ten severe droughts in a long-term record. The drought is then asigned a magnitude based on resulting values
Drought Severity Index (DSI), (Bryant *et al*., 1992)	M	Accumulated monthly deficit of precipitation, temperature and ground water conditions data
National Rainfall Index (NRI), (Gommes and Petrassi, 1994)	M	The total annual precipitation agianst its long term average are correlated. It reveals patterns and abnormalities of yearly precipitation on continental scale.
Drought frequency index (DFI), (Gonzalez and Valdes, 2006)	M	The index uses the mean frequency of of reccurrence for assessing the magnitude of drought in an area.
Ground water resource index (GRI), (Mendicino *et al*., 2008)	H	A simplified distributed water balance model is applied. It considers river basin geo-physical characteristics that affect hyrological response to precipitation
Water balance drought derived index (Vasiliades *et al*., 2011)	H	A water balance model is used to simulate runoff and the index is created by developed by normalizing and standardizing the generarted runoff to the mean runoff
Vegetation outlook (VegOut),	Aggregate	It combines the cliamte information and RS data of

(Tadesse and Wardlow, 2007)		current vegetation conditions with oceanic index data and environmental information including land cover type, irrigation status, soils and ecological settings to predict future outlook of vegetation condition
Ration Vegetation Index (RVI), (Pearson and Miller, 1972)	RS	The index is computed from the relation: $RVI = \frac{NIR}{R}$ where the NIR is near infra red and R is red bands repsectively
Weighted Diffrence Vegetation Index (WDVI), (Clevers, 1988)	RS	The index is a function of near infra-red (NIR), rainfall (R) and slope of soil line γ as expressed in the relation: $WDVI = NIR - \gamma R$
Perpendicular Vegetation Index (PVI), (Richardson and Wiegand, 1977)	RS	this index is calculated for the function: $PVI = Sin(a)NIR - Cos(a)R$ where a is the angle between the soil line and the NIR
Difference Vegetation Index (DVI), (Lillesand and Kiefer 1987; Ray, 1994)	RS	The index is a function of near infra red band and the rainfall computed from the relation: $DVI = NIR - R$
Soil Adjusted Vegetation Index (SAVI), (Huete, 1988)	RS	It is a function of near infrared band, red band and soil adjustiment fcator as per the expression: $SAVI = \frac{NIR - R(1+L)}{(NIR + R + L)}$ where L is the soil adjustment factor to account for soil variation
Infrared Percentage Vegetation Index (IPVI), (Crippen, 1990)	RS	It is calculated from the following function: $IPVI = \frac{NIR}{NIR + R} = \frac{NDVI + 1}{2}$
Atmospherically Resistant Vegetation Index (ARVI), (Kaufman and Tanre, 1992)	RS	It based on atmoshperic variables as per the relatiion: $RB = R - \gamma(B - R) = \frac{NIR - RB}{NIR + RB}$ where γ is a correlation parameter which is optimum at γ =1
Anormally Vegetation Index (AVI)/Chen *et al*., 1994	RS	It uses the annual ndvi to assess vegetation dynamics and surface dryness of land
Cubed Ratio Vegetation Index (CRVI), (Thenkabail *et al*., 1994)	RS	The $CRVI = \left(\frac{NIR}{MIR}\right)^3$ MIR is the landsat-5 thematic mapper mid infrared
Simple Ratio Water Index (SRWI), (Zarco-Tejada and Ustin, 2001)	RS	Uses NDVI applied in MODIS with feedback loops to minimize atmospheric and soil bias that is present in NDVI and other DIs
Vegetation Tempertaure Condition Index (VTCI), (Wang *et al*., 2006)	RS	Uses NDVI and LST and is computes from: $VTCI = \frac{LST_{NDVIi\max} - LST_{NDVIi}}{LST_{NDVIi\max} - LST_{NDVIi\min}}$ where $LST_{NDVIimax}$ and $LST_{NDVIimin}$ are the maximum and minimum land surface temperature of the pixels in the study region and lstndvii is the land surface temperature of the pixel

Table 15A: Examples of supervised ANN

Broad classification	Sub-class	Specific types
Feedforward Neural Network	Linear	▪ Hebbian NN ▪ Perceptron ▪ Adaline
	Mulilayer perceptron	▪ Back propagation ▪ Radial basis function
	Classification only	▪ Learning Vector Quantification ▪ Probability neural network (PNN)
	Regression only	▪ General Regression Neural Networsk
Feed back Neural Networks		▪ Back-propagation ▪ Elman network ▪ Recurrent back propagation ▪ Time delay neural network (TDNN) ▪ Real time neural network ▪ Hopfield neural network

Table 16A: Summary of part of the data plotting in matlab

Duration (months)		Dependent variables	
		(SPI Observed and SPI Forecasted)	Precipitation (mm)
x1=[1	x2=[1	y1=[-2.000000 -2.419800000	y2=[0
2	2	-1.804132587 -1.982741713	0.65
3	3	-0.5475678 -0.360177701	4.82
4	4	-0.5475678 -0.460177701	4.33
5	5	-0.695221696 -0.764048644	0.52
6	6	-1.84330607 -2.025793371	0.15
7	7	-1.954799828 -2.514832501	0.24
8	8	-1.927679724 -1.611852002	0.93
9	9	-1.719758932 -1.489001507	6.39
10	10	-0.074472665 -0.098184546	7.35
11	11	0.214808437 0.127360745	0.19
12	12	-1.942746449 -2.413507835	0.01
13	13	-1.996986655 -2.419468833	0.22
14	14	-1.933706414 -2.125143349	0.49
15	15	-1.852346104 -2.035728368	15.7
16	16	2.730951356 3.001315540	21.22
17	17	4.394317693 4.829355144	0.62
18	18	-1.813172622 -1.992676711	0
19	19	-2.000000000 -2.198000000	1.25
20	20	-1.623331898 -1.878404176	8.88
21	21	-1.623331898 -1.784041756	6.4
22	22	0.675850194 0.874275936	0.05
23	23	-0.07145932 -0.078533793	3.68
24	24	-1.984933276 -2.18144167	2.68
25	25	-0.891089109 -0.979306931	3.45
26	26	-1.19242359 -1.310473526	9.83
27	27	-0.96039604 -1.105547525	8.39
28	28	0.962117951 1.057367628	0.93
29	29	0.528196298 0.658048773	3.26
30	30	-1.71975893 -1.890015067	0.42
31	31	-1.01764959 -1.118396901	0.13
32	32	-1.87343951 -2.05891003	3.55
33	33	-1.960826517 -2.154948343	7.4
34	34	-0.930262591 -1.022358588	1.68
35	35	0.229875161 0.252632802	0.42
36	36	-1.493758071 -1.641640121	0
37	37	-1.873439518 -2.05891003	1.84
38	38	-2.00000000 -2.19800000	9.43
39	39	-1.445544554 -1.588653465	5.94
40]	40]	0.841584158 0.92490099]	0.03]

Table 17A: Annual average weight parameters and the non-exceedance probabilities of the SWSI

month	precipitation (mm)	Average monthly Weighting parameter for SWSI				Average Non-exceedance probability			
		a	*B*	*c*	*d*	PN_{rn}	PN_{sf}	PN_{rs}	PN_{df}
Jan	3.321	0.0325	0.1217	0.6721	0.1737	0.6298	0.9361	0.4784	0.3884
Feb	1.6674	0.0268	0.0893	0.7054	0.1785	0.6507	0.9558	0.4840	0.3142
Mar	0.1101	0.0649	0.0758	0.6834	0.1759	0.4096	0.9633	0.4901	0.2670
Apr	6.7602	0.1910	0.1327	0.5396	0.1366	0.1535	0.9129	0.4904	0.2128
May	7.1243	0.1014	0.2250	0.5274	0.1462	0.2762	0.8476	0.4801	0.0532
Jun	3.1200	0.0208	0.1542	0.6505	0.1745	0.5780	0.9194	0.5223	0.1171
Jul	0.1099	0.0198	0.1057	0.6908	0.1837	0.6231	0.9493	0.5330	0.2074
Aug	0.3250	0.0244	0.0886	0.7003	0.1867	0.6007	0.9583	0.5372	0.2607
Sep	0.0005	0.0214	0.0844	0.7050	0.1892	0.6862	0.9594	0.5291	0.3085
Oct	4.9018	0.0702	0.0789	0.6692	0.1817	0.4543	0.9597	0.4926	0.5107
Nov	10.2346	0.1000	0.1517	0.5897	0.1587	0.4035	0.9039	0.4887	0.5585
Dec	8.1245	0.0375	0.1453	0.6441	0.1730	0.5688	0.9218	0.4745	0.6277

Key: a,b,c, d and PN_{rn}, PN_{sf}, PN_{rs} and PN_{df} are weight parameters and probability of non-exceedance for rainfall, stream flow, reservoir levels and dam inflows respectively

Table 18A: Monthly average streamflow of the gauged stations used in the study

Month	Average Streamflow (m^3/s) at Gauged stations					
	4AB05	4AD01	4BC02	Maragua (ID)	Thiba (ID)	4DA10)
Jan	1.5538		14.6725	9.9112	3.4482	10.1985
Feb	0.8015		13.4871	7.5241	3.0565	7.9876
Mar	0.7354		15.7390	7.3303	2.7166	7.7111
Apr	2.2727		45.4088	20.8813	6.1387	16.4643
May	2.7747		76.6992	29.7100	19.8789	43.2541
Jun	1.0557		37.9182	15.3258	12.4697	29.1793
Jul	0.7413		21.1989	9.5936	5.8657	15.7863
Aug	1.0998		17.7013	7.6427	4.4726	13.3713
Sep	0.9768		15.1034	6.5295	3.6266	11.3016
Oct	1.7069		19.9140	9.7869	6.0722	13.1231
Nov	2.9718		37.4919	18.8439	7.5825	18.5222
Dec	2.0621		26.1597	14.1239	5.2174	14.4553

Table 19A: Estimated of precipitation at a station Naro-moru (N-M) using inverse distance weighting (IDW) technique

Station point	P (mm) Jan-1980 precipitation	Δx	Δy	d^2= Δx2+ Δy2	$w_i = \frac{1}{d^2}$	Product =wxp
N-M	___	___	____	___	_____	__-_
A	3.90	25.40	27.26	1388.27	0.000072	0.018290
B	62.70	19.25	51.50	3022.81	0.000331	0.006368
C	7.01	51.00	32.84	3679.47	0.000271	0.013860
D	1.29	26.80	2.70	725.530	0.001378	0.036939
E	3.80	7.50	2.50	62.500	0.016000	0.120000
F	1.24	22.40	50.70	3072.250	0.000325	0.007290
Total					0.019027	0.202755

Key A=Kerugoya DWO, B=Nyeri, C=Embu, D=GP3, E=GP4 and F=GP9

$$P_{N-M} = \frac{\sum w_i \times p_i}{\sum w_i} = \frac{0.202755}{0.019027} = 10.6565 mm$$

