UNIVERSITY OF READING

Department of Geography & Environmental Science
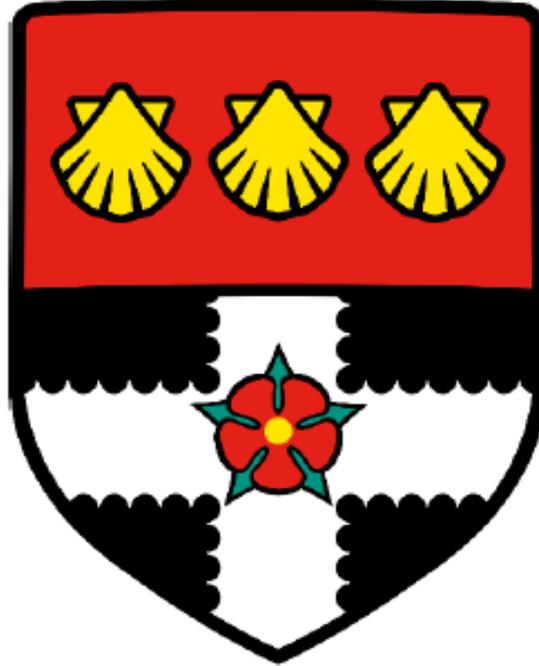
Improving flood modelling and forecasting in Kenya

Maureen Anyango Wanzala

Declaration

I can confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Maureen Anyango Wanzala

Abstract

The frequency of floods is rising and constitutes one of the main causes of detrimental consequences arising from natural disasters, not only in Kenya, but across the globe. The anticipation, forecasting and accumulation of science-based evidence of floods is a vital component in managing, preparing for and mitigating the impacts of severe events, from local to international scales. This research aims to explore ways to improve flood modelling and forecasting at the national scale. To achieve this, a multi-stage approach is adopted, first, to understanding the key aspects to consider when selecting an appropriate model for flood forecasting and modelling, secondly, to understanding the application of global reanalysis precipitation datasets to hydrological modelling as potential alternatives due to data scarcity challenges, and thirdly, to undertaking an analysis of trend detection in floods and possible shifts in flood timing.

Five criteria are applied to a hydrological model selection framework following a filter sequence; an evaluation of twelve potential models is performed, and four potential model candidates are selected for flood applications for a Kenyan national forecasting centre. Model selection has shown that not all models are good at capturing and/or representing the important processes relevant to flood generation and a single model would not be applicable to the entire country, due to stark differences in the hydroclimatic characteristics of catchments, and model developments and upgrades should allow incorporation of such differing characteristics. Four reanalysis precipitation datasets are evaluated for their ability to be used for hydrological modelling. The choice of precipitation input is found to be the dominant component of the hydrometeorological modelling chain, creating the need to aggregate both sensitivity indices and performance statistics. Improvements have arisen from the introduction of ERA5 as a source of meteorological data. Performance varies by season and catchment, with wetland catchments obtaining relatively better scores compared to those in the semi-arid regions. Examination of trends in river flow series identified statistically increasing trends in annual floods for stations in proximity to each other, which is evidence of a spatially coherent pattern. and increasing flood frequency across Kenyan catchments, with observations showing a shift in timing and variability in flood occurrences in most parts of the country.

This research has explored and provided an enhanced understanding of the avenues of improving flood modelling and forecasting in Kenya in terms of models, data, and historical flooding trends as well as seasonality and shifts in flood timing, which can inform future developments and operational flood forecasting for the end-users of an early warning system that can help mitigate the effects of floods in data-scarce regions such as Kenya.

Dedication

This PhD thesis is dedicated

- To my lovely *daughter – Amor Racheal Tamara* for her endurance and sacrifice of her playtime whilst I was busy working on my research every day of the week up to and including weekends.

- To my M*um – Elizabeth, Dad – Evans* and my *sister – Lydia* for your sacrifices and selfless support to me in my time of need.

- To the *African women in STEM – "I believe that fortitude is key. More than anything, be consistent. Go at it. Go at it. Go at it. When you succeed, don't forget the responsibility of making someone else succeed with you."* Antonia Novello.

Acknowledgements

Table of contents

List of figures

xi

List of tables

List of abbreviations

AFDM	African Flood and Monitor
ACMAD	Africa Centre for Meteorological Applications for Development
ANN	Artificial Neural Network
AMV	Atmospheric Motion Vector
AMAX	Annual MAXimum daily mean river flow
ASALs	Arid and Semi-Arid Lands
CFSR	Climate Forecast System Reanalysis
CHC	Climate Hazard Centre
CoF	Coefficient of Efficiency
COVID-19	Corona Virus Disease of 2019
CCD	Cold Cloud Duration
CC	Correlation Coefficient
CSC	Commonwealth Scholarships Commission
CHIRPS	Climate Hazards Group Infrared Precipitation with Station
CSR	Clear-Sky Radiance
DHI	Danish Hydraulic Institute
DEM	Digital Elevation Model
EA	East Africa
ENSO	El Niño-Southern Oscillation
EM-DAT	Emergency Events Database
EnKF	Ensemble Kalman Filter
ECVs	Essential Climate Variables
EFAS	European Flood Awareness System
ECMWF	European Centre for Medium-range Weather Forecasts
EVOFLOOD	Evolution of Global Flood Risk
FP	Factor Prioritization
FF	Factor Fixing
PE	Potential Evapotranspiration
POT	Peak-over threshold
EWS	Early Warning Systems
FATHUM	Forecasting for AnTicipatory HUManitarian action
FEWSNET	Famine Early Warning Systems Network
FFC	Flood Forecasting Centre
FFS	Flood Forecasting System
FEWS	Flood Early Warning Systems
GFFMS	Galway River Flow Forecasting and Modelling System
GFFS	Galway Flow Forecasting System
GIS	Geographical Information System

GeoSFM	Geospatial Streamflow simulation model
GA	Genetic Algorithm
GCM	General Circulation Model
GLUE	Generalized Likelihood Uncertainty Estimation
GCRI	Global Climate Risk Index
GCOS	Global Climate Observing System
GLOFAS	Global Flood Awareness System
GHMs	Global Hydrological Models
GSA	Global sensitivity analysis
GSMaP	Global Satellite Mapping of Precipitation
GUI	Graphical User Interface
GDP	Gross Domestic Product
GR4J	modele du G ` enie Rural ´a` 4 parametres au pas de temps Journalier
HBV	Hy- drologiska Byråns Vattenbalansavdelning
HRUs	Hydrological Response Units
HYPE	Hydrological Predictions for the Environment
IOD	Indian Ocean Dipole
ITCZ	Intertropical Convergence Zone
IMERG	Integrated Multisatellite Retrievals for Global Precipitation Measurement (
INRAE	France's National Research Institute for Agriculture, Food and Environment
ISDR	International Strategy for Disaster Reduction
JMA	Japan Meteorological Agency
JRA-55	Japanese 55-year Reanalysis (JRA55),
KMD	Kenya Meteorological Department
KRCS	Kenya Red Cross Society
KGE	Kling Gupta Efficiency
LSMs	Land Surface Models
LVB	Lake Victoria Basin
LRQ	Latonyanda River Quaternary
LPM	Linear Perturbation Model
LVGF	Linearly Varying Gain Factor Model
LSA	Local sensitivity analysis
MJO	Madden-Julian Oscillation
MAM	March-April-May
MK	Mann -Kendall
MERRA	Modern-Era Retrospective analysis for Research and Applications

MAE	Mean Absolute Error
MRLP	Mean Residual Life Plots
MODFLOW	Modular finite-difference Flow model
MSI	Model Suitability Index
MI	Multiple Index
MSWEP	Multi-Source Weighted-Ensemble Precipitation
MCA	Multi Criteria Analysis
NDMA	National Disaster Management Agency
NMHS	National Meteorological and Hydrological Services
NERC	Natural Environment Research Council
NAM	Nedbør-Afstrømnings-Model
NSE	Nash-Sutcliffe efficiency
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NET	North Equatorial Trough
NWP	Numerical Weather Predictions
OND	October- November-December
AFFS	Pan-African Flood Forecasting System
PDM	Probability Distribution Model

PERSIANN-CCS Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System

PET	Potential Evapotranspiration
QDM	Quantile Delta Mapping
QBO	Quasi-Biennial Oscillation
RMSE	Root Mean Square Error
SSTs	Sea Surface Temperatures
SA	Sensitivity Analysis
SCE	Shuffled Complex Evolution
SHEAR SSC	Science for Humanitarian Emergencies and Resilience Studentship Cohort
SVATS	Soil–Vegetation–Atmosphere Transfer schemes
SLM	Simple Linear Model
SMAR	Soil Moisture Accounting and Routing
SST	Split sample testing
SSE	Sum of squared errors
SWAT	Soil Water Assessment Tool
TSI	Total Sensitivity Index
TOPMODEL	TOPography based hydrological Model
TRMM	Tropical Rainfall Measuring Mission

TMPA 3B42RTTRMM-based Multi-satellite Precipitation Analysis 3B42 Real-Time ()

UH	Unit Hydrograph

UNCEF	United Nations International Children's Emergency Fund
UNFCCC	United Nations Framework Convention on Climate Change
USGS EROS	U.S. Geological Survey Earth Resources Observation and Science
UK	United Kingdom
VIC	Variable Infiltration Capacity
WATCH	Water and Global Change
WRA	Water Resource Authority
WMO	World Meteorological Organization

List of equations

Chapter 1

1 Introduction

The introductory chapter of this research thesis provides an overview of the research background. Firstly, background information on flooding events and flood risk in Kenya is presented (Section 1.1), with a discussion of the factors (natural and human induced) contributing to extreme events across the country (Section 1.2), and flood early warning and response at global and regional level (Section 1.3). Secondly, the chapter sets out the motivation of this research by outlining some of the challenges, associated with flood modelling and forecasting in data-scarce regions, confronting Kenya (Section 1.4). This will be followed by an account of the research aims, objectives and research questions to be addressed. Finally, Section 1.5 provides an outline of the remainder of the thesis.

1.1 Flood risk in Kenya

According to the Global Climate Risk Index (GCRI), Kenya is ranked among the countries with the highest climate risk in the world. Kenya is prone to climate and weather extremes, particularly droughts and floods (Eckstein *et al*., 2019), with the latter being the most common climatic extreme and the leading hydro-meteorological disaster in East Africa (EA) (Huho and Kosonei, 2014). Floods commonly affect the low-lying areas located in Kenya's nine densely populated drainage basins, with Lake Victoria Basin (LVB) being the worst affected (Huho and Kosonei, 2014). This contributes to relatively high exposure and vulnerability levels, with much of the population affected in some way (Kilavi *et al.*, 2018; Otiende, 2009).

Kenya has experienced several major flood events recently (EM-DAT & Owour, 2000). Floods occur during rainy months, with flash and riverine floods occurring after torrential rains (Wanzala and Ogallo, 2020). The typical population affected per event is estimated to be approximately 70,000 (Parry *et al.*, 2012). Between 1964 and 2021, 20 major flood events were recorded, with the floods of 1961, 1997–1998, 2002, 2003, 2006, 2010, 2012 and 2018 having particularly high impact (Table 1), and being declared national disasters (Parry *et al.*, 2012; Kilavi *et al.*, 2018). During the first quarter of 2010, flash floods claimed the lives of 73 people and 1,864 livestock countrywide. Over 3,375 households were displaced, affecting 14,585 people. In addition, at least 16 bridges were destroyed in Rift Valley province (Kenya Floods, 2010). Flash floods caused by the 2012 long rains killed 84 people, displaced around 30,000 and affected over 280,000 (Kenya Floods, 2012). Following the long rains season of 2018 at least 311,164 people were displaced by floods across 40 counties in Kenya and, according to the National Disaster Management Agency (NDMA), this particular flood event claimed the lives of approximately 150 people (OCHA, 2018; Kenya Floods, 2019). The most recent severe flood was during the 2019 enhanced short rains season which affected most parts of the country. According to the Kenya Red Cross Society (KRCS), more than 160,000 people, including nearly 18,000 households, were displaced countrywide by floods or landslides, with approximately 330,000 people affected and 132 deaths reported, including those of 72 people whose houses were buried in a landslide in West Pokot County in North-Western Kenya (Kenya Floods, 2019; UNICEF, 2019).

Flooding in Kenya occurs annually, with an annual average economic loss to the country approximated at 5.5% of gross domestic product (GDP) (Connor, 2015) from damage to roads, water systems, buildings, and communication networks; direct costs of treatment for waterborne diseases; and crop and livestock loss (Warner and Van der Geest, 2013). With the changing climate, enhanced precipitation is expected to increase the frequency of major flood events in Kenya (Huho and Kosonei, 2014). For instance, from October to December 2019, Kenya experienced one of the wettest short rains seasons on record, with rainfall totals ranging up to 400 percent of the average (FEWSNET, 2019; Kenya Floods, 2019). This led to countrywide floods resulting in a huge economic loss from damage to roads, energy supplies, and buildings infrastructure (Njogu, 2021). Additionally, some of these extreme rainfall events are influenced by phenomenon such as -Southern Oscillation (ENSO) causing El Niño floods (see Table 1), and detailed description of these phenomena in Section 1.2.

1.2 Factors contributing to extreme events and increased vulnerability in Kenya

Precipitation is arguably the most important driver of catchment hydrological response, particularly of floods, but is hard to estimate, like other components of the hydrological cycle (Tapiador *et al.*, 2012; Beck, van Dijk, *et al.*, 2017). Accurate measurement of precipitation is made difficult by the high spatio-temporal variability of rainfall events (Vischel *et al.*, 2011), especially in the tropics, where most of the rainfall is controlled by mesoscale convective systems. This is also compounded by the general scarcity of precipitation gauging networks, except for a few dense networks operated by dedicated organizations or programmes (Gosset *et al.*, 2013).

There is a great deal of variability in the spatial and temporal patterns of Kenyan rainfall (Asnani and Kinuthia, 1979; Ogallo, 1984). These variations are caused by the presence of large inland lakes, complex topography, the Indian Ocean to the east, and the seasonal migration of the Intertropical Convergence Zone (ITCZ) (Ayugi *et al.*, 2018). The ITCZ, though not properly defined over land, migrates through this region twice as the sun moves overhead (Nicholson, 2018), and together with the North Equatorial Trough (NET) and the Indian Ocean Dipole (IOD), control the seasonal behaviour of the migration of the precipitation zone (Nyenzi, 1988). The seasonal rainbelts shift northwards and southwards with the ITCZ, with rainfall patterns in some locations modified by local features such as lakes and highlands (Ongoma, Chen and Omony, 2018). Throughout the year, the northern coastal strip, mountains, and areas near water bodies receive a significant amount of precipitation. In addition, rainfall events over Kenya is greatly influenced by weather phenomena such as El Niño -Southern Oscillation (ENSO) (Ayugi *et al*., 2020; Ojara *et al.*, 2021; Onyutha, 2016) and play a major role in extreme rainfall events and inter-annual variability (Ongoma *et al*., 2015). For example, the warm phase of ENSO/El Niño results in unusually heavy rainfall, causing rare floods like the 1997/1998 occurrence (Takaoka, 2005).

Near the equator, there are two separate rainy seasons (bimodal), whereas further south and in the northeast, the country has a single lengthy summer rainy season (unimodal) (Camberlin and Okoola, 2003; Camberlin *et al.*, 2009). The bimodal rainfall seasons occur during March-April-May (MAM) and October- November-December (OND) and are commonly known as the 'short

and 'long' rains respectively (Omondi *et al.*, 2014; Ongoma *et al.*, 2015; Kilavi *et al.*, 2018) . This seasonal pattern is caused by the rainfall migrating northward at a slower rate than it migrates southward (Nyenzi 1988).

Table 1: History of flood occurrences and the affected areas in Kenya (Huho et al.,2014; Kenya Floods, 2010-2019)

Year	Area of occurrence	Impacts
2019	West Pokot landslide and widespread flooding	330,000 people affected, 132 killed including 72 landslide victims
2018 **	Mandera, Garissa, Isiolo, Tana River, Taita Taveta, Western Kenya and parts of Nairobi city	186 people killed, 300,000 displaced, 800,000 people affected
2017	Tana River (Taita Taveta), Mombasa	5 people killed, 5,000 displaced
2016	Nairobi, Northern Kenya (Wajir, Marsabit, Turkana)	13,129 people affected, 16 people killed in a collapsed building in Huruma estate, Nairobi.
2015	Widespread	15 people killed and thousands displaced Infrastructure destroyed
2014	Narok town, Nairobi City	Property and infrastructure destroyed
2013	Tana River County	82,000 people displaced
2012**	Nyanza/ Western (Nzioa River basin, Nyando Basin	84 people killed, 30,000 displaced Approximately 280,0000 people affected
2010 **	Budalang'i, Tana River, Turkana	73 killed, 14,585 people affected
2008	Rift valley, Kitale, Makueni, Kibwezi, Budalang'i	24 people killed, 2,396 affected
2006 **	Widespread	7 deaths, 6,500 people displaced
2004	Nyeri, Othaya, Kihuri	5 people killed
2003 **	Nyanza, Western and Tana River Basin	60,000 people affected
2002 **	Nyanza, Busia, Tana River Basin	150,000 people affected
1997-1998***	Widespread	1.5 million people affected
1985	Nyanza/Western	10,000 people affected
1982	Nyanza	4,000 people affected
1961***	Athi River, Lamu, Tana River Basin, Nzioa and Nyando Basin	

*** El Niño flood, ** National disaster

Most flood events in Kenya occur during the two rainy seasons. The huge impacts of these floods are partly due to overreliance on seasonal forecasts (with lead times of one month ahead and longer) by users of climate information (such as government authorities, planners and policy makers (Graham *et al.*, 2011). These forecasts are uncertain and have low skill in forecasting

hydrological variables at the timescales relevant for floods in most parts of the world (e.g., Shukla et al., 2019; Mishra et al., 2019; Coelho et al., 2010). In Eastern Africa seasonal forecasts are correlated with large-scale climate drivers (Nicholson, 2017), such as the El Niño-Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD) modes and the Madden- Julian Oscillation (MJO) (MacLeod et al., 2021). This can sometimes be helpful in anticipating floods:, for example, a positive IOD in 1997, 2005 and 2019 is likely to have caused the enhanced rains during those years (Bethel and Dusabe, 2021). A positive IOD event in the Indian Ocean is associated with anomalously warm Sea Surface Temperatures (SSTs) in the western Indian Ocean, near to Eastern Africa, and anomalously cold SSTs in the eastern Indian Ocean (Saji et al., 1999). Westerly winds in the middle equatorial Indian Ocean move moisture away from Eastern Africa during the long rains season (OND). During extreme positive IOD events, the north-central Indian Ocean experiences strong low-level easterly wind anomalies, which weakens the westerly flow that normally transports moisture away from Eastern Africa, resulting in wetter conditions over Eastern Africa and drier conditions in the central and eastern Indian Ocean basin (Wainwright et al., 2021). However, the long rains have less spatial and temporal coherence and a weak correlation with most of these large-scale drivers. As a result, the long rains season is far less predictable than the short rains season (Dutra et al., 2013), hence most parts of the country are adversely impacted during this particular season, e.g., by the MAM 2018 rainfall events (Kilavi et al., 2018).

Flood preparedness and early action in the face of an incipient flood event depend on correct flood forecast information and timely dissemination of early warning messages, to relevant organizations involved in flood management and to vulnerable communities (WMO 2003; Njogu, 2021). Kenya, like many other countries, is currently grappling with the global COVID-19 pandemic (Lone and Ahmed, 2020), with much attention and many resources diverted to tackle its impacts. For example, by September 2020, the number of deaths due to covid was approximately 1,000, according to a study by Ombajo et al. (2020), and rose by 20 percent by June 2022 (WHO, 2022; Ombajo et al., 2022); the Global Fund requested the Ministry of Health (MoH) to reallocate USD 5million to the COVID-19 response (Barasa et al., 2021). Yet at the same time, floods continued to kill people in their thousands. For example, thousands of people had been displaced in Kenya due to flooding in Lakes Baringo and Bogoria, which disrupted many livelihood activities and social infrastructure, eroded farmlands, and threatened cross-contamination of water between the two lakes (Aura et al., 2020). As the flood waters recede, some semblance of 'normal' life' returns: the "flood memory" is short and the recovery of affected communities is quickly forgotten by the authorities, news media and the international community (Njogu, 2021), until the onset of another flood event. People forget floods quickly and so there is an inadequate understanding and perception of flood risks in Kenya both by the authorities and those affected. News of devastation, such as broken bridges (Kenya , 2021) and massive landslides caused by the raging floodwater, hit mainstream and social media for only a matter of hours and then other topics take their place, and it becomes a forgotten story (Nyakundi, Mwanzo and Yitambe, 2010).

Compare, for example, the flooding events in Europe (e.g., Germany and the UK in 2021) and Africa (e.g., Kenya, Nigeria, Somalia). There is a remarkable difference in terms of response and recovery, forecasting, the dissemination of information, and media coverage. For instance, flood

forecasts for Germany were issued by the European Flood Awareness System (EFAS) before the event, and bulletins sent to relevant authorities (Cloke, 2021). Despite this, there were still fatalities (NBC News, 2021), but there was also a significant media presence and there were subsequent discussions about what went wrong and what lessons could be learnt for the future (Da Costa, 2021). Parliamentary inquiries and post disaster "Lessons learned" exercises are important components of improving resilience to future floods (Stephens and Cloke, 2014; HM Government, 2016). Even though flood disasters are common in Kenya, no such initiative has been brought forward by Kenyan authorities and could be alluded to the incapacity to deal with issues of such magnitude yet.

In Kenya there are distinct challenges and limitations in the production and dissemination of flood-specific early warning to relevant authorities, and in associated media coverage, as well as in response and recovery efforts (Wanzala and Cloke, 2021). The connection between the media and science communication, knowledge and behaviour are beneficial in reducing the impacts of natural hazards such as floods. For example, the media promote an increased understanding of the onset and evolution of extreme events, the mobilization of response efforts, a discussion of these subjects with others, and the facilitation of online debates on matters surrounding climate change and their potential impacts on communities (Barnston *et al.*, 2010; Weisheimer and Palmer, 2014). Therefore, the media are critical in the dissemination of information: they should work more closely, collaboratively, and vibrantly with other actors than at present. Mass media can also play an important role in influencing inquiries and policy change. The visibility of the Europe 2021 flood event in news, discussions and media interviews with climate scientists contributed to the triggering of an official parliamentary committee inquiry into the flood catastrophe (Cloke, 2021)

1.3   Flood early warning and response in Kenya

Flood Forecasting (FF) is an important tool in reducing vulnerabilities and flood risks as well as improving the resilience of communities through well informed response and adaptation measures. Reliable flood forecasts (with medium to seasonal timescales) have recently become possible in many parts of the developed world. This is due to the progressive advancements in Numerical Weather Predictions (NWP), because of the expansion of data assimilation techniques (English *et al.*, 2013), improvements in precipitation datasets (Nowak and Hodson, 2013), higher computing power (Bauer *et al.*, 2015), improvement in numerical modelling techniques (Yamazaki *et al.*, 2011), and the incorporation of ensemble modelling (Cloke and Pappenberger, 2009).

However, in many developing countries, there is still no tangible evidence of the benefits resulting from these advancements. Mr Mwai, the deputy director at the Kenya Meteorological Department (KMD), noted that "presently, FF in Kenya is limited due to inadequate tools (models), model source codes, data, personnel, forecast lead-time, and inadequate documented research to back up operations", which hinders informed flood preparedness actions. With the advancements in computing, various hydrological models have been developed to improve flood modelling and forecasting in many parts of the world. However, the application of these tools is challenging, for several reasons, such as inadequate observational data, lack of informed forecast model skills, uncertainty in forecast skills, limited applications due to geographical constraints and models

without open-source codes (Thiemig, De Roo and Gadain, 2011). For instance, in Kenya, the Kenya Meteorological Department (KMD) runs an operational flood forecast system only in Nzioa catchment out of the nine flood prone catchments. The model adopted for this system is the Soil Moisture Accounting and Routing Model (SMAR) incorporated in the Galway Flow Forecasting System (GFFS) (O'Connor, 2005)(see Section 2.4.3 for detailed description of SMAR model). The choice and use of the SMR model was an entirely subjective matter, mainly driven by the project funding following the push to implement a FF system in Nzioa and for its simplicity and comparatively low data requirements. Further detailed discussion of FF system in Kenya can be found in Section 4.5.5 of the thesis.