Table 20A: Summary of the drought indices for the study

Drought index	Input variables	Drought type
SWSI	P,Q reservoir level, Reservoir Inflow, dam inflow	Hydrological
SDI	Q	Hydrological
SPI	P	Meteorological
EDI	P	Meteorological
PDSI	P, T, AWC	Agricultural
SMDI	SWC	Agricultural
Generic **NDI**	P,Q, reservoir level, sm	Hydrological, Meteorological, Agrcicultural

Table 21A: Drought severity in 1980 computed for upper Tana River basin

Station name	Station ID	Drought severity, S
Kamburu	4EB01	4.023
Yatta furrow	4CC03	4.651
Nyamindi	4DA10	3.581
Rupingazi	4DC03	3.248
Sagana	4AC03	2.354
Tana sagana	4BC02	2.288
Amboni	4AB05	1.041
Gura	4AD01	1.051

Table 22A: SWSI critical values of the performance measures extracted from Figure 22B

Cross sectional point	Graphs	Lead time (x)(months)	Performance value (y)	Values at R2=0.50	
					Threshold Value
P1	NSE and RMSE	12	0.59	Index	range (y)
P2	R2 and RMSE	13	0.62	R2	0 to 0.50
P3	d1 and RMSE	14	0.64	RMSE	0 to 0.72
P4	MAE and NSE	14	0.56	NSE	0.49 to 0.99
P5	MAE and R2	14	0.57	MAE	0 to 0.68
P6	MAE and d1	15	0.59	d1	0.52 to 1.00

Tabe 23A: Regression equations for filling data for defined stations

S.No	Regression equation	Correlation coefficient (R^2)
1	$P_E = 0.736P_{SFCF} + 7.2$	0.882
2	$P_{SFCF} = 1.2P_{Ke} - 33.9$	0.891
3	$P_{Ki} = 0.878P_{Ke} - 4.8$	0.908
4	$P_{Ke} = 1.58P_{Ke} - 14.4$	0.831

Key: P is precipitation (mm) , while subscripts E, SCFC Ki and Ke stands for embu, Sagana FCF, kerugoya and kirinyaga stations

Table 24A: Steps in formulation of $SWSI_m$ model for upper Tana River basin

i) By considering the regression equation for Kamburu (ID 4ED01) hydrometric station plotted for $SWSI_p$ against the normalized $SWSI_n$

$$y = 11.082x - 5.4096 \quad (1)$$

ii) Factoring 5.4096 outside bracket yields

$$y = 5.4096[2.04(x) - 1] \quad (2)$$

iii) Guided by the original equation which has a denominator, writing the value 5.4096 in form of a fraction shows that

$$5.4096 = \frac{1}{0.1849} \quad (3)$$

iv) Thus substituting the right hand side of Equation 3 in Equation (2) yields

$$y = \frac{[2.04(x) - 1]}{0.1849} \quad (4)$$

v) But from the SWSI the x is an expression defined by:

$$x = a \times PN_{rn} + b \times PN_{sf} + c \times PN_{rs} + d \times PN_{df} \quad (5)$$

vi) Substituting the right hand expression of equation (5) into equation (4) and y with SWSI yields:

$$SWSI = \frac{\lfloor 2.04(a \times PN_{rn} + b \times PN_{sf} + c \times PN_{rs} + d \times PN_{df}) - 1 \rfloor}{0.1849} \quad (6)$$

vii) Repeating this process for all the other stations gave an equation of the same form. Thus it is concluded that the general equation takes the form:

$$SWSI_m = \frac{\lfloor k_1 \times (a \times PN_{rn} + b \times PN_{sf} + c \times PN_{rs} + d \times PN_{df}) - k_2 \rfloor}{k_3} \quad (7)$$

Table 26A: The selected drought indices for different drought types as applied in the study

Drought index	Drought type for which the index is adopted
SWSI	Hydrological
EDI	Hydrological
SPI	Meteorological
EDI	Meteorological
SMDI	Agricultural
PDSI*	Agricultural

*Index could also be used for hydrological drought

APPENDIX B: FIGURES

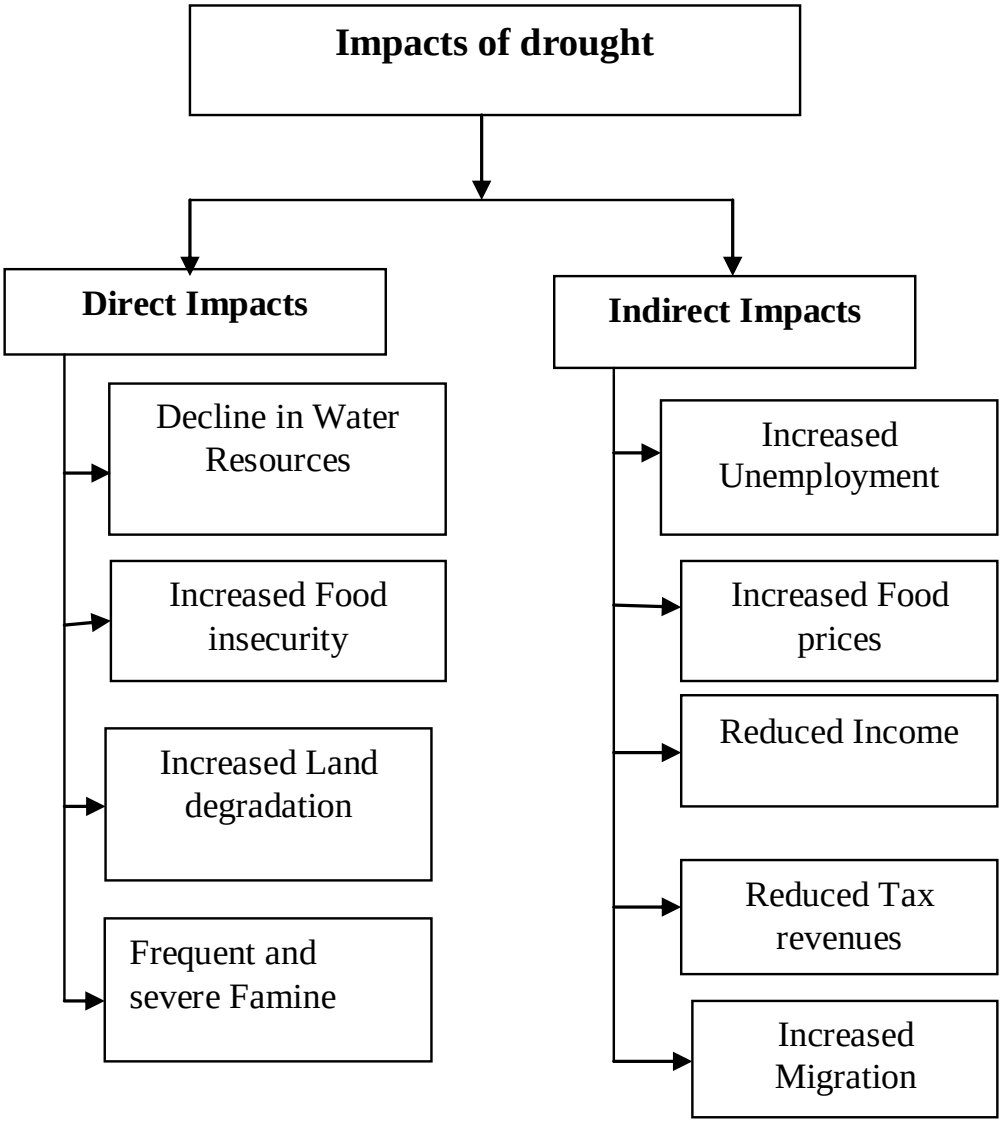


Figure 1B: Categories of impacts of drought

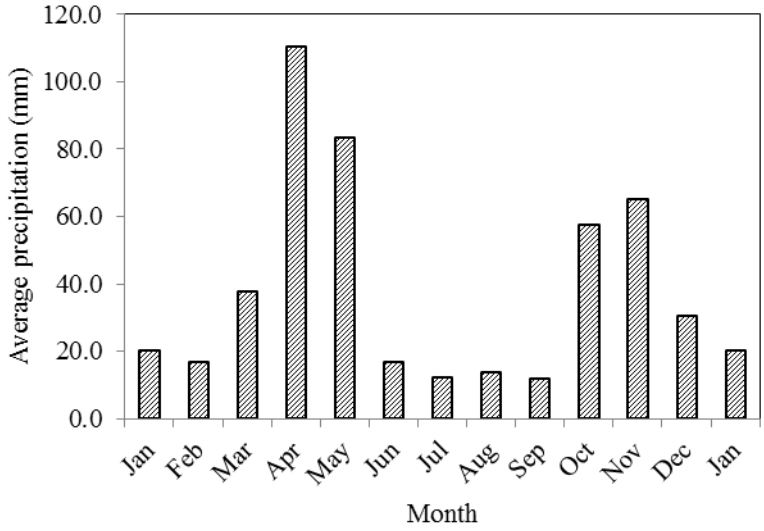


Figure 2B: Bimodal distribution of the precipitation for upper Tana River basin

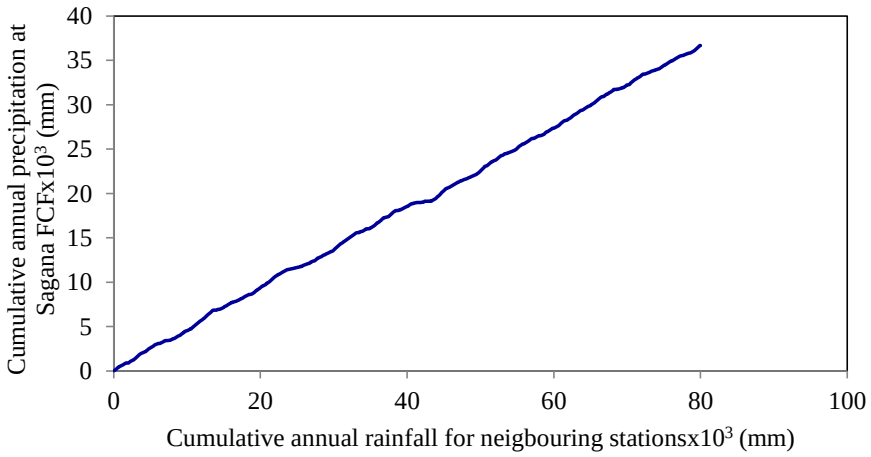


Figure 3B: Double mass curve based on precipitation upper Tana River basin

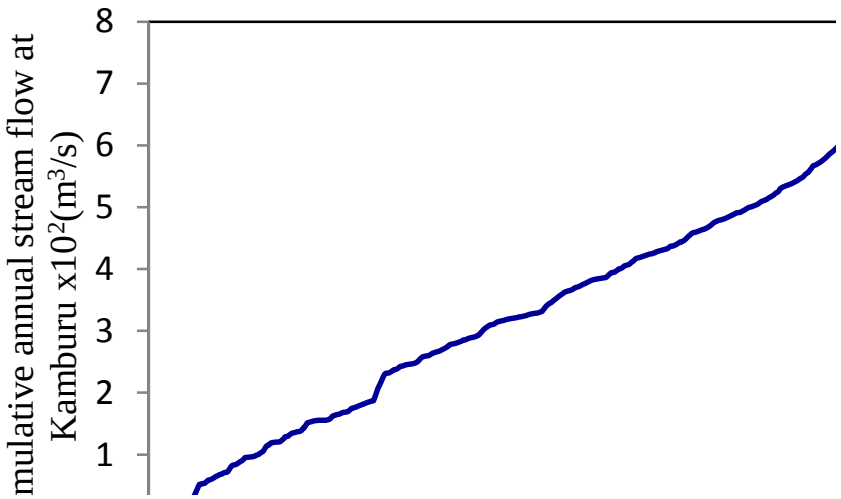


Figure 4B: Double mass curve based on strem flow for upper Tana River basin

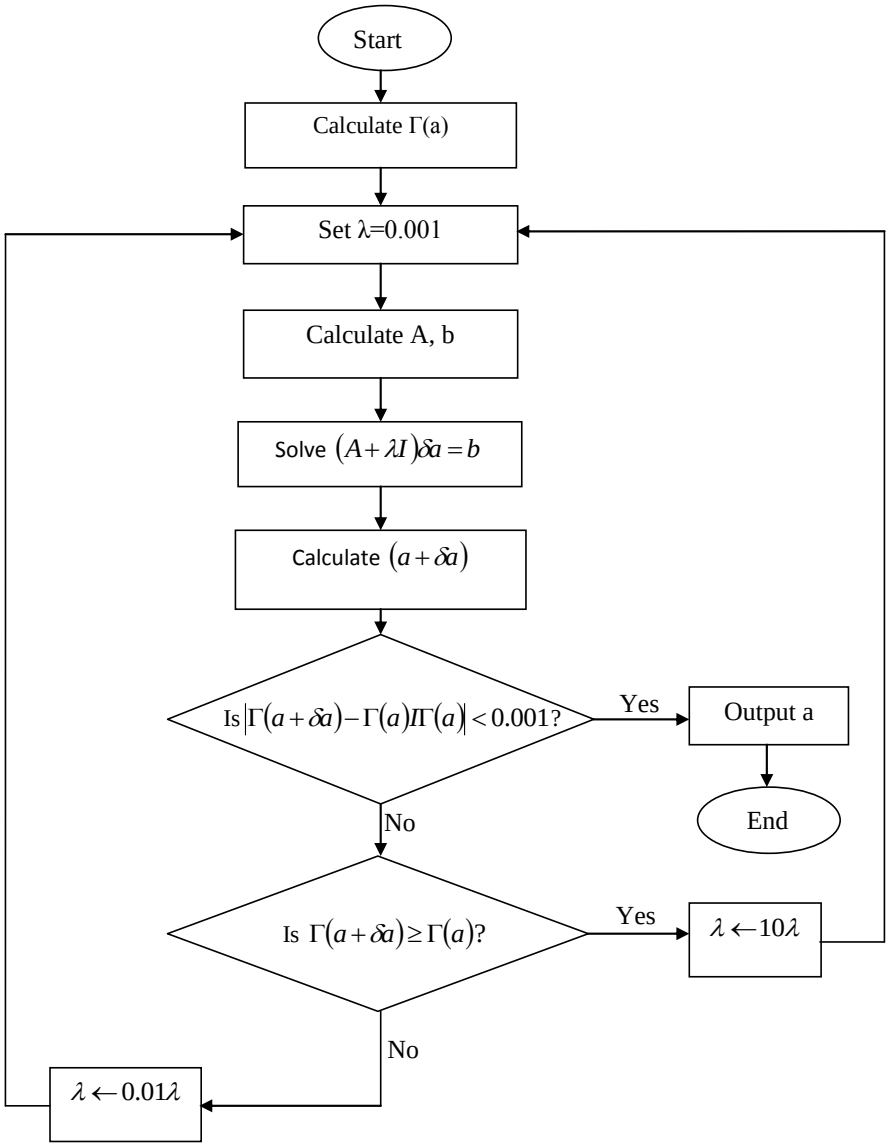


Figure 5B: Flow chart of Levernberg-Marquardt algorithm (Source Toushmalani, 2013)

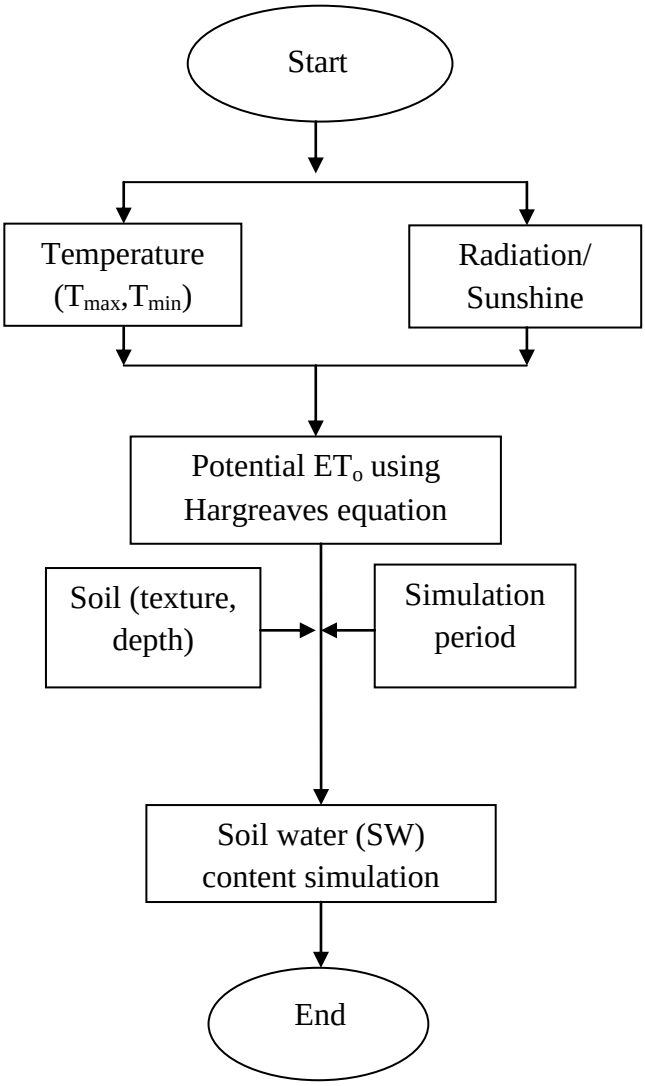


Figure 6B: Flow chart showing the steps used in soil water (SW) simulation in Aquacrop model

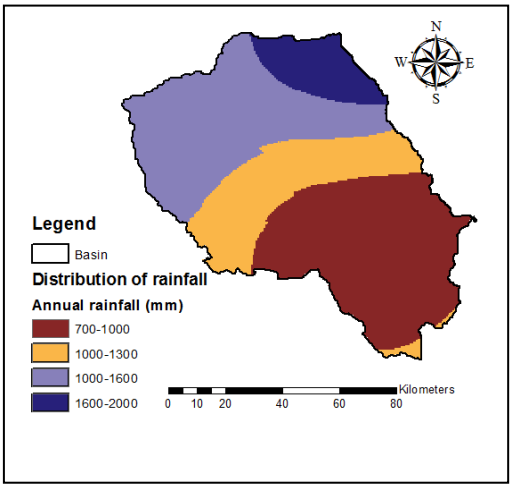


Figure 7B: spatial distribution of annual precipitation in the upper Tana River basin

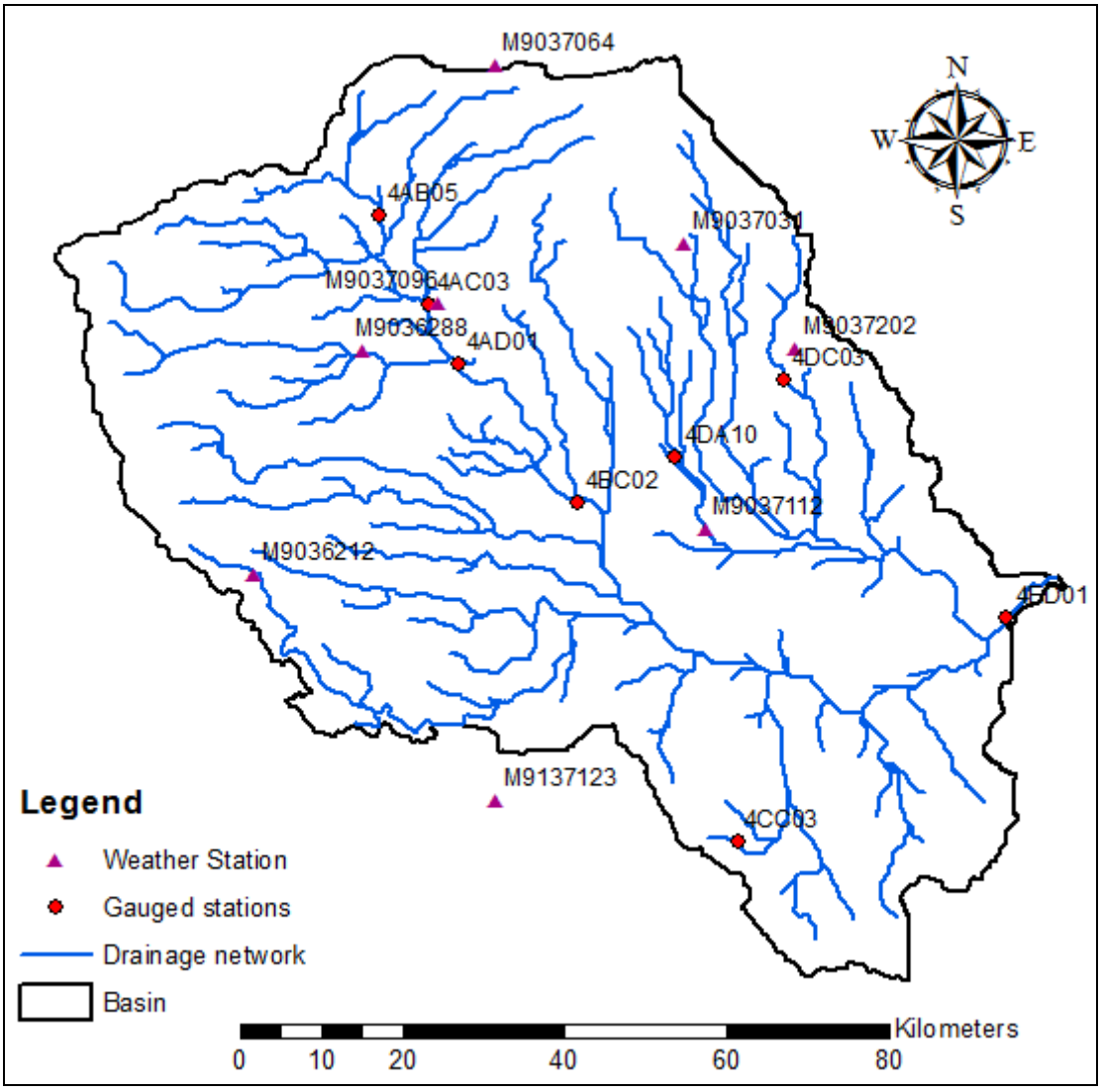


Figure 8B: the location of hydro-metric stations in upper Tana River Basin based on IDs