Flood Early Warning Systems (FEWS) are important in the context of flood incident management activities including evacuation, protection of vital infrastructure, relief and rescue operations, emergency medical support etc. An early warning system has four parts: forecasting, transformation of the forecast into a warning, dissemination of the warning to local decision makers and users, and translation of the warning into remedial action (ISDR, 2004). According to the Kenya Meteorological Department (KMD), a flood-specific EWS is currently operational only in one out of the nine flood prone catchments in Kenya: the Nzoia River in Western Kenya (Thiemig, De Roo and Gadain, 2011), as it was the only catchment to have a reliable long time series of observational discharge data. No published documentation on this forecast modelling approach exists, and so there is neither information explaining the informed choice of the modelling approach, nor a skill assessment of the model used in this catchment. Extensive research in other flood prone catchments in Kenya is severely hampered by the lack of reliable observational data (precipitation and discharge data) (Njogu, 2021) and thus there is no understanding of how the hydrology in these catchments is changing through time. In addition, a review of the public experience of early warnings suggests that lead times are often insufficient for effective action (Opido *et al.*, 2017), and there is consequently a need for flood forecasts with a longer lead time.

One of the ways by which improved flood EWS is supported in flood vulnerable regions such as Kenya is through global initiatives, such as the Global Flood Awareness System (GLOFAS), which is part of the Copernicus Emergency Management Service for floods, (www.globalfloods.eu). GloFAS provides 30 days extended lead-time flood forecast information using ensemble meteorological forecasts from the European Centre for Medium-range Weather Forecasts (ECMWF) for the major river basins across the world (Alfieri *et al.*, 2013). GloFAS, however, has only recently (2019) added operational monitoring points on the major rivers in the flood prone areas in Kenya. There is thus a new opportunity to explore the use of GloFAS in Kenya.

1.4 Forecasting and modelling floods

1.4.1. Challenges associated with data scarcity

Forecasting floods in data-scarce regions is very challenging. Data at different spatial and temporal timescales is vital for the processing and production of actionable flood forecast information (Zanchetta and Coulibaly, 2020). However, with the current observational data scarcity issues, combined with a lack of human and computational resources, network connectivity, and adequate

maintenance resources, there is often a preference among operational agencies for simpler modelling tools (Mr Mwai, personal communication). Such tools require less data but necessarily have a simpler process representation and may not yield the detailed information required for forecasting floods. These subjective choices of which model to use for flood forecasting, based on the requirement for simplicity in undertaking the modelling, have the potential to affect forecasting efforts significantly. It is therefore necessary to establish an objective pre-selection criterion for models, which informs the first objective of this research, further detailed in Chapter 4.

Models cannot be run in practice if the necessary input data are not available, and extra efforts should be made to estimate the input data from freely available sources. Other sources of precipitation data, such as those from satellite remote sensing, are now available, but they come with their own uncertainties, including random and systematic errors (Sun *et al.*, 2018; 2020). This, combined with model uncertainties, may result in a worse level of performance. Additionally, some of the datasets are at a very coarse scale and downscaling to finer scales to run the model in a fine grid size leads to extra work and may not result in better performance statistics. Therefore, these datasets suffice as best alternatives but first, assessment of their performance in area-specific applications is crucial, which was adequately addressed by the second objective of this research, further detailed In Chapter 5.

There is a growing concern that major flooding events in many parts of Kenya in the past decade are indicative of the effects of a changing climate (Wainwright *at al* 2021; Wanzala and Ogallo, 2020). Understanding flood characteristics such as frequency, magnitude and timing is important for informing policy for disaster risk management, infrastructure design and agriculture, amongst other hydrological applications (Rosner, Vogel and Kirshen, 2014; Bezak, Brilly and Šraj, 2016). In addition, consideration of the trends in flood data series may result in more accurate flood timing, magnitude and frequency estimations (Berghuijs *et al.*, 2017, 2019; Mangini *et al.*, 2018; Sa'adi *et al.*, 2019). Trend analysis can be used to investigate whether there is any evidence of an increase in river floods in the observational river discharge data, which is adequately addressed by third objective and further detailed in Chapter 6.

Therefore, the aim of this research is to contribute to a better understanding of the uncertainties in models and datasets intended to inform flood modelling and applications in data-scarce regions, and provide evidence of change in the frequency and magnitude of flood events to inform and improve flood modelling and forecasting in Kenya.

1.4.2. Research objectives and questions

Research Objective 1: - Design and propose an objective model pre-selection criterion for a Kenyan national flood centre.
- **Research Question 1.1**. What are the important factors to consider in objective choice of models for flood applications in Kenya amidst data scarcity issues and varied physiographic settings?
- **Research Question 1.2**. How practical and effective is the proposed pre-selection criterion?

Research Objective 2: - Evaluate the utility of different reanalysis precipitation datasets for hydrological modelling through performance statistics and parameter identifiability to establish their influence on the catchment streamflow simulations.

- **Research Question 2.1:** - Can reanalysis datasets be used in hydrological modelling and how are the model parameters impacted by different precipitation datasets?
- **Research Question 2.2.** How does the performance of a hydrological model vary with different reanalysis datasets and is a higher resolution dataset any better in simulating streamflow?
- **Research Question 2.3.** Which of the model parameters is most sensitive and how does this vary across the different precipitation datasets in simulating catchment streamflow?

Research Objective 3: - Assess the historical trends in flood series and possible shifts in flood timing across Kenyan catchments.

- **Research Question 3.1**. What are the historical trends of flood events in the observed flood series in Kenya and how has this changed over time?
- **Research Question 3.2.** What are the trends in the GR4J model simulated discharge across Kenya for the period of 1981 – 2016. How do they compare with the trends in the observations?
- **Research Question 3.2.** Has there been a shift in the occurrence and timing of floods across Kenya?

1.5. Thesis Structure

The thesis continues with a detailed literature review, description of data and methods used, three research papers which form the main body of the thesis, discussion, and conclusions as illustrated in Figure 1.1.
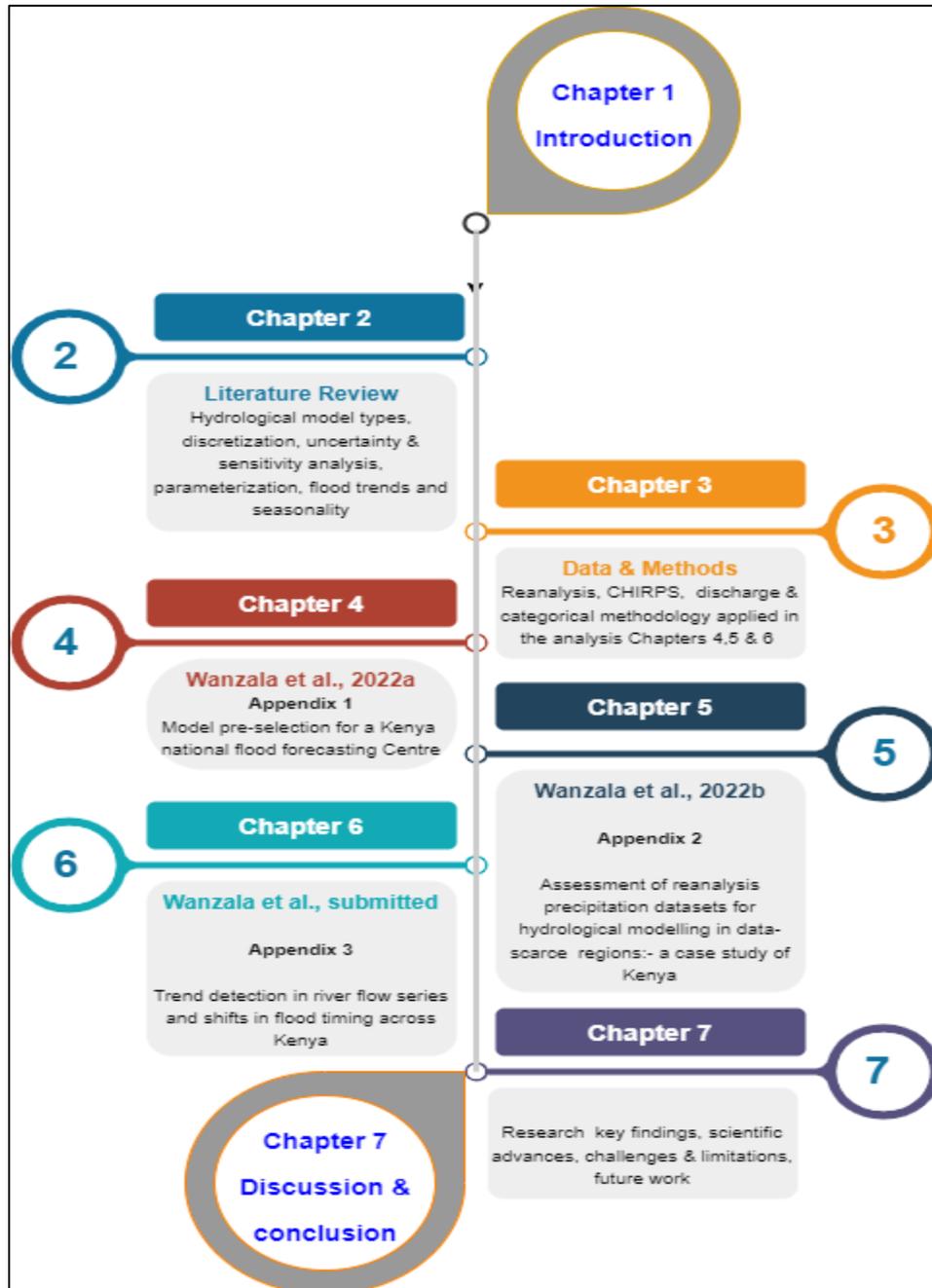
Figure 1.1:- Thesis structure: small, coloured circles indicate the main chapters with 4, 5 and 6 being the analysis chapters. The speech bubble at the top and the bottom introduces and concludes my thesis respectively.

Chapter 2 is a comprehensive literature review on hydrological models (2.1 & 2.4) and classes (2.2), catchment discretization (2.3) model parameterizations (2.6), input datasets for hydrological modelling (2.5) and trends and seasonality of floods (2.7). These form the literature context for Chapters 4, 5 and 6 of the thesis.

Chapter 3 covers data (3.1) and methods (0) used to achieve the objectives in 1.4.2.

Chapter 4 is the first paper presented in this research and addresses the first objective (I). It provides an in-depth model selection framework for a national flood forecasting centre, with focus on Kenyan hydroclimates and challenges associated with data scarcity. The research proposes modelling tools that can be adopted for operational flood forecasting that can help mitigate the effects of floods in data-scarce regions such as Kenya.

Chapter 5 is the second paper and addresses the second objective (II). It assesses the performance of reanalysis datasets relative to observations at catchment scale considering both the performance statistics and uncertainty quantifications**.** The findings are important in informing future applications of reanalysis products for setting up hydrological models that can be used for flood forecasting, early warning and early action in ungauged catchments in Kenya.

Chapter 6 is the third paper and addresses objective (III) on the trends in river flow series and possible shifts in flood timing and predictability. The conclusions provide evidence of patterns in trends in frequency and the magnitude of flood events for catchments in proximity.

Chapters 7 contain the discussions (7.1) and key messages of the research, with its challenges and limitations (7.2), presented in the three papers, scientific advances of the research findings (7.3) the key conclusions (7.4) and recommendations for future work (7.5).

The three research papers presented in this thesis have been reformatted as thesis chapters but exactly correspond to the work as published. The published versions of Chapters 4 and 5 are provided in Appendix **A2** and **A3**. At the time of submission, Chapter 6 was still at the reviewing stages of publication, but the online review copy can be found in Appendix **A4**. Statements of authors' contributions are given at the beginning of each of the chapters.

Chapter 2

2 Literature Review

This chapter presents a detailed review of the different types of hydrological models and their characteristics, catchment discretization, input datasets for hydrological modelling and their evaluation, uncertainty and sensitivity analysis of model parameters, and the concepts of equifinality in modelling, model parameterization, non-stationarity, calibration, and validation.