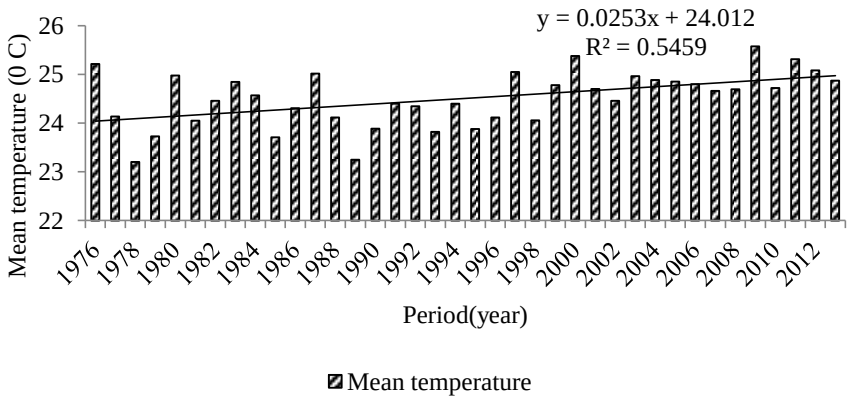


Figure 9B: Mean monthly temperature at Embu hydrometric station

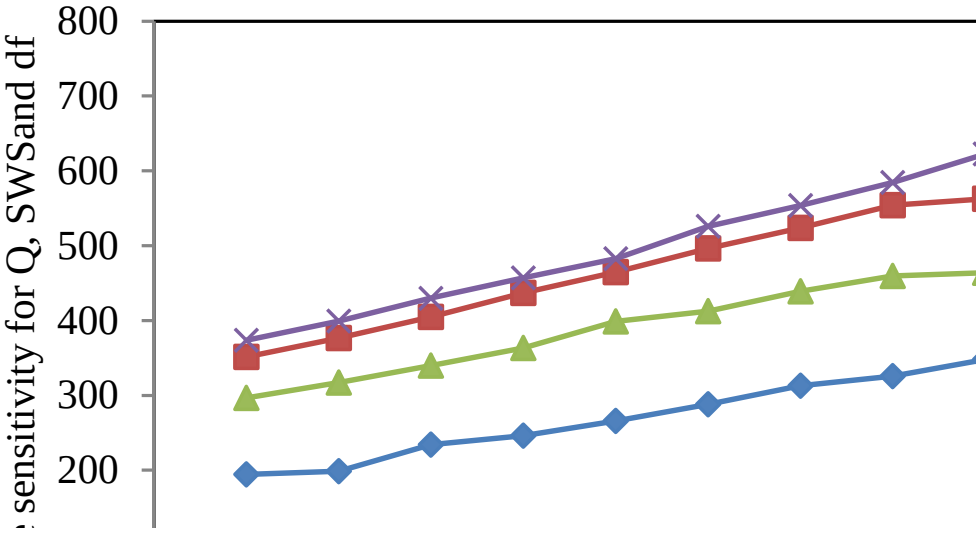


Figure 10B: Absolute sensitivity of NDI for a decrease in parameter values at MIAD

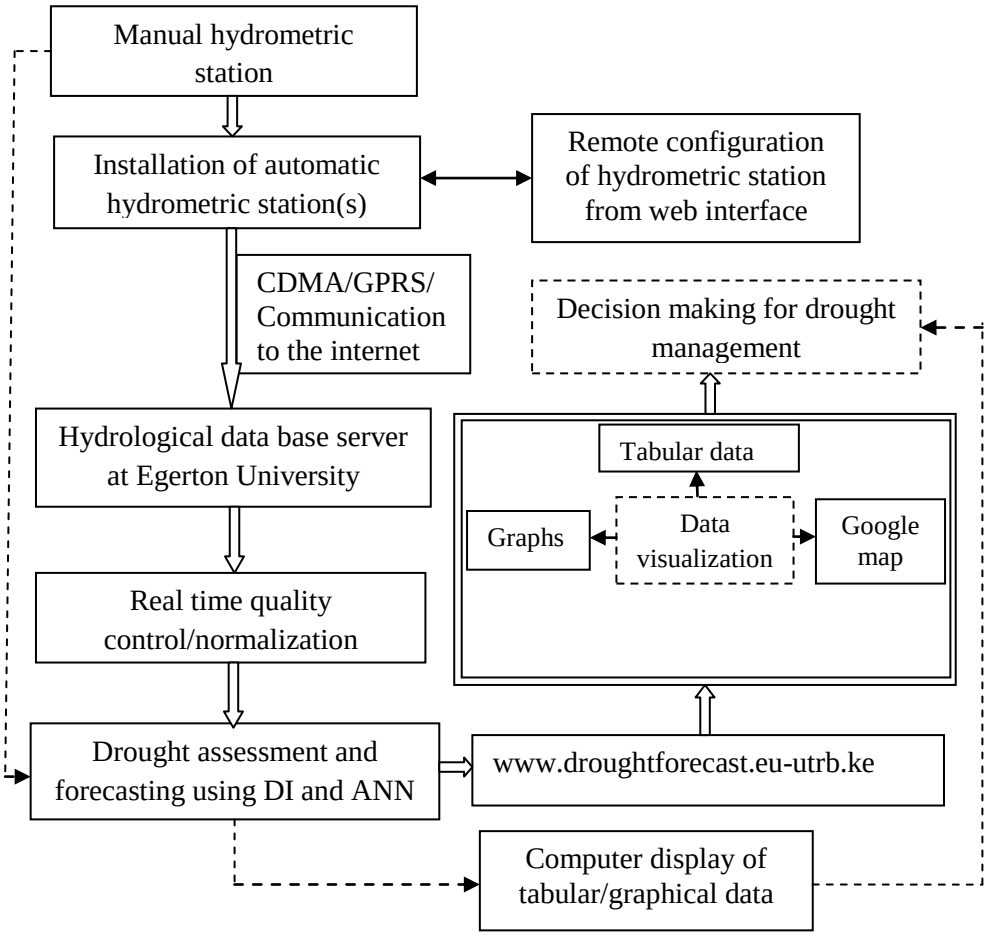


Figure 11B: Telemetric data transmission for drought forecasting and management

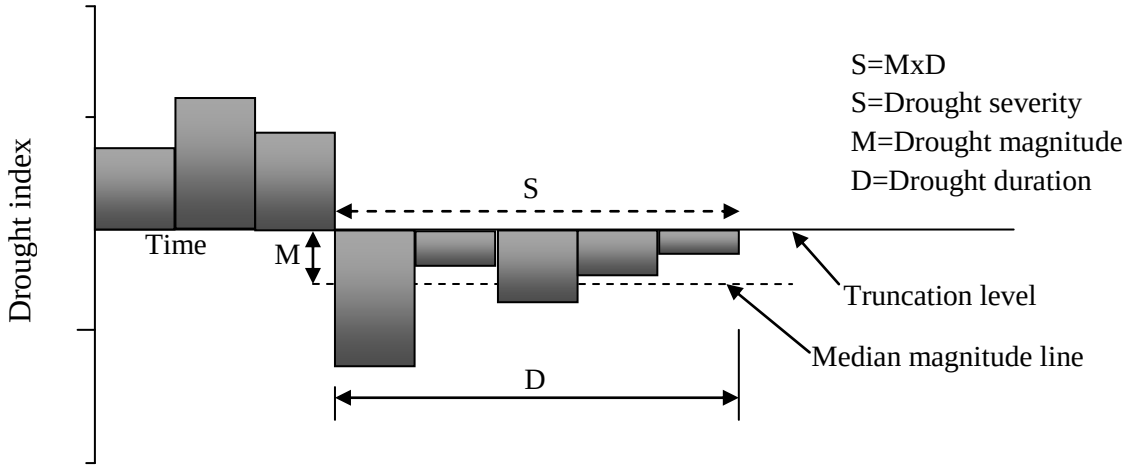


Figure 12B: Relationship between the magnitude, duration and severity of a drought event

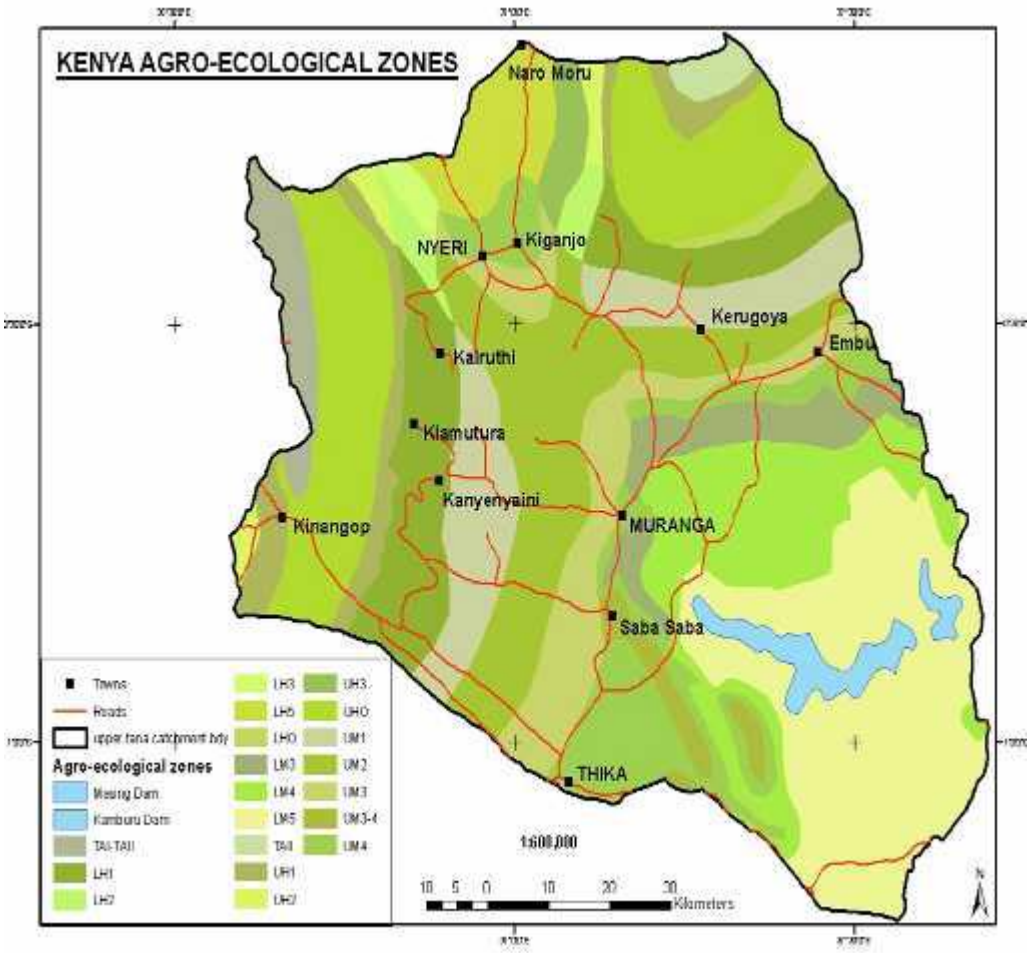


Figure 13B: agro-ecological zones in the upper Tana River basin (Source; GoK, 2012)