2.1 Hydrological models

Hydrological models are important tools for understanding natural processes in water resources (Jyathilake, 2019), usually predicting hydrological variables such as river flow or soil moisture. A model can be defined as a set of constructs, derived from explicit assumptions about how a system responds to specified inputs (Beven, 2012). For example, a hydrological model would model the interaction of inputs (e.g., climate) with the system (e.g., catchment), to produce an output (e.g., the outflow hydrograph) through representation of the physical, biological and chemical characteristics of a catchment and simulation of the natural hydrological processes (Jain & Singh, 2019). Hydrological models therefore incorporate runoff generation and river routing processes to simulate hydrological response to meteorological variations (Sutanudjaja *et al.*, 2018). They can be used to represent a real-world hydrological system to better understand various water and environmental processes, predict system behavior and provide consistent impact assessment (Beven, 2012; Devia *et al*., 2015; Towner *et al*., 2019), particularly where data is scarce (Sutanudjaja *et al.*, 2018). They are used in a range of hydrological applications such as short to extended-range flood forecasting (Alfieri *et a*l., 2013; Emerton *et al*., 2018,., climate assessment (Tamm, Luhamaa and Tamm, 2016; Hattermann *et al.*, 2017; Lu *et al.*, 2018), hazard and risk-mapping (Artan *et al.*, 2001; Ward *et al*., 2015), drought prediction (Van Huijgevoort *et al.*, 2014), and water resource assessment) (Dessu *et al*., 2016; Mutie, 2019; Praskievica and Sang, 2009; Sood and Smakhtan, 2015).

At the global scale, a distinction can be made between Land Surface Models (LSMs) and Global Hydrological Models (GHMs), although the division between these classes is becoming less and less distinct as LSMs become more complex and there is a move towards earth system modelling (Harrigan, Cloke and Pappenberger, 2020). Whereas LSMs describe the vertical exchange of heat and water, GHMs are not only focused on water resources and the lateral transfer of water (Haddeland *et al.*, 2011; Trambauer *et al.*, 2013), but also parameterize ground water processes for simulating ground water recharge and water table depth from the perspective of water resources assessment (Koirala *et al.,* 2014). Different hydrological models make use of different inputs, variables at different temporal timescales (e.g., rainfall, air temperature) and static variables (e.g., soil characteristics, topography, vegetation, hydrogeology, and other physical parameters) (Devia and Dwarakish, 2015). In some cases, little input data exists, and the model can be used to estimate the runoff and river flow in ungauged catchments. The choice and application vary across different classes of models and a comprehensive review of some of the models is covered in Section 2.4 of this thesis. This review is important as it provides more information about the different

classifications of hydrological models and their differing parameters and data requirements, as well as applications in various hydrological studies in different parts of the world. The reviewed models also provide the contextual information for the model evaluation presented in Chapter 4, which aids in the pre-selection process of suitable models for operational flood forecasting and flood applications at national level. The subsequent sections of this chapter provide a comprehensive review of different classes of models (Section 2.2), an account of catchment discretization in models (Section 2.3) and a brief overview of the hydrological models and their characteristics (Section 2.4), before considering input datasets for hydrological modelling (Section 2.5), and model parameterization (Section 2.6).

2.2 Classification of hydrological models

Models can be classified . into two broad categories, based on: - (i) mathematical representation (conceptual, empirical and physically based models) and (ii) spatial configuration (lumped, semi-lumped, distributed and semi-distributed), but not all models fit in these classifications (Devia *et al*., 2015; Quang, 2016). Hydrological models can be initialized in a continuous or event-based mode. Continuous model initialisation requires running a model in a warm-up period to let the model states reach values that no longer depend on arbitrarily chosen initial values. Usually, a standard one climatic year is used, but this warm-up depends on the model and the catchment memory of past conditions (Berthet *et al.*, 2009). However, in some catchments influenced by large aquifers, the warm-up period may take several years (Le Moine, 2008). Running a model in continuous state, however, requires a long continuous precipitation time series, which may be difficult to provide. In addition, it may be difficult to gather a long enough data series before issuing the first forecast at new locations (Berthet *et al.*, 2009). Event-based model initialisation requires a separate method to derive the initial values of model states. For example, if the model states reliably represented measurable physical quantities, recent measurements or values based on climatology would be solutions, as shown in the findings of Brocca *et al.*(2009) that assimilating soil moisture measurements into the event-based SCS-CN model can be useful for flow simulation on a small catchment. A detailed description of model classification can be found in Section 2.2.1. Summary characteristics, advantages and disadvantages of physically based, conceptual and empirical and lumped, semi-distributed, and distributed models are shown in Table 2 and Table 3 respectively. Detailed information of some of these models' categories and their application is found in Section 2.4. An example of general classification is represented in Figure 2.1, which follows the description of model classification terminologies in Jajarmizadeh, Harun and Salarpour (2012)(but not all models can fit in this classification. The subsequent sub-sections provide descriptions of different classes of models.

2.2.1 Based on mathematical representation

2.2.1.1 Conceptual models

Conceptual (also known as grey-box) models incorporate conceptual representation of catchment information by means of simple concepts, such as representing soil storage as a leaky bucket, and are characterized by parameters that usually have no direct, physically measurable identity (Wheater et al., 2007). They incorporate conceptual representations of preceding information and

important processes and are thus easy to run and calibrate. Parameters need to be calibrated, which requires fitting them to an observed data set to obtain an appropriate set of parameter values, using either a manual or automatic procedure (Wheater et al., 2007); 2007)see further details in Section 2.6.3.The model structures are characterized by catchment physical features and their semi-empirical nature. The hydrological unit is a combination of a transport routing system and storage for elements such as surface, soil and groundwater. Routing modules may include one or more linear or non-linear stores. Model parameters vary depending on the process descriptions and the number of processes; thus, the models are moderately complex.

2.2.1.2 Empirical models

Empirical (also known as black-box) models are based on observations or experience and do not contribute to physical understanding of the catchment (Quang, 2016). They make use of mathematical equations which are derived from the input and output data series, and hence valid only within boundaries, e.g., the unit hydrograph method. The functional relationships between the model inputs and outputs are built from the regression and correlation equations applied within the set boundaries. Artificial neural networks and fuzzy regression are some of the machine learning techniques used in hydro informatics methods.

2.2.1.3 Physically based models

Physically based models compute flows and energy fluxes from physical equations (Beven and Kirby, 1979; Booker & Woods, 2014). They can be semi-distributed or fully distributed in terms of discretization and parameter allocation. These models explicitly incorporate theoretically derived understanding of the catchment characteristics using hydrologic state variables and fluxes (Beven, 2010; Beven, 1989; Fatichi et al., 2016). These models can be used in a closure of assumed forms of the laws of conservation of mass, energy, and momentum at temporal scales in reference to underlying physical processes. When applied spatially at catchment scale, physically based distributed models can incorporate the space–time variability of the primary forcing, such as precipitation and radiation, and variations of land-surface properties (e.g., soils, vegetation, and topography) at the sub-catchment scale, while resolving the subsurface domain in horizontal and vertical directions (Fatichi *et al.*, 2016). Thus, they can be easily applied at regional scales.

The complex and heterogeneous internal conditions within a catchment are often difficult to apply and solve. Therefore, distributed models are useful in this case, as they describe internal states and fluxes whilst considering preservation of mass, energy, and momentum budgets. These models are good at explaining specific variables at the local scale that can be simulated only with detailed representations, such as snowmelt (Johnson, Zhang and Downer, 2013), land management (Mutie, 2019), landslide occurrence(Anagnostopoulos, Fatichi and Burlando, 2015) and snowpack evolution (Lehning *et al.*, 2006).

2.2.2 Based on spatial configuration

2.2.2.1. Lumped models

In lumped models the parameters, inputs, and outputs are spatially averaged and take a single value for the entire catchment. Lumped models have a range of advantages in their application at the catchment scale. i) They give a general picture of how the catchment works as a whole, since it is hard to predict the response of any natural land-surface to a given rainfall event. ii) Each of the building blocks of any given distributed model is itself a lumped rainfall-runoff model. However, models of this kind have a limitation in the sense that, if an event parameter cannot be uniquely identified, then it cannot be linked to catchment characteristics, and that is a major problem in application to ungauged catchments. Similarly, it is difficult to represent catchment change if the physical significance of the parameters is unclear. In between is a semi-lumped model which may adopt a lumped representation for individual sub-catchments (Wheater *et al*., 2007).

2.2.2.2. Semi-distributed models

In semi-distributed models, parameters, inputs, and outputs are allowed to vary spatially. It divides the whole catchment into Hydrologic Response Units (HRUs) based on other variables in addition to land use and land cover, soil type, and slope and simulates the various hydrological processes in each HRU. Further description of HRUs can be found in section 2.3 of this thesis. models. Compared to lumped models, semi-distributed, and distributed models better account for the spatial variability of hydrologic processes, input, boundary conditions, and watershed characteristics. While lumped models like GR4J (Perrin, Michel and Andréassian, 2003) present an entire river basin in one unit, the spatial variability of the basin is represented by distributed models like the MIKE SHE model (Abbott *et al.,* 1986).

2.2.2.3. Fully distributed models

In fully distributed rainfall-runoff models, the catchment is divided into a number of cells whose physical properties are assumed homogeneous and characterised by a series of reservoirs linked vertically and representing different hydrological processes. In each cell, these hydrological processes are represented by linearised approximations of the non-linear differential equations that govern the movement of water (Beven 2012).
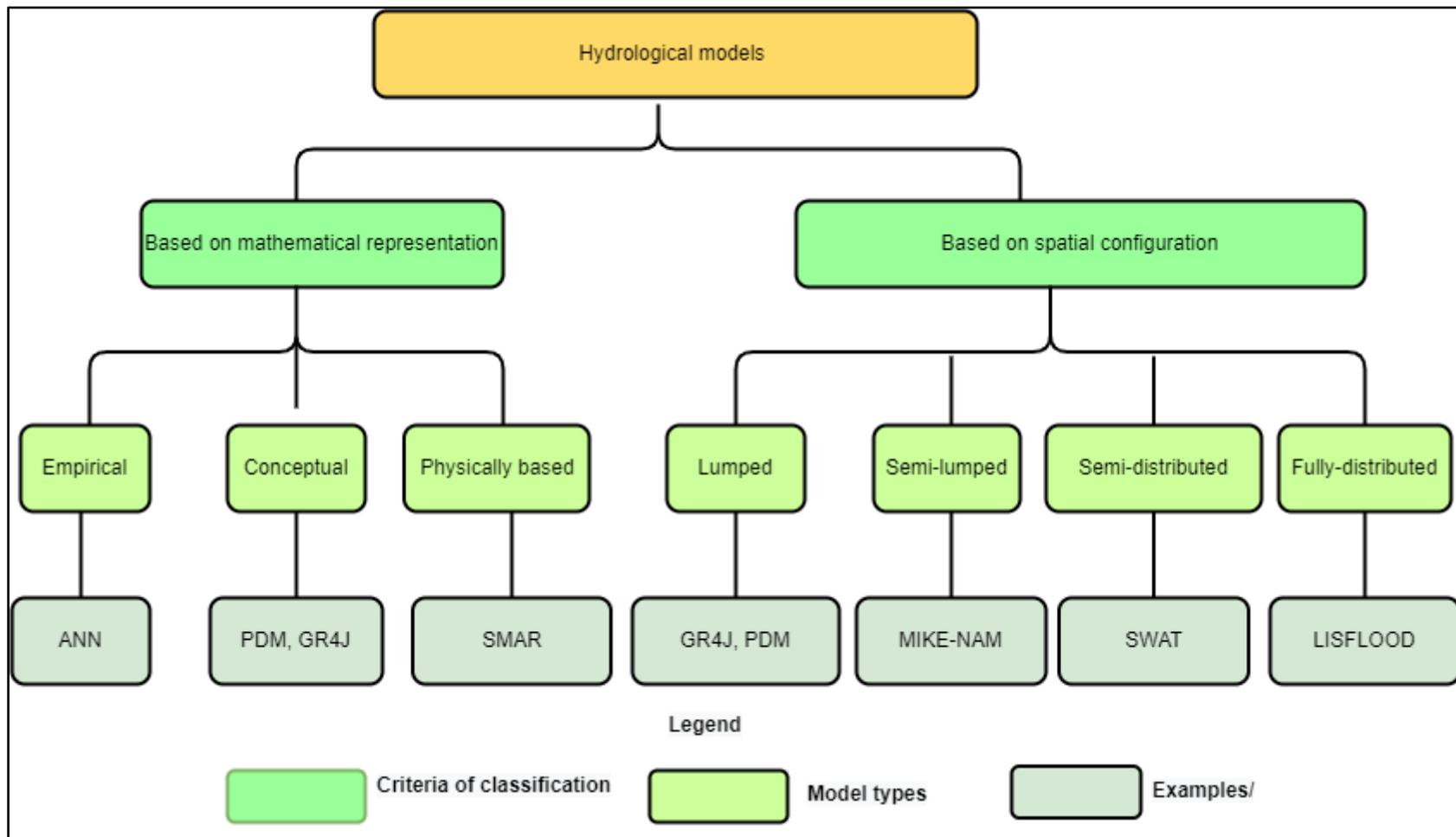
Figure 2.1:- Example of ways of classifying hydrological models.

All these classifications make use of common forcing data such as rainfall, potential evapotranspiration, pumping schedules, irrigation schedules, etc., and each model has distinct parameters. Model parameters are quantities intended to characterize the inherent properties of a modelled system (both observational and physical) (Kavetski, 2019), and vary from one model to another. Lumped conceptual models, such as GR4J (Perrin *et al*., 2003) have just a handful of parameters (four), representing production store, groundwater exchange, routing store, and time base of the unit hydrograph. Semi-distributed physically based models like SWAT (Arnold *et al.,* 1998) may have many parameters describing soil hydraulic properties, crop growth rates, surface lags, etc., and describe processes, parameter estimation approaches, time scales, and spatial resolution of input data and simulations by spatial discretization, where a catchment is divided into Hydrological Response Units (HRUs) (Xu *et al*., 2012; Haddeland *et al*., 2011), which is explained in Section 2.3 of this thesis.

2.3. Catchment discretization and hydrological response units (HRUs)

HRUs are "distributed, heterogeneously structured areas with common land use, soil and geology controlling their unique hydrological dynamics": the dynamics of hydrological processes within an HRU vary only by a small amount compared to the dynamics among different HRUs (Flügel, 1995; Savvidou *et al.,* 2016). For a better representation of catchment processes, HRUs can be classified into topographic based (Leavesley *et al.*, 1983) and homogeneous HRUs (Flügel, 1995). A homogeneous HRU requires threshold specifications for land cover, soil and slope classes, which are then used to delineate HRUs (Neitsch et al., 2011). The process involves the division of catchments into several sub-catchments, which are further divided into discontinuous landmasses through aggregate delineation. A user defines thresholds for soil types, slope ranges and land use within each sub-catchment. A spatial overlay scheme based on the Geographical Information System (GIS) is then used to produce "unique combination" HRUs with homogeneous characteristics. The resultant HRUs represent percentages of the sub-catchment area whose contributions differ from the entire catchment responses. Topographic-based HRUs permit the simulation of lateral fluxes within the hillslope and vertical fluxes into the soil due to their ability to integrate the topological flow routing between HRU subareas into the process-based distribution concept (Savvidou *et al.*, 2016). For catchments with areas less than 200 km^2, homogeneous HRUs are much preferred as they provide a better representation of the catchment processes (Bongartz, 2003).

However, in most distributed hydrological models, a catchment is assumed to be an assembly of discrete entities in the form of grid cells, with different properties, that contribute differently to its responses (Efstratiadis *et al.*, 2008). Thus, HRUs denote spatial elements of pre-determined geometry, while the parameterization of the hydrological processes is dictated by the model discretization (Daniel *et al.*, 2011; Nalbantis *et al.*, 2011).

Table 2:- Summary Characteristics of Physically based, Conceptual and Empirical Models

Classification/ Descriptive characteristics/	Physically based Models	Conceptual Models	Empirical Models
Model classification	Mechanistic or white box models.	Parametric or grey box models.	Data based or metric or black box models
Description of hydrological processes	Based on spatial distribution, evaluation of parameters describing physical characteristics	Based on modelling of reservoirs and includes semi empirical equations with a physical basis.	Involve mathematical equations, derive value from available time series.
Data requirement for model parameterization	Require data about initial state of models and morphology of catchment.	Parameters are derived from field data and calibration.	Little consideration of features and processes of system.
Computer and power to run the models	Complex models. Require human expertise and computation capability.	Simple and can be easily implemented in computer code.	High predictive power, low explanatory depth.
Length of hydrometeorological data for calibration of parameter sets and scale & model examples	Suffer from scale related problems. SHE or MIKESHE (Abbott *et al.*, 1986), GeoSFM (rtan. et al., 2004), VIC (Liang *et al.*, 1994).	Require much hydrological and meteorological data GR4J (Perrin *et al*., 2003), HBV model (Berglöv *et al.*, 2009), TOPMODEL Beven and Kirby, 1979), PDM (Moore, 2007).	Cannot be generated to other catchments. ANN, unit hydrograph.
Validity range in application	Valid for a wide range of situations, e.g., applied to any set of data over which physical laws are valid.	Calibration involves curve fitting; makes physical interpretation difficult.	Valid within the boundaries of given domain and calibration data range.

Table 3: Summary Characteristics of Lumped, semi-distributed and distributed models

Characteristic/ classification	Lumped Models	Semi distributed Models	Distributed Models
Variability of input parameters within the catchment	Input parameters do not vary spatially within the basin.	Input parameters are allowed to vary partially in space.	Input parameters are allowed to vary fully in space at a resolution chosen by the user.
Response of output to individual sub-basins	Response or output is evaluated only at the outlet without considering the response of individual sub-basins.	Response is evaluated by dividing the basin into several smaller sub-basins.	Response is evaluated by dividing whole basin into small sub-basins.
Ease of simulating event-based processes	Not applicable to event-based processes.	Lie between lumped and distributed models.	Applicable to event-based processes.
Physical meaning of parameters	Parameters do not represent physical features of hydrologic processes; model parameters area weighted average	Require less data than the fully distributed models	Consider the hydrologic processes taking place at various points in space and define the model variables as functions of the spatial dimensions.
Data requirement	Low data requirements.	Lower data requirement than distributed models.	Require large amount of data.
Simplicity	Easy to use.	Lie between.	Require expertise.
Accountability of point hydrological process	Prediction results are at catchment outlet only	Lie between.	Prediction results can be obtained at any location and time.
Computational speed and capacity	Simple in nature and require minimal computational time.	Lie between requiring more computation speed and time	Requires much computational time due to being cumbersome in nature.
Common examples	Examples: GR4J (Perrin *et al*., 2003), PDM (Moore, 2007), NAM (Nielsen and Hansen, 1973), SMAR (O'Connor and Zhang, 1970).1970).	Examples: SWAT (Arnold *et al*., 1998), GeoSFM (Artan. *et al*., 2004), HBV-96 (Berglöv *et al*., 2009).	Examples: LISFLOOD (van der Knijff *et al*., 2010), MIKE SHE (Abbott *et al*., 1986), TOPMODELv3 (Beven and Kirby, 1979).
Treatment and inclusion of the physical process at prediction stage	They do not consider governing processes during result predictions.	Lie i between	Consider detailed governing physical processes down the modelling chain.
Accuracy in simulating observations	Not very accurate.		Highest accuracy is achieved if accurate data are available.

2.4. Overview of hydrological models

This section provides a detailed review of several hydrological models. This section is important because the information is used to develop a hydrological model selection framework with criteria (Chapter 4), and the models considered in the evaluation are those outlined in this section. Im model selection, sampling is done from a many available models. In this thesis, convenient sampling (which models are the authors familiar with) was partially considered and much weight given to judgemental sampling (models are included based on expert judgement). I sampled the models from a predetermined number of models widely applied to Kenyan studies as well as trialled in the FF prototypes, thus lower weighting factor given to convenient sampling. Judgement sampling is much considered after the detailed evaluation presented in Table 9 in Section 4.6.2 in Chapter 4 of the thesis. Also, Chapter 5 looks at evaluation of precipitation datasets for hydrological applications, in which the GR4J model (Section 2.4.1 and detailed processes outlined in 3.3.2.2.1) is calibrated and validated over Kenyan catchments. The sub-sections here provide summary descriptions, with their applications, of twelve hydrological models that have been applied globally across different catchments. In some models, a description of the parameters, variables and processes is given, and a summary of their real meaning outlined in Table 4. Key references of the models are also summarised in Table 8. The strengths and weaknesses, processes, and data etc. of each model are summarised in Table 9.

2.4.1. GR4J

The GR4J model, known as *modèle du Génie Rural à 4 paramètres au pas de temps Journalier* (Perrin, Michel and Andréassian, 2003; Ficchì, Perrin and Andréassian, 2019), is a simple daily continuous lumped conceptual rainfall-runoff model with only 4 parameters. The GR4J model is the modified version of the GR3J model which originally focussed on the soil moisture compartment. Typical meteorological input data used in the GR4J model are precipitation (P) and evapotranspiration (E) at daily time steps (Figure 2.2). P is an estimate of the areal catchment rainfall that can be computed by any interpolation method from available rain gauges and PE is estimated from the average value of evaporation. A schematic diagram for the GR4J model, illustrating input data, processes and parameters and a corresponding FORTRAN code is available on the *Cemagref* Website. The *airGR* package in Coron (2019) has now been released to ease the implementation of the GR models.

There are four free main parameters to be optimized in GR4J model, namely: maximum capacity of production store (X1, mm), groundwater exchange coefficient (X2, mm), maximum capacity of non-linear routing store (X3, mm), and time base of the unit hydrograph (X4, days) (Ficchì, Perrin and Andréassian, 2019) and fixed parameters, whose values were set in Perrin, Michel and Andréassian (2003) (

Table **5**). The processes represented include interception, evaporation, groundwater, runoff, and routing components. A production store and a routing store are the main conceptual reservoirs in the model. The production store controls the water infiltration into the ground following the characteristics of a conceptually unsaturated reservoir. The role of the routing store is to disconnect the groundwater reservoir and contribute to the water routing out of the system. Water is routed

linearly by two-unit hydrographs (UH1 and UH2), with a fixed contribution of 90% and 10% respectively. The conceptual nature of the model coupled with this routing method are useful in simulating the time lag between a rainfall event and the corresponding peak discharge (Perrin, Michel and Andréassian, 2003; Arnal, 2014).

Due to its very simple design, GR4J can be calibrated relatively quickly, and so various versions of this model have been used in a range of studies worldwide for different applications (Oudin *et al.*, 2004; Berglöv *et al.*, 2009; Tian, Xu and Zhang, 2013; Van Esse *et al.*, 2013; Humphrey *et al.*, 2016; Mostafaie *et al.*, 2018; Ficchì, Perrin and Andréassian, 2019; Lerat *et al.*, 2020). For instance, Mostafaie *et al.* (2018) compare different multi-objective functions in calibrating a conceptual rainfall runoff model in the Danube River Basin. Van Esse *et al.* (2013) compare model structures to assess their effectiveness in different climates between the calibration and validation periods, in catchments with flashy flows, and in catchments with unexplained variations in low flow measurements within French catchments. Humphrey *et al.* (2016) apply a hybrid monthly streamflow forecasting approach to explore the simulation of soil moisture from the GR4J model to represent initial catchment conditions in a Bayesian artificial neural network (ANN) statistical forecasting model in Southeast Australia catchments. Results indicate a good performance in forecasting high flows in terms of the accuracy of the median forecasts, as well as reliability and resolution of the forecast distributions.
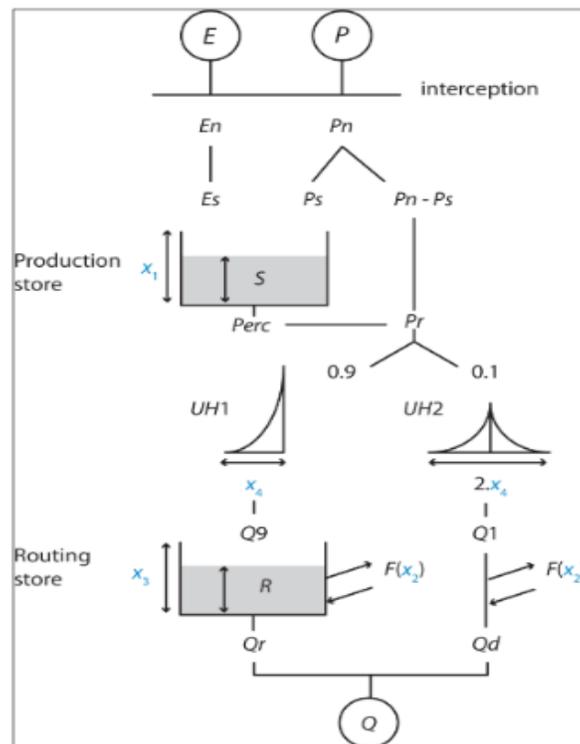
Figure 2.2: Schematic representation of the GR4J rainfall-runoff model (Source: Perrin et al., 2003). *P* is rainfall depth; *E* is potential evapotranspiration estimate; *Q* is total streamflow; X_i are the model parameters; all other letters are model variables or fluxes summarized in Table 4.