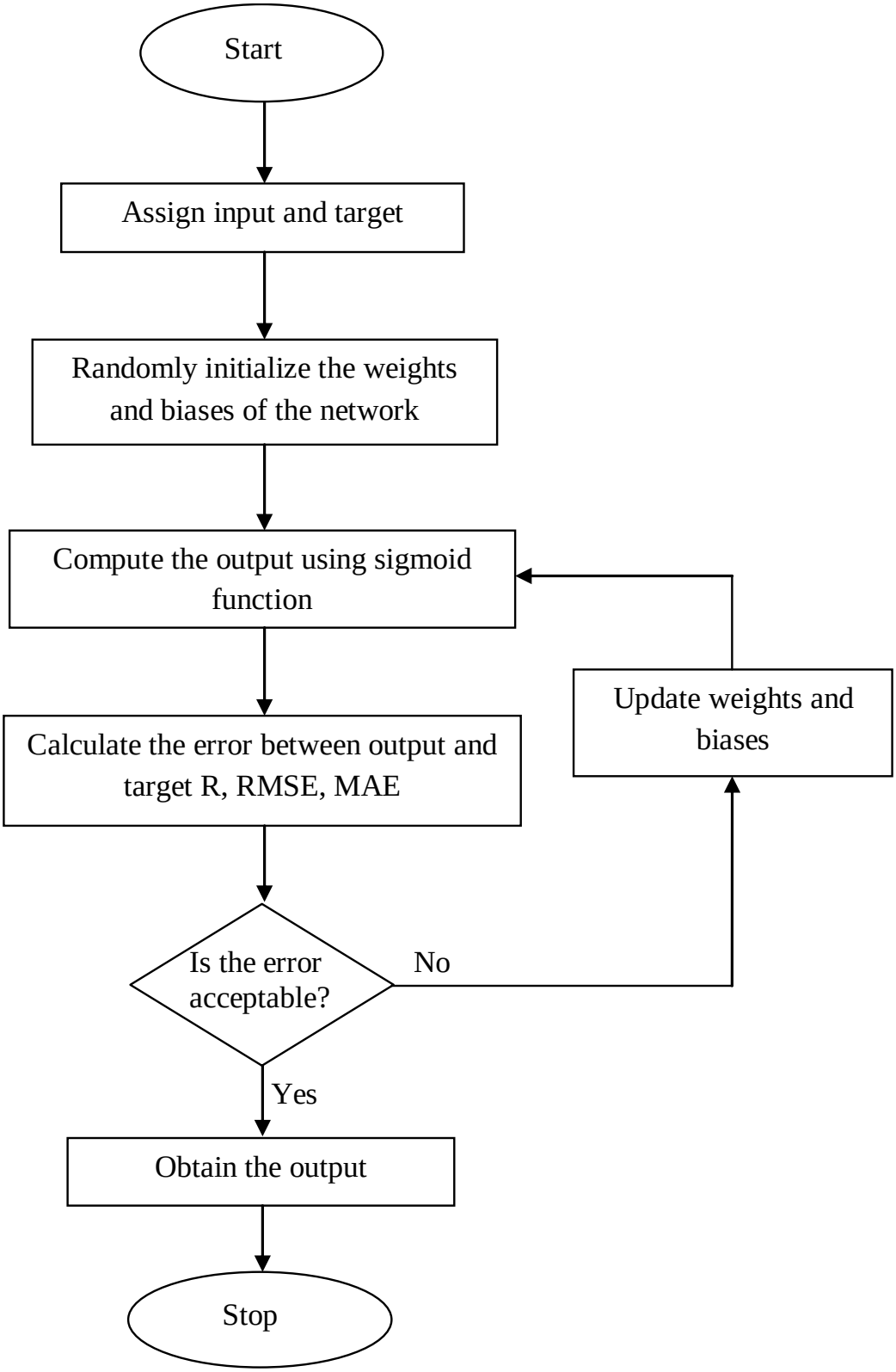


Figure 14B: Flow chart for back-propagation training

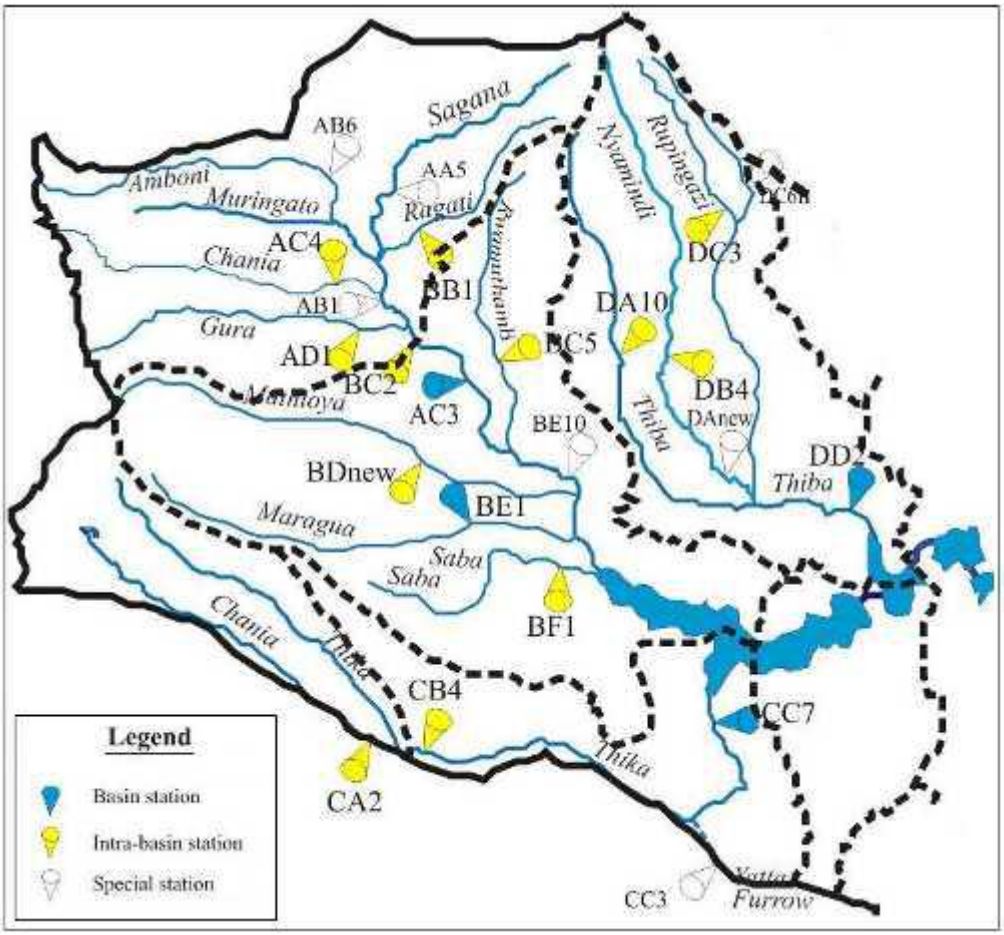


Figure 15B: The sub-basins in the upper Tana River basin (Source,;GoK, 2012)

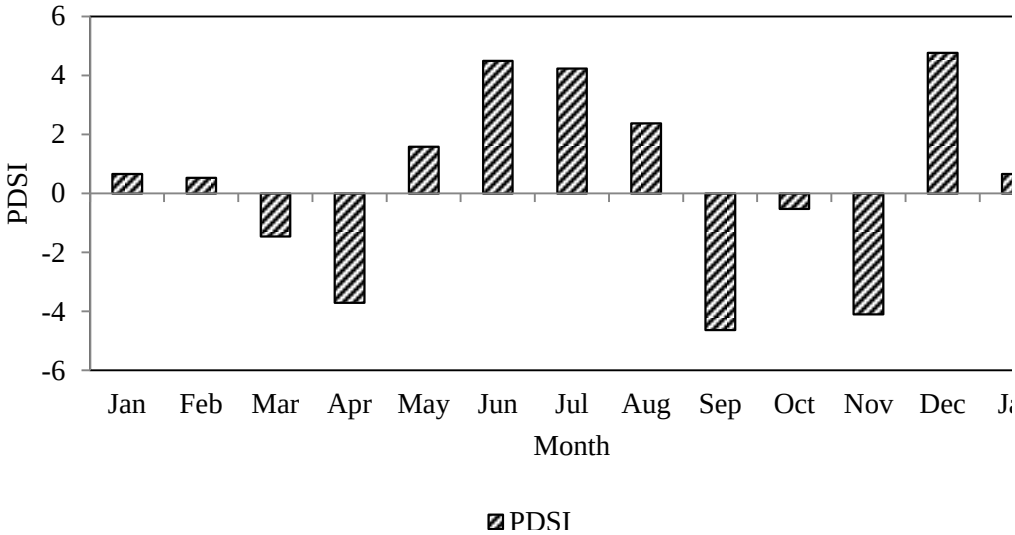


Figure 16B: Monthly average PDSI at Amboni hydrometric station

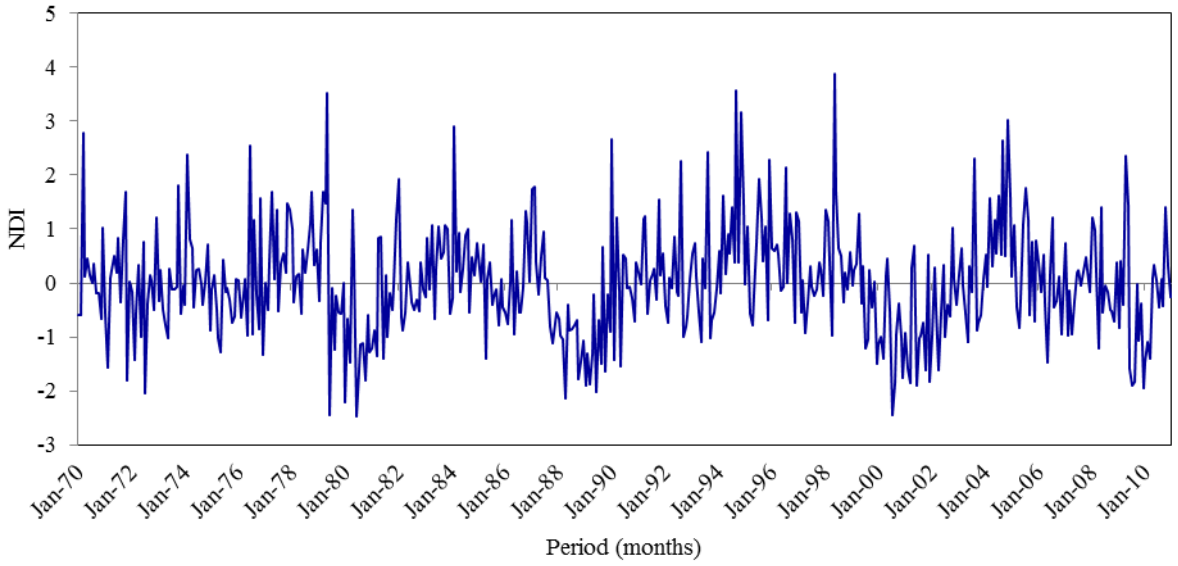


Figure 17B: NDI time series for Sagana F. C. F. station

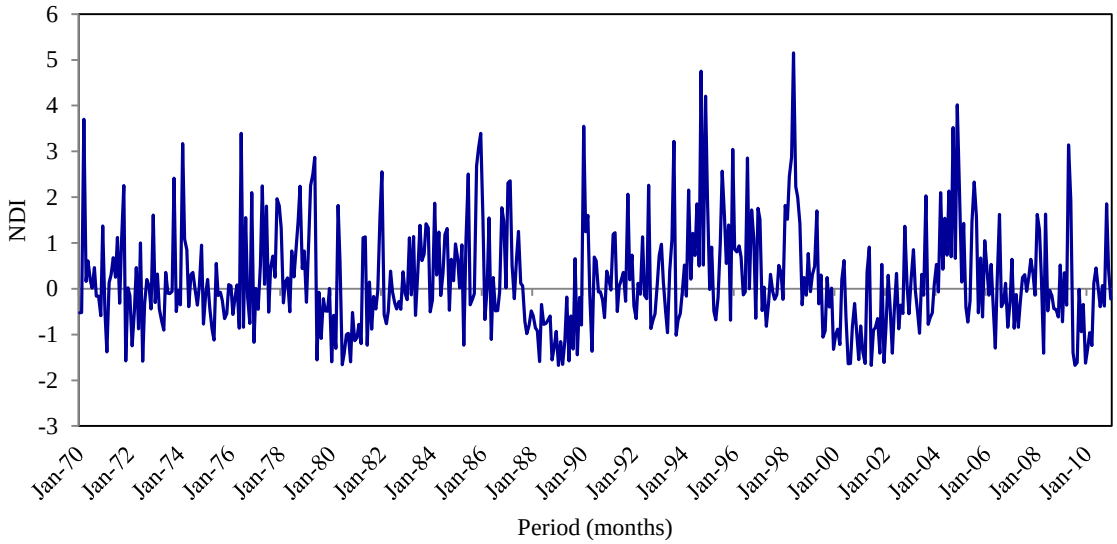


Figure 18B: NDI time series for Naro-moru F.G.P. station

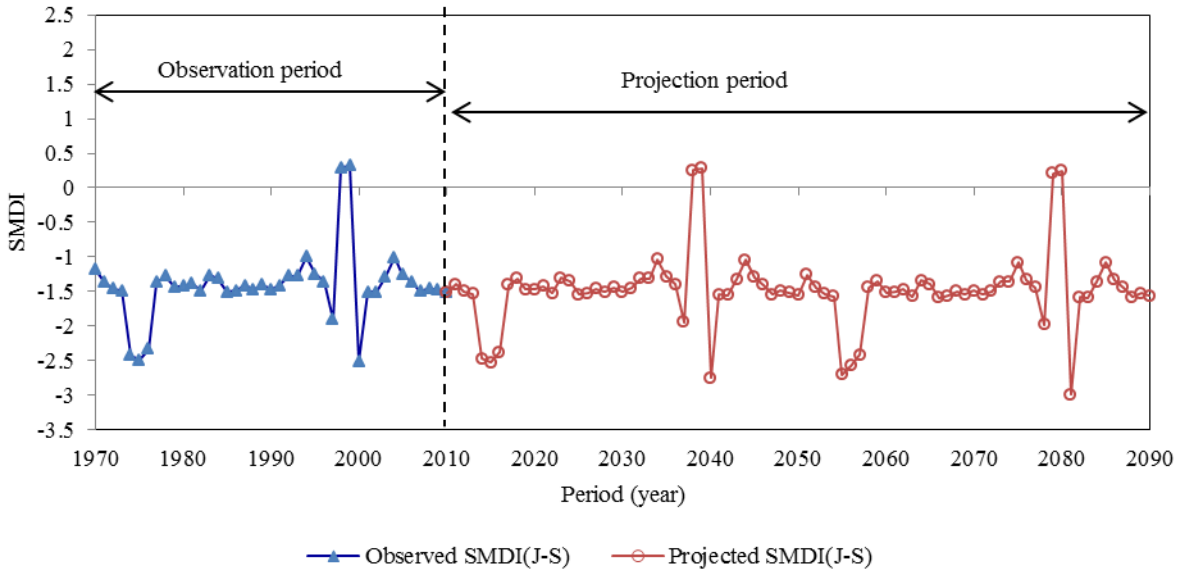


Figure 19B: observed and projected SMDI at Naro-moru F.G.P. station

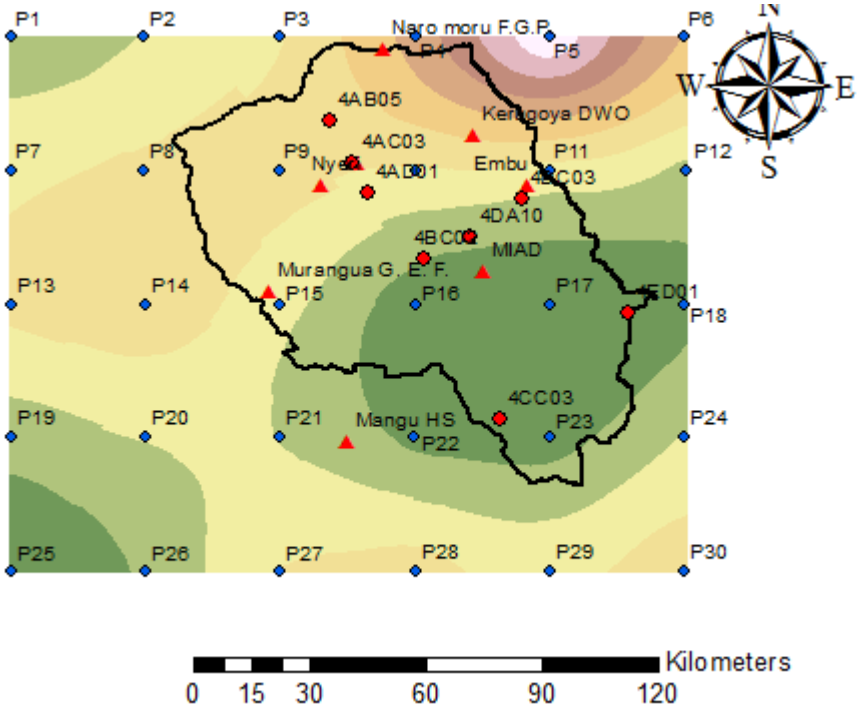


Figure 20B: The location of hydro-metric and global data stations in upper Tana River basin