The GR4J model has been used in several studies over a range of catchments in Africa. For instance, Kodja *et al.* (2018) analysed the performance criteria of the GR4J model for reproducing high water flows in the Ouémé catchment in West Africa. The results showed that GR4J overestimated the streamflow during the low water period and underestimated flow at high water. Ruelland *et al.* (2010) simulated hydro-climatic variability over Sudano-Sahelian catchments and found that GR4J provided a near approximate estimate of cumulated discharge. Traore *et al.* (2014) used the GR4J and GR2M models to evaluate the availability of water in the basin of Koulountou Rivern, a tributary of the Gambia River, and the results show a good performance,ion depicted by a high NSE value. Tegegne *et al*. (2017) compared three models (GR4J, IHACRES and SWAT) in the Upper Blue Nile Basin to assess their usefulness in water resources assessment. Results show that GR4J performed best for the simulation of high, mid-range, and dry flows in a range of catchments.

Consequently, the GR4J model was adopted for further analysis in this thesis and a detailed description of the model processes and equations are provided in Section 3.3.2.2.1 of Chapter 3.

2.4.2. NAM

NAM (Danish: Nedbør-Afstrømnings-Model) is a lumped conceptual rainfall–runoff model developed at the Technical University of Denmark (Nielsen and Hansen, 1973) and later incorporated into the MIKE11 river modelling system, which includes a user interface and a multi-objective autocalibration scheme (Madsen, 2000). The schematic representation of the NAM model is shown in Figure 2.3.
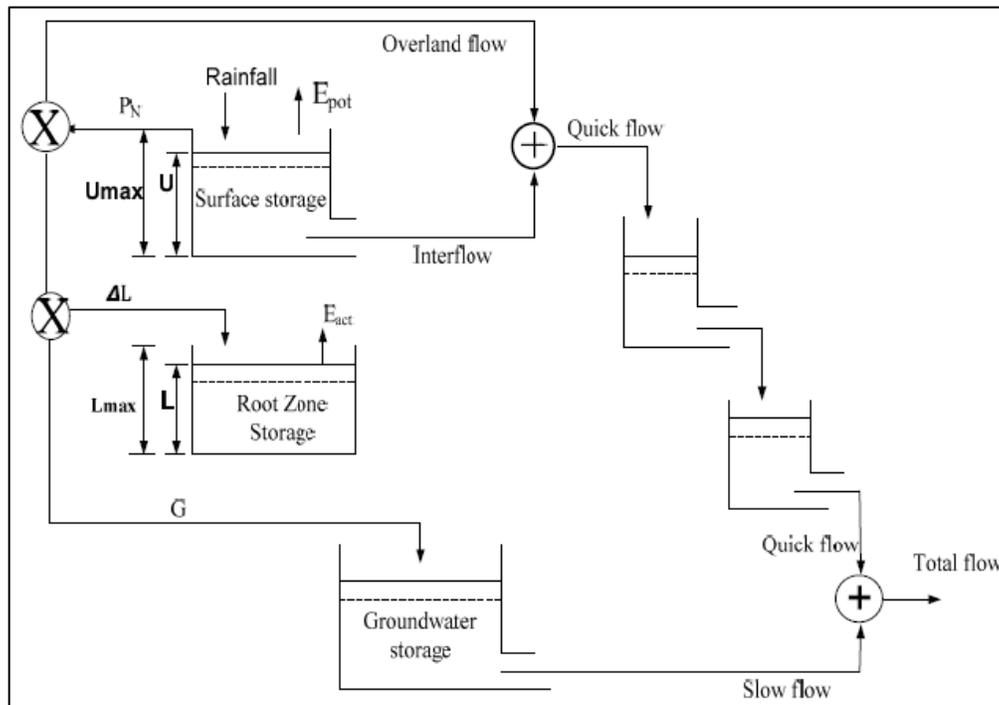
Figure 2.3: Schematic overview of the NAM model (Nielsen and Hansen 1973).

Flow in NAM is divided into overland flow (surface flow), interflow (subsurface flow) and base flow. The model contains four water storage components: surface storage, snow storage, root zone storage and groundwater. In addition to these storages the NAM model also incorporates the other depletions with the help of irrigation and ground water pumping modules. The NAM model has nine model parameters which are used to route water through the mutually connected storages (Nielsen and Hansen, 1973). The snow storage is considered only in cases where snowmelt contributes considerably to runoff; of the other three main storages, the upper zone storage represents vegetation, depressions and near surface (cultivated) soil, and the lower zone storage represents the root zone and the main soil horizons, while the groundwater storage represents water-bearing rocks.

The basic meteorological data inputs into the NAM model are precipitation, potential evapotranspiration, and temperature at daily timesteps. The model utilizes these to produce output in the form of catchment runoff, subsurface flow contributions to the channel, and information about other elements of the land phase of the hydrological cycle, such as soil moisture content and groundwater recharge (Agrawal and Desmukh, 2016). Like GR4J, NAM is simple and easy to calibrate and so it has been applied in several studies around the globe to simulate flow values in respect of rate, timing, and volume. For example, it has been used for design flow and stage computations (Rahman *et al.*, 2011), reservoir inflows (Razad *et al.*, 2018), climate change scenarios and impact projections (Vansteenkiste *et al.*, 2014), rainfall-runoff modelling (Doulgeris *et al.*, 2011; Lin *et al.* 2014; Singh *et al.*, 2014) and ecosystems and water resource management (Doulgeris *et al.*, 2012). For flood applications, the NAM model has been used in several studies. For example, Katuva *et al.* (2018) used NAM to assess water allocation and hydrological simulation in the Mukurumudzi River Basin in Kenya. Through an abstraction survey conducted along the entire length of the river followed by a low flow study and water balance modelling, they established remarkable differences in simulated and observed flows at source and mouth during the dry months of July and August. Wara *et al.* (2014) assessed climate change impacts on the water cycle and river flow regime of the Nyando River Catchment using the NAM model. They established that climate change would result in stream-flow alteration and increased frequency and intensity of flooding, with consequences for water availability and agricultural productivity. Odiyo *et al.* (2012) conducted a study on the Latonyanda River Quaternary catchment using the MIKE 11 NAM model. Results show that observed and simulated flow at catchment scale correlated well except for underprediction of peak events and a few low flows. In addition to this, a few overpredictions occurred due to illegal irrigation abstractions as they reduce the observed values and are not captured in the modelling. The study of Nguyen and Tran (2010) on the Ben Hai River basin, which combines auto calibration with the trial-and-error approach of the MIKE 11 NAM model, concludes that the good agreement between simulation and observation of the hydrograph's shape and total flow volume indicates the model parameters are consistent. An additional detailed review on applications of the NAM model can be found in Agrawal and Desmukh (2016).

2.4.3. Soil Moisture Accounting and Routing Model (SMAR)

The SMAR model is a lumped conceptual rainfall-evaporation-runoff model introduced by O'Connor and Zhang (1970) and modified by Khan (1995) and Liang (1992). SMAR consists of

a water balance component and a routing component: the former simulates in a very simplified manner the various physical phenomena governing the runoff generation process, and the latter transforms therunoff generated by the water balance component into discharge at the catchment outfall (Tan and O'Connor, 1996). A schematic diagram showing the structure of the standard SMAR model, incorporating the modifications suggested by Khan (1995) and Liang (1992), Figure 2.4.
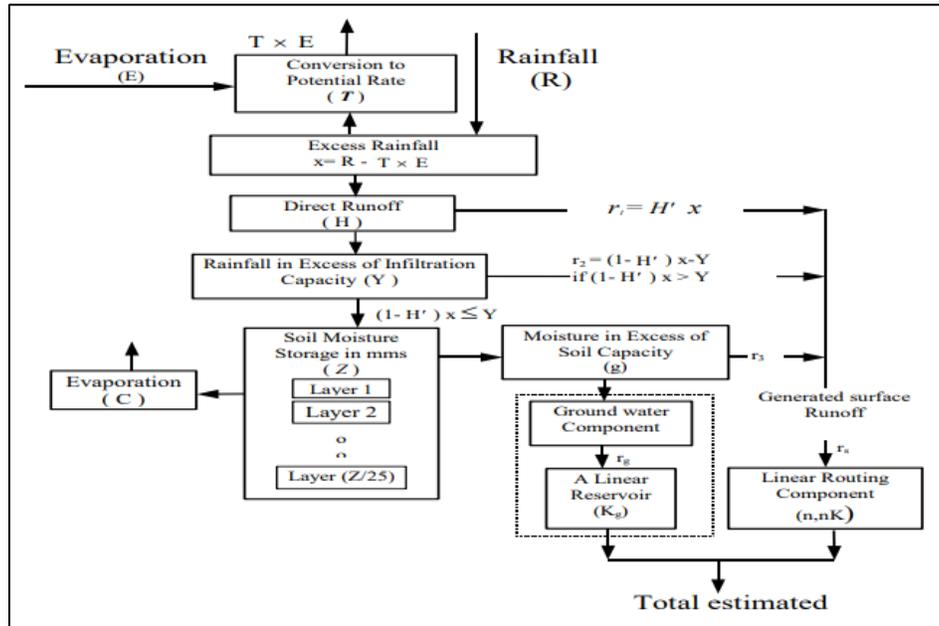
Figure 2.4: Schematic diagram of SMAR model (Liang 1992; Khan, 1995).

The main meteorological inputs of the SMAR model are rainfall and temperature at daily timescales. In total, the SMAR model has nine parameters, namely Z, T, H, Y, C, G, N, NK, and K_g. Z is the combined water storage depth of the layers, T is a parameter (less than unity) that converts the given evaporation to potential, C Evaporation decay parameters, facilitating the lower evaporation rates from the deeper layers, H is the direct runoff coefficient, Y is the maximum infiltration capacity, N is the shape parameter of the Nash gamma function model, a routing parameter, NK is the scale parameter of the Nash gamma function model, a routing parameter, G is the groundwater weighting parameter and Kg is the storage coefficient of the linear reservoir, a routing parameter. Some of these parameters may be fixed at appropriately chosen values, while the values of the rest are usually estimated empirically by optimization to minimize the objective function in the form of the sum of the squares of the errors between the observed and estimated flows (Kachroo, 1992; Goswami and O'Connor, 2010). Detailed technical description of the SMAR model, optimization of the parameters and updating modules can be found in Zhang (1970).

For flood applications, SMAR has been applied to several studies across the globe, due to its simplistic nature. For instance, the SMAR-AR Model was used for real-time river flow forecasting on the Blue Nile catchment at a location near the Sudanese-Ethiopian border (Shamseldin *et al.*, 2009). The SMAR-AR model discharge forecasts were compared with those of SAMFIL for the flood season (August-September) and the discharge forecasts of the SMAR-AR model proved to be far more reliable than those of the SAMFIL. An assessment of the applicability of the Galway

River Flow Forecasting and Modelling System (GFFMS) for Lake Tana Basin, Ethiopia, which included the SMAR model, established that the SMAR model gives reliable forecast results with an NSE value of 0.78 (Dessalegn *et al.*, 2017). Fazal *et al.* (2005) records the assessment of the potential of the SMAR model to estimate groundwater recharge using only rainfall, evaporation and groundwater level data in Miyakojima Island, Japan. Using the Genetic Algorithm (GA) optimization technique, the authors established that the estimated recharge by the SMAR model was 45% of the mean annual rainfall and concluded that the model could be a viable choice since it could estimate dependable recharge with a minimum of input data. Three flood forecasting techniques were assessed to determine the most efficient model for flood forecasting in the Nzoia Basin, Kenya: SMAR-LTF, ANN-NARX and LPM-LTF; SMAR -LTF and ANN-NARX were the better performers (Gathee and Odera, 2015)..