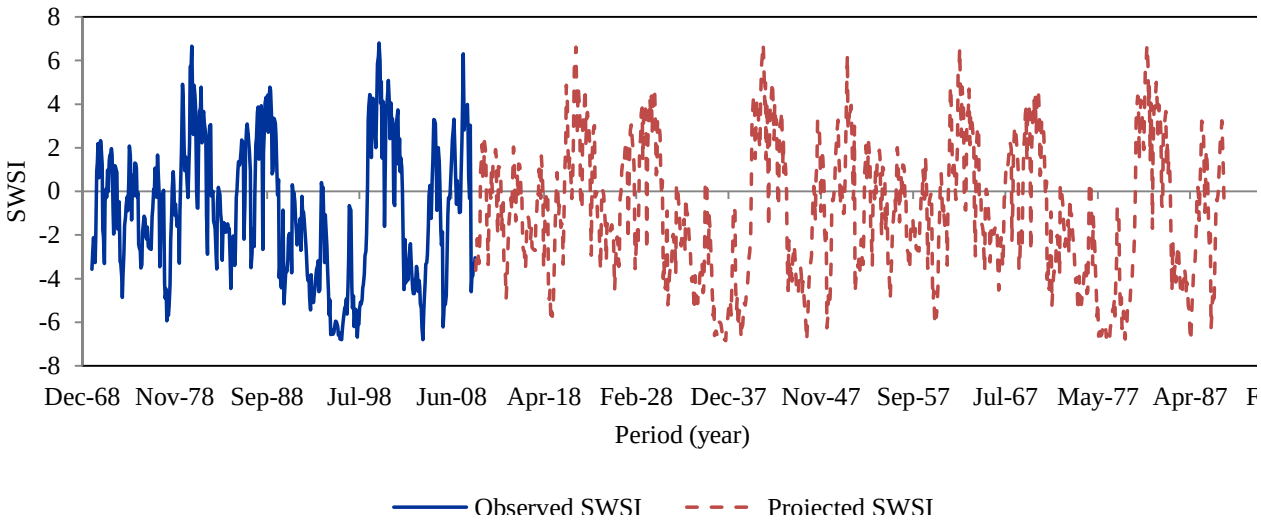


Figure 21B: Observed and projected SWSI at Amboni hydrometetric station

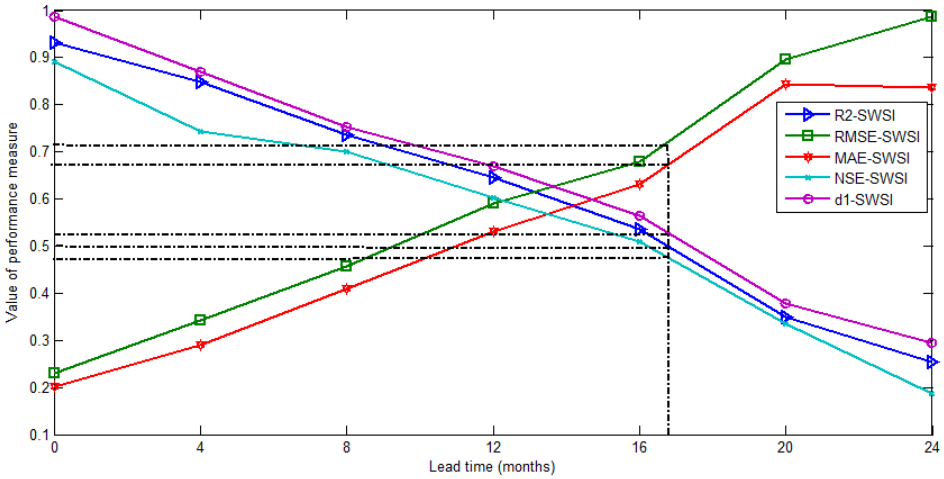


Figure 22B: Comparison of SWSI forecasts at different lead times

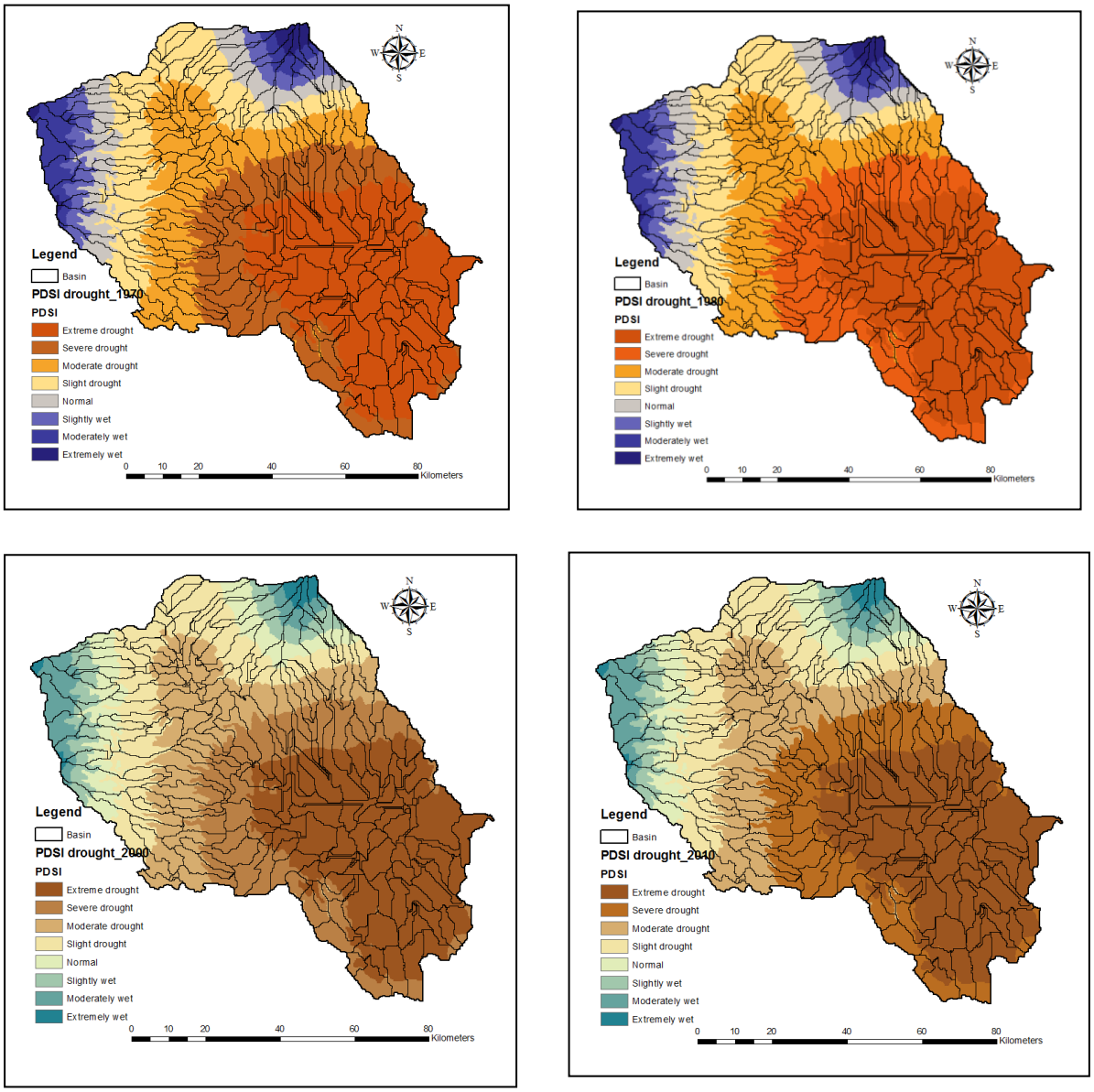


Figure 23B: The spatially distributed drought based on PDSI for upper Tana river basin

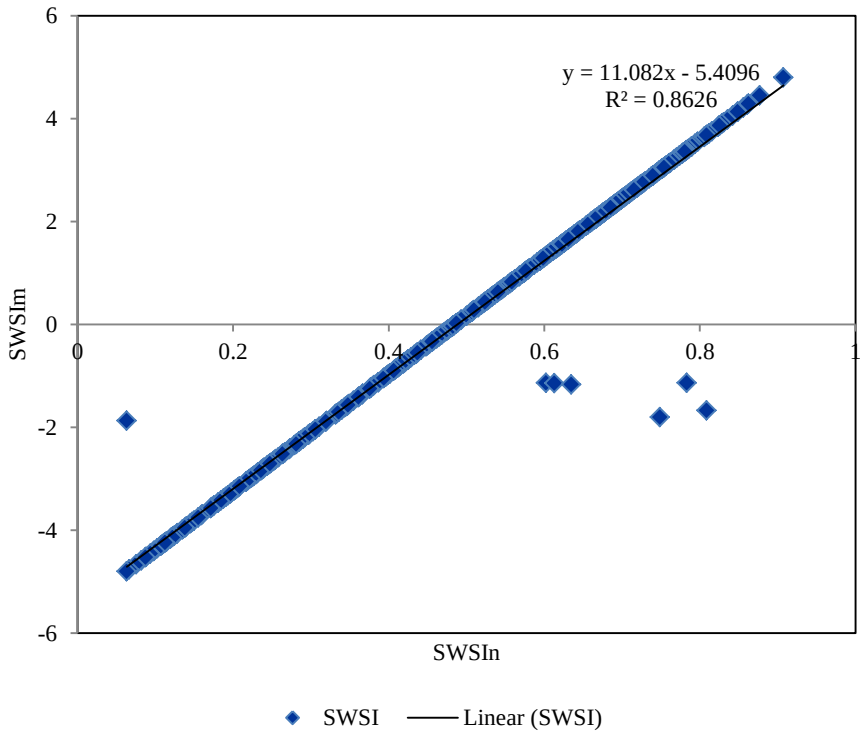


Figure 24B: Regression of SWSn and SWSIm for kamburu hydrometric station

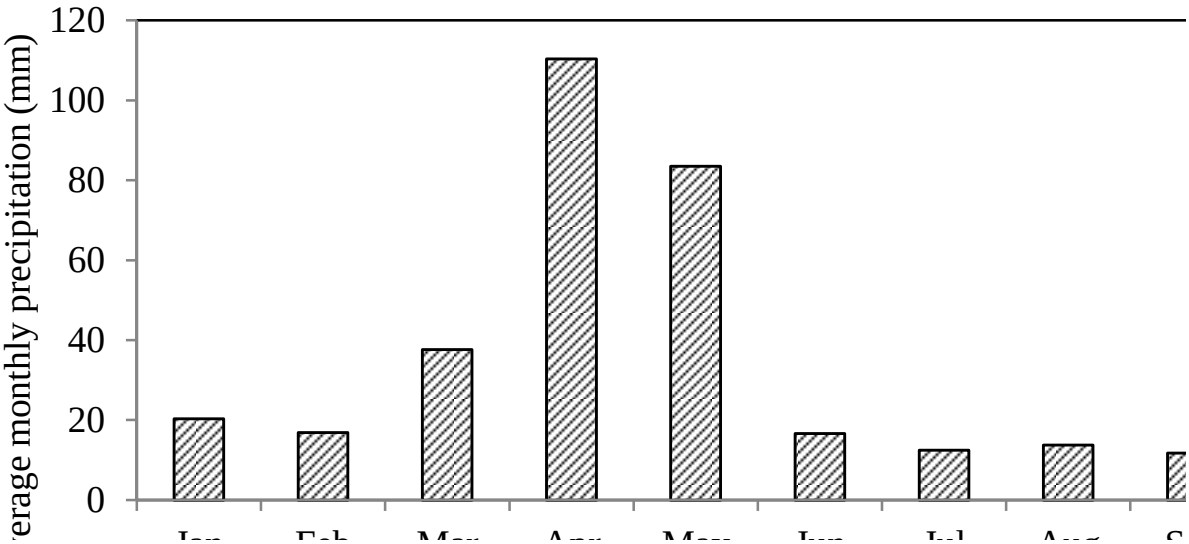


Figure 25B: Average monlthly precipitation for upper Tana River basin

H_{Jan}=

R	rs	df	swc
2.54	1055.44	695.95	2.986
1.52	1051.37	704.68	4.505
13.21	1052.27	506.36	5.691
0.76	1055.46	757.94	4.511
1.52	1051.27	694.46	5.533
0.76	1055.16	663.40	5.855
8.13	1055.26	648.50	6.248
17.27	1051.37	758.36	6.566
64.77	1056.66	1236.06	6.436
41.91	1039.40	1487.95	6.237
4.06	1020.33	1497.94	4.869
7.37	1036.40	771.96	5.351
156.46	1045.48	723.40	8.632
0.76	1054.27	739.10	5.300
0.76	1054.52	676.95	5.193
42.67	1055.01	735.01	6.156
3.30	1046.16	858.28	6.105
106.43	1055.78	734.40	7.999
0.00	1045.73	561.50	6.074
0.00	1024.37	697.29	6.074
136.91	1054.69	1006.65	11.083
17.27	1054.76	756.01	6.450
0.76	1054.96	547.01	4.276
32.77	1055.15	697.10	7.177
32.77	1056.34	482.60	7.177
43.18	1054.61	685.70	3.434
38.10	1056.58	1542.8	6.339
55.88	1055.96	1359.4	8.073
72.39	1055.42	920.90	8.531
4.32	1056.16	821.28	5.562
0.00	1038.26	738.10	5.752
50.80	1026.13	759.50	7.119
0.00	1040.54	702.50	6.020
32.34	1051.04	736.50	6.451
32.34	1055.14	649.50	7.163
42.62	1056.65	711.50	3.473
37.60	1053.79	668.50	6.335
55.15	1050.53	730.70	8.047
71.45	1053.24	911.10	8.498
4.26	1042.09	1339.90	5.569
0.00	1033.42	702.27	5.756

Figure 26B: Observed data matrix for January at Sagana FCF hyrometric station
(Columns 1, 2, 3 and 4=monthly average rainfall, R(mm), reservoir level, rs(m asl), dam inflow df, (MCM) and Soil-water content (SWC) (cm)