2.4.4. Probability Distribution Model (PDM)

The Probability Distribution Model is a conceptual rainfall-runoff model which transforms potential evaporation and rainfall data to flow at a catchment outlet (Moore, 2007). The most commonly used PDM model configuration comprises a probability-distributed soil moisture storage, a surface storage, and a groundwater storage component (Moore, 2007). A probability-distributed soil moisture storage component is based on a Pareto distribution of soil moisture storage capacity over a catchment and separates direct runoff from subsurface runoff. A surface storage component transforms direct runoff to surface runoff using a two-linear reservoir cascade formulated as a transfer function with dependence on two past outputs and current and previous input. The groundwater storage component receives drainage water from the distributed soil moisture storage as input and contributes the ground water component of total runoff as output. The schematic representation of the PDM model is shown in Figure 2.5.
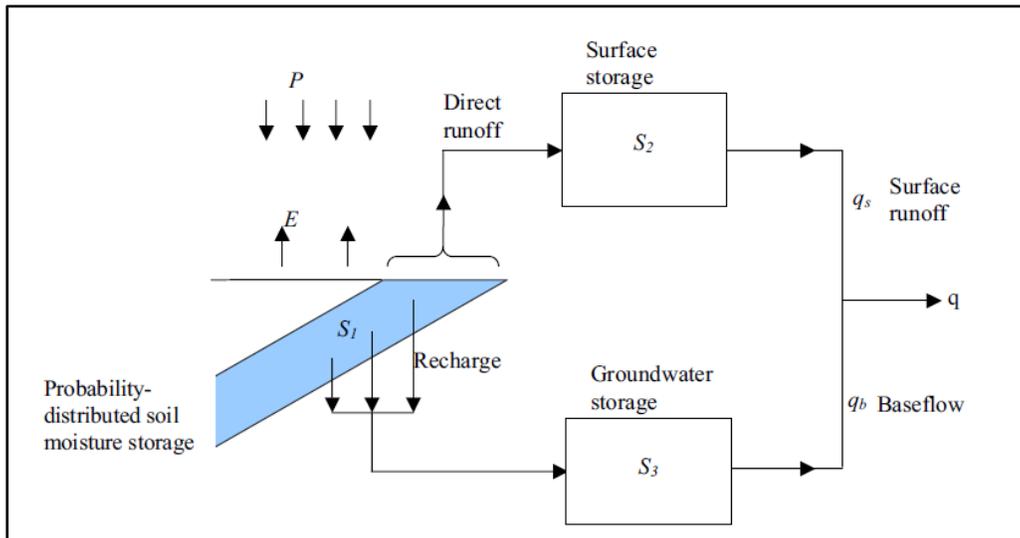
Figure 2.5:The PDM rainfall runoff model: s_1 is the recharge, s_2 is the surface storage, $s3$ is the groundwater storage, q_s is the surface runoff, q_b is the baseflow, E is evapotranspiration and P is precipitation (Moore, 2007)

Runoff production at a point in a catchment is controlled by the absorption capacity of the soil (taken as canopy and surface detention) to take up water. A runoff production model which integrates point runoffs can be formulated, taking into consideration variation in the catchment storage capacities and the spatial variation of capacity described by probability distribution. The groundwater recharge from the soil moisture store passes into subsurface storage. The outflow from the surface and subsurface storages and any compensation releases from the reservoir or constant abstractions form the model output (Moore, 2007). The main meteorological inputs into the PDM model are rainfall and evapotranspiration at hourly time steps. The PDM has several parameters detailed in Moore (2007).

PDM has been applied in studies in different parts of the world. For example, Bennet *et al*. (2016) assessed the performance of models including the PDM when calibrated with hourly rainfall disaggregated from daily forcing for streamflow forecasting applications, to show an improvement in model performance over mesoscale catchments. Willems *et al.* (2014) assessed the performance of NAM and PDM models in capturing the peak and/or low flow extremes, and changes in these extremes, for impact investigations on such hydrological extremes in Grote Nete and Nyando rivers in Belgium and Kenya respectively. Their findings reveal that identification of the model structure in a case-specific way does not lead to higher accuracy than the traditional NSE, which do not necessarily reflect the model performance for high and low flow extremes, and sub-models or sub-flows. A study by Srivastava *et al*. (2014) showed a poor performance of dynamically downscaled daily precipitation from ERA interim using the conceptual PDM in simulating the observed streamflow for the Brue catchment (UK). Ngaina (2014) assessed the capability of PDM using the Ensemble Kalman Filter (EnKF) to forecast flood events over the Nzoia sub-basin, Kenya. Results show a better performance in terms of RMSE and the Coefficient of Efficiency (CoF), and thus potential for improving flood forecasting to enable the management of flood related risk on a real time basis over the sub-basin.

2.4.5. Hydrologiska Byråns Vattenbalansavdelning- 96 (HBV-96)

HBV-96 is a conceptual semi-distributed hydrological model, for continuous calculation of runoff, originally created at the Swedish Meteorological and Hydrological Institute (SMHI) in 1996 from a re-evaluation of the lumped HBV hydrological model (Lindström *et al.*, 1997; Berglöv *et al.*, 2009; Arnal, 2014). The model's name is an abbreviation of Hydrologiska Byråns Vattenbalansavdelning (Hydrological Bureau Water Balance Department). The schematic representation of the HBV -96 model and the representative model parameters and routines are illustrated in Figure 2.6.
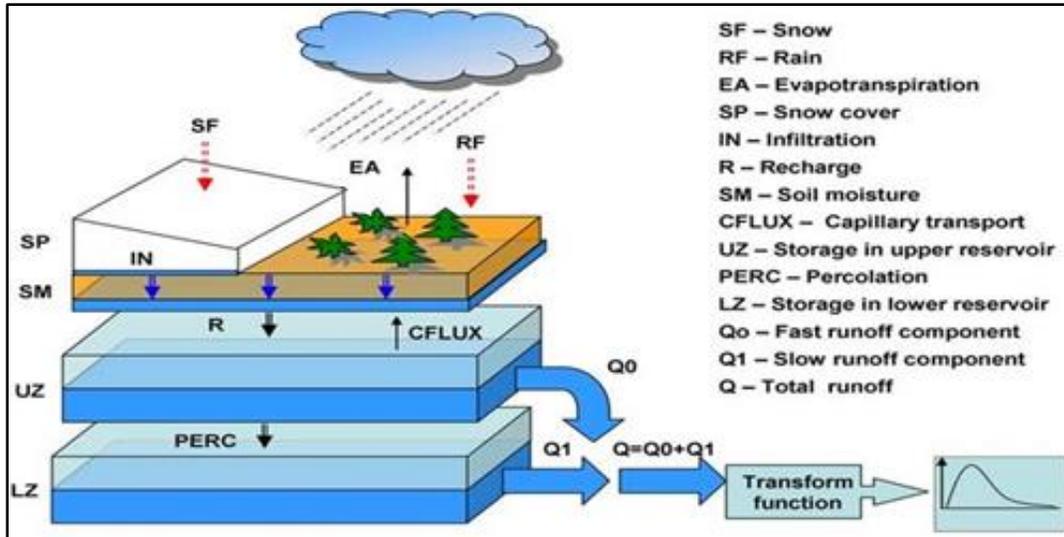
Figure 2.6: Schematic representation of the HBV-96 model (after Lindström et al. 1997) with routines for snow, soil, and runoff response represented by *SP, SM* and *Q* respectively.

HBV-96 runs at hourly time steps using the following input meteorological data: hourly precipitation and air temperature and potential evapotranspiration monthly mean averages. Sub-basins are subdivided into zones defined by elevation and spatial distribution of vegetation. Four different land use classes are used within the model: open areas, forests, lakes and glaciers (Berglöv *et al.*, 2009). Lakes have a significant impact on runoff dynamics and the routing in major lakes is therefore modelled explicitly. It has a simple interception storage for forested areas, but interception is neglected for open areas. The response function of the model transforms excess water from the soil moisture routine to discharge to each sub-basin. It consists of two reservoirs connected in series by constant maximum percolation rate and one transformation function (Lindström et al., 1997)

The structure of HBV-96 follows multiple routines, namely i) the precipitation routine, ii) the soil moisture routine and iii) the runoff response routine. The latter is composed of two zones, an upper non-linear reservoir and a lower linear reservoir. The upper zone reservoir generates a quick runoff flux, while the lower reservoir produces as an output a baseflow component (Arnal, 2014); they are then summed before being subjected to the MAXBAS transformation function, a triangular weighting function that is used as a routing function to compute simulated runoff (Bergstrom, 1995; Lindström *et al.*, 1997; Berglöv *et al.*, 2009). Over the years, the HBV-96 model has been under development and this has led to improvements, such as the data interpolation method, evaporation calculations and precipitation and temperature updating for forecasting (Berglöv *et al.*, 2009).

The HBV model has been used in global and regional flood forecasting studies in different parts of the world without the modification of its structure. A number of studies have shown that HBV-96 produces good flood forecast simulations because of its simple design and low number of model parameters, so is applicable to an extensive number of cases, especially large sub-basins such as the Rhine basin (Berglöv *et al.*, 2009; Arnal, 2014). More specifically, HBV-96 has been applied

in assessment of satellite precipitation datasets as alternatives to hydrological model inputs to simulate streamflows (Worqlul *et al.*, 2017), and model effects of deforestation and land use changes on streamflow (Muli, 2007; Rientjes *et al.*, 2011), behavioural changes in water levels and streamflow (Birundu & Mutua., 2017), and impacts of climate change on hydrological response (Booij *et al.*, 2011; Nobel, 2011; Abraham *et al.*, 2018).

2.4.6. TOPography based hydrological MODEL (TOPMODEL)

The first lumped version of the TOPMODEL model was developed by (Beven and Kirby 1979). This version of TOPMODEL treats a catchment as a set of homogeneous sub-catchment units in hydrologic response through the channel network method and so they have to be modelled separately (Beven et al., 1984). This version of the model is driven by a set of model parameters including maximum interception storage, infiltration storage level, overland flow velocity, field capacity, subsurface flow, channel velocity and sub-catchment topographic constant. Due to the complexity of the model and the numerous model parameters to be calibrated, there has been developments in the structure and physics of the TOPMODEL. The schematic representation of the lumped version of TOPMODEL adopted from Beven and Kirby (1979) is as shown in Figure 2.7.
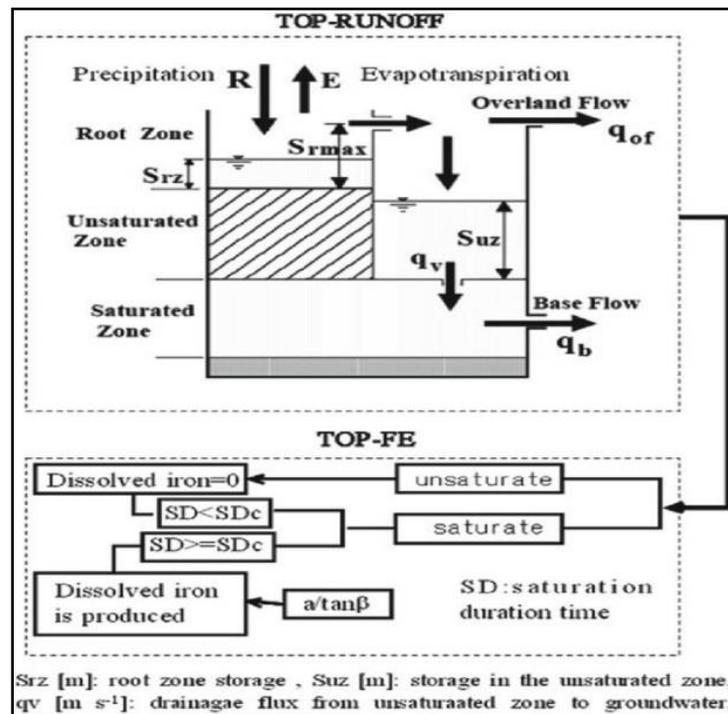
Figure 2.7: Schematic diagram of the components of TOPMODEL (Beven and Kirkby 1979).

Gao et al., (2015) developed the physically based spatially distributed version of the TOPMODEL. Unlike the lumped version, this uses grid cells as computational units, downscaling runoff production equation from catchment to cell scale, derived from the original version. A new module on the basis of the multiple-direction flow theory of Quinn et al.,(1991) and the Darcy-Weisbach equation is used to describe the overland flow (Gao, Holden and Kirkby, 2017). The flow direction

depends on the depth of water, surface roughness, topography, velocity and slope. Within this module, a stochastic algorithm is used to describe the routing of overland flow. TOPMODEL has two main components: the water balance at the soil surface useful for runoff prediction and the routing component, useful for transfer of runoff to the basin outlet. The transfer component is divided into two phases: the first representing the transfer along the slopes towards the drainage network and the second representing transfer along the drainage network to the basin outlet.

The meteorological input datasets for TOPMODEL include rainfall and potential evapotranspiration at daily or hourly timesteps. Five parameters optimized in TOPMODEL include; M - the parameter of the exponential transmissivity function or recession curve (m), Ln (T_o) - the natural logarithm of the effective transmissivity of the soil when just saturated in which a homogenous soil throughout the catchment is assumed (m^2/h), SR_{max} - the soil profile storage available for transpiration, i.e. an available water capacity (m), SR_{int} - the initial storage deficit in the root zone (m) and ChVel - routing velocity for scaling the distance over area – also called network width function whereby a linear routing is assumed (m/h) (Holden *et al.*, 2008; Gao, Holden and Kirkby, 2017). The physical processes represented are vegetation interception capacity, surface runoff from saturation and infiltration excesses. Vegetation interception capacity is represented by a reservoir with a capacity of SR_{max}. The water is extracted from the reservoir at the potential evapotranspiration rate; when net precipitation is more than the capacity SR_{max} it reaches the soil and forms the input for the subsequent model components. The saturated hydraulic conductivity of the soil follows the negative exponential law verses the depth, assuming the water table is parallel to soil surface.