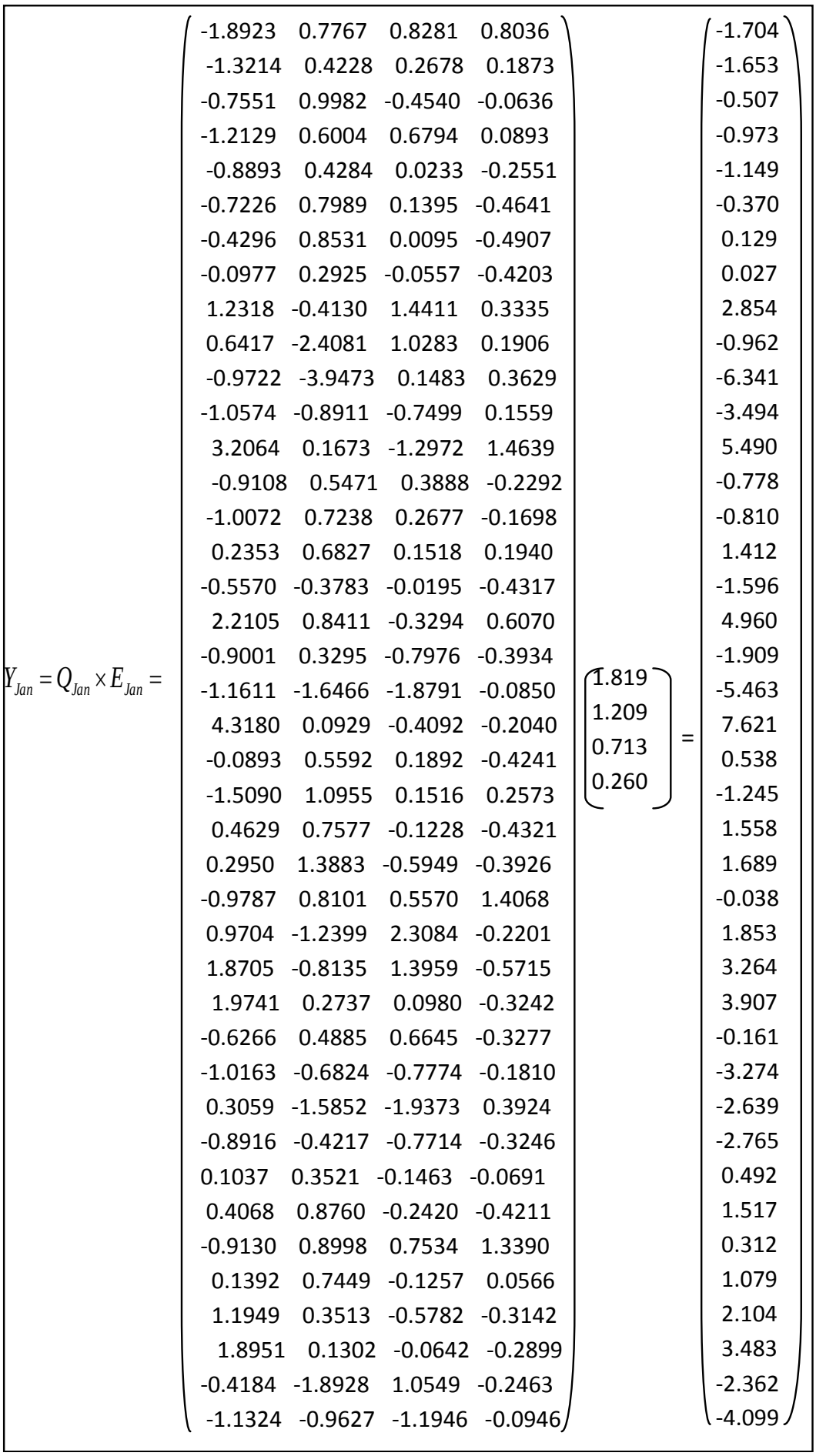


Figure 27B: PC1 for the month on January at Sagana FCF hydrometric station

$$NDI_{Jan} = \frac{Y_{Jan}}{\sigma_{Jan}} = \frac{1}{2.857} \begin{pmatrix} -1.7040 \\ -1.6527 \\ -0.5065 \\ -0.9730 \\ -1.1492 \\ -0.3695 \\ 0.1294 \\ 0.0271 \\ 2.8543 \\ -0.9622 \\ -6.3406 \\ -3.4943 \\ 5.4899 \\ -0.7776 \\ -0.8101 \\ 1.4119 \\ -1.5964 \\ 4.9602 \\ -1.9091 \\ -5.4634 \\ 7.6212 \\ 0.5384 \\ -1.2452 \\ 1.5582 \\ 1.6892 \\ -0.0382 \\ 1.8534 \\ 3.2644 \\ 3.9069 \\ -0.1607 \\ -3.2743 \\ -2.6386 \\ -2.7652 \\ 0.4920 \\ 1.5172 \\ 0.3119 \\ 1.0788 \\ 2.1044 \\ 3.4831 \\ -2.3617 \\ -4.0992 \end{pmatrix} = \begin{pmatrix} -0.5964 \\ -0.5785 \\ -0.1773 \\ -0.3406 \\ -0.4022 \\ -0.1293 \\ 0.0453 \\ 0.0095 \\ 0.9991 \\ -0.3368 \\ -2.2193 \\ -1.2231 \\ 1.9215 \\ -0.2722 \\ -0.2836 \\ 0.4942 \\ -0.5588 \\ 1.7361 \\ -0.6682 \\ -1.9123 \\ 2.6676 \\ 0.1884 \\ -0.4358 \\ 0.5454 \\ 0.5912 \\ -0.0134 \\ 0.6487 \\ 1.1426 \\ 1.3675 \\ -0.0562 \\ -1.1461 \\ -0.9235 \\ -0.9679 \\ 0.1722 \\ 0.5311 \\ 0.1092 \\ 0.3776 \\ 0.7366 \\ 1.2192 \\ -0.8266 \\ -1.4348 \end{pmatrix}$$

Figure 28B: NDI values for January at Sagana FCF hydrometric station

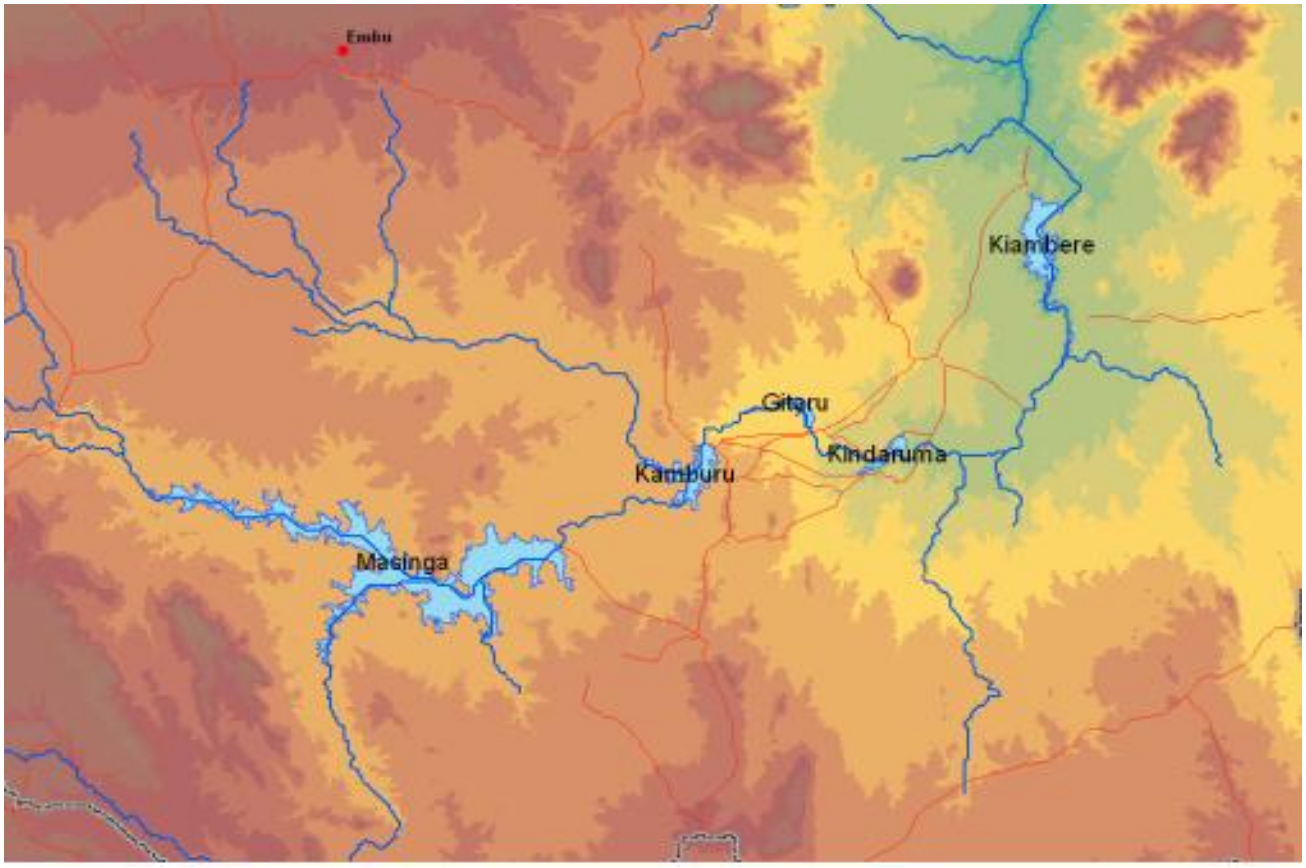


Figure 29B: The location of hydropower stations in the upper Tana river basin (Droogers, *et al*., 2006**)**

Appendix C: List of selected published articles

S.No	Description of the article
1.	Wambua, R. M., Mutua, B. M. and Raude, J. M. (2014). Drought Forecasting Using Indices and Artificial Neural Networks for Upper Tana River Basin, Kenya-A Review Concept, *Journal of Civil & Environmental Engineering*, 4(4): 1-12.
2.	Wambua, R. M., Mutua, B. M. (2014). Drought Indices Assessment For Sustainable Water Resources Management-A Case Review For Upper Tana River Basin, Kenya, *International Journal of Engineering Research and Technology (IJERT)*, 3(10): 429-438.
3.	Wambua, R. M., Mutua, B. M. and Raude, J. M. (2014). Performance of Standardized Precipitation Index (SPI) and Effective Drought Index (EDI) in Drought forecasting using Artificial Neural Networks (ANNs) for upper Tana River basin, Kenya, *International Journal of Engineering Research and Technology (IJERT)*, 3(11): 547-556.
4.	Wambua, R. M., Mutua, B. M. and Raude, J. M. (2015). Stochastic drought forecasting exploration for water resources management in the upper Tana River Basin, Kenya, Handbook of research on computational simulation and modelling in Engineering, Section 3; Civil and Environmental Engineering, Chapter 17: 508-539.
5.	Wambua, R. M., Mutua, B. M. and Raude, J. M. **(**2015). Detection of drought-risk areas using trends of spatial and temporal drought characteristics in upper Tana River basin, Kenya, *Hydrological sciences journal*, submitted 2015; *Under review*
6.	Wambua, R. M., Mutua, B. M. and Raude, J. M. (2015). Spatio-temporal Drought Characterization for the Upper Tana River Basin, Kenya Using Standardized Precipitation Index (SPI), *World Journal of Environmental Engineering, 3(4):111-120.*
7.	Wambua, R. M., Mutua, B. M. and Raude, J. M. (2015). Hydrological drought frequency estimation using stream flow drought index and modified Gumbel method in upper Tana River basin, *International Journal of current research and review*, 7(22): 42-51.
8.	Wambua, R. M., Mutua, B. M. and Raude, J. M. (2016) Prediction of Missing Hydro-Meteorological Data Series using Artificial Neural Networks (ANN) for Upper Tana River basin, Kenya, *American Journal of Water Resources,* 4(2): 35-43.