The model had been used in flood forecasting and water related applications. For example, flood forecasting in ungauged catchments in humid temperate climate (Beven et al., 1984; Gumindoga et al., 2011), flash flood forecasting (Serrat-Capdevila, Valdes and Stakhiv, 2014), impacts of water abstraction such as reservoirs retention and discharging capacity (Peng *et al.*, 2016), water resources management (Chen et al., 2012) and climate change studies (Dietterick, Lynch and Corbett, 1999; Cameron, Beven and Naden, 2000).

2.4.7. Soil Water Assessment Tool (SWAT)

Soil Water Assessment Tool (SWAT) is an operational, process based, semi-distributed conceptual model that has been incorporated into an ArcGIS interface (ArcSWAT; Di Luzio et al. (2002). SWAT operates on a daily time step and is used to predict the impact of land management on water, sediment and agricultural chemical yields in large ungauged basins (Arnold *et al.*, 1998; Neitsch *et al.*, 2005). The SWAT model requires five basic input data sets: topography, soil, land use, climatic data and management data. The climate data includes daily precipitation, temperature, solar radiation, wind speed and relative humidity (Malag, Bouraoui and Roo, 2018). The schematic representation of the SWAT model, and the associated processes is illustrated in Figure 2.8.
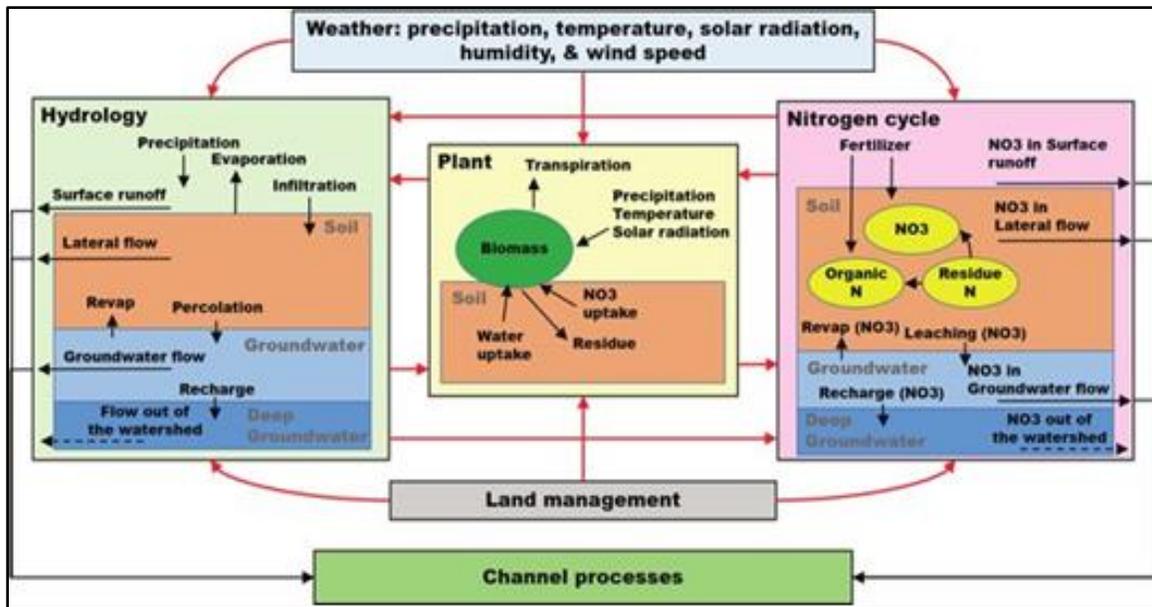
Figure 2.8: Schematic of SWAT model components (adapted from Neitsch et al., 2005).

In SWAT, a catchment is divided into multiple sub-basins, which are further subdivided into hydrologic response units (HRUs) that consist of a unique combination of soil, land use/cover, and slope (Tamm, Luhamaa and Tamm, 2016). The SWAT model simulates sub-basin components which are classified into eight major divisions: hydrology, weather, sedimentation, soil temperature, crop growth, nutrients, pesticides, and agricultural management. The overall hydrologic balance is simulated at daily time steps for each HRU, summarized at the sub-basin level, and then routed through the stream network to the catchment outlet (Abbott *et al*., (1986a; 1986b)). A command structure is used for routing runoff and chemicals through a catchment through streams and reservoirs, adding flows, and inputting measured data on point sources, thus the model can simulate a basin subdivided into grid cells or sub-watersheds (Tamm, Luhamaa and Tamm, 2016).

SWAT is currently being utilized in several large area projects. Intercomparing and hydrological modelling studies have shown relatively good streamflow simulations using SWAT. SWAT has global and regional applications over a range of large-scale catchments due to its simple design, readily available inputs and computational efficiency (Awan *et al.*, 2016; Xu *et al.*, 2016; Le and Pricope, 2017). For example, Denmark, Karlson *et al*. (2016) modelled the combined effect of land use and climate changes on hydrology and evaluated the sensitivity of the results from three hydrological models, NAM, SWAT and MIKE SHE. Results showed similar performance during calibration, the mean discharge response to climate change varied up to 30%, and the variations were even higher for extreme events (10th and 99th percentile). Land use changes appeared to cause little change in mean hydrological responses and little variation between hydrological models. Alemayehu *et al*. (2018) evaluated the efficiency of Climate Forecast System Reanalysis (CFSR) and Water and Global Change (WATCH) rainfall in simulating the observed streamflow SWAT model in the Mara Basin (Kenya/Tanzania). Studies by Baker and Miller (2013), Githui *et*

al. (2009), Hunink *et al.* (2013) and Sang *et al.* (2005) used the SWAT model in Kenyan catchments to assess the impact of land use changes and reservoirs on water resource management.

2.4.8   Geospatial Streamflow Simulation Model (GeoSFM)

The GeoSFM is a physically based semi-distributed hydrologic model developed by the U.S. Geological Survey Earth Resources Observation and Science (USGS EROS) Center (Artan. et al., 2004; 2001; 2007; Asante et al., 2008). The GeoSFM is a catchment-scale hydrologic model and simulates runoff processes using remotely sensed global data sets using a few parameters and variable input data (rainfall and evapotranspiration). The model consists of two components: a Graphical User Interface (GUI) component and a rainfall-runoff simulation component. The GUI component is useful in input data preparation and visualization of model outputs and is run within a Geographical Information System (GIS). The schematic illustration of the GeoSFM process map model is shown in Figure 2.9 and detailed description of the processes can be found in the manual at (https://pubs.usgs.gov/of/2007/1441/pdf/ofr2008-1441.pdf).
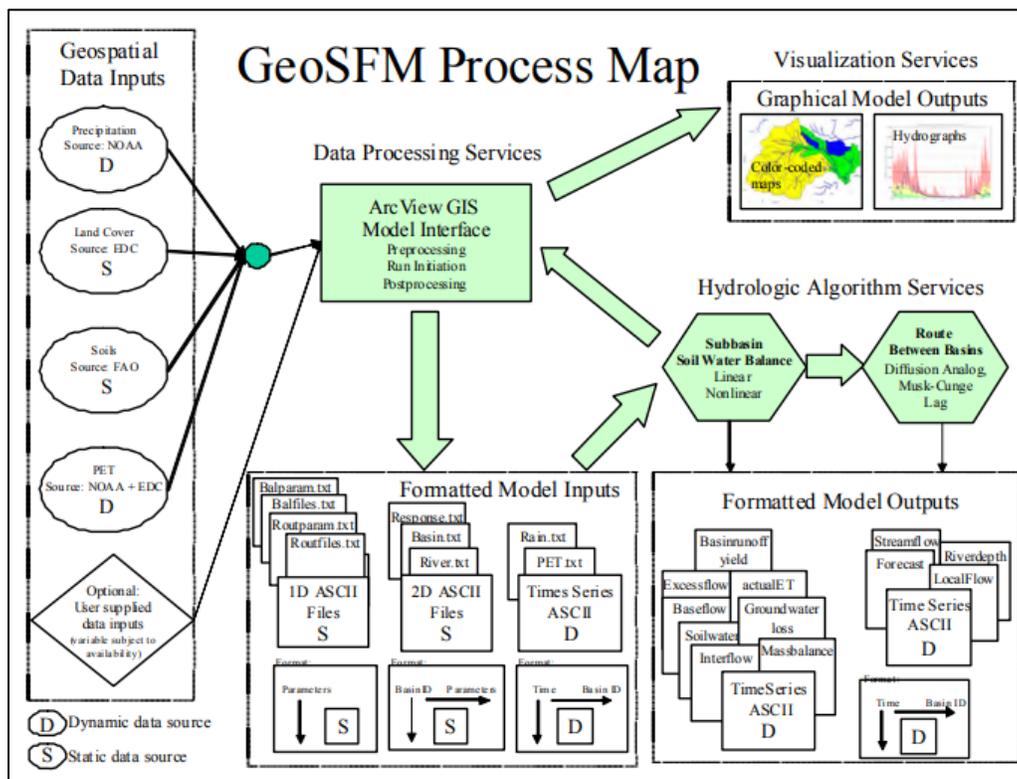
Figure 2.9: Process map and system diagram for the Geospatial Stream Flow Model. (Asante *et al.*, 2008)

Normally, the catchment is subdivided into several sub-catchments based on the digital elevation model (DEM) data in the pre-processing module. Most of the model parameters, topography, landcover and soils, derived from the continental-scale datasets, have a physical meaning. The rainfall-runoff component of the GeoSFM model is characterized by three main modules. These modules include the water balance, catchment routing and distributed channel routing. Daily water

balance calculations which determine how much water enters the stream network from each sub-catchment are a subject of sub-catchments in the water balance module. Also, in this module, the soil is treated in two zones: an active soil layer characterized by active soil– vegetation–atmosphere interaction processes, and the groundwater zone. The active soil layer is further divided into two: an upper thin soil layer where evaporation and transpiration both predominantly occur and a lower soil layer where only transpiration takes place. Excess precipitation runoff, direct runoff from impermeable areas of the basin, rapid subsurface flow (interflow) and base flow contribution from groundwater are the main catchment runoff mechanisms within the GeoSFM model (Asante *et al*., 2008).

The water balance module produces a runoff which is routed in two phases: from the sub-basin level to its outlet, then through the main river channel network. In the latter, the subsurface runoff is routed using a set of two conceptual linear reservoirs whereas the surface runoff routing is enabled through a diffusion wave equation modified for use in a GIS environment (Olivera and Maidment, 1999). To determine the rate at which runoff is transported to the catchment outlet from the generation point, the DEM and the land cover are used. The linear Muskingum–Cunge scheme (Artan *et al*., 2001; Asante *et al*., 2008) is useful in routing the water within the river channel network.

The GeoSFM model has been applied and tested in a range of catchments across the world, such as the Nile River, to assess the effects of land management practices (e.g., irrigation and water abstraction (Mutie *et al*., 2006; Serrat-Capdevil *et al*., 2006), to assess the suitability of remotely sensed data for streamflow simulations) (Artan *et al*., 2007), in flood forecasting and streamflow simulations (Kiluva *et al.*, 2011; Dessu *et al.*, 2016) and climate change and risk assessment (Blanc and Strobl, 2013; . These studies found some comparable results in model simulations and the ability of GeoSFM to effectively account for water balance.

2.4.9. MIKE SHE

The MIKE SHE model is a physically based, surface-subsurface integrated, coupled, fully distributed hydrological model developed by the Danish Hydraulic Institute (DHI; DHI 2002) on the basis of the SHE (Système Hydrologique Européen) code (Abbott *et al*., 1986; Ma *et al*., 2016).. In terms of structure, MIKE SHE represents spatially distributed catchment parameters, climate variables and hydrological processes. This is achieved through an orthogonal grid network and column of horizontal layers at each grid square in the horizontal and vertical, respectively. The flexibility in MIKE SHE's process-based framework allows each process to be solved at its own relevant spatial and temporal scale. Detailed developments in the MIKE SHE model, and associated parameters can be found in Ma et al. (2016). M. The schematic representation and processes of the MIKE SHE model are shown in Figure 2.10.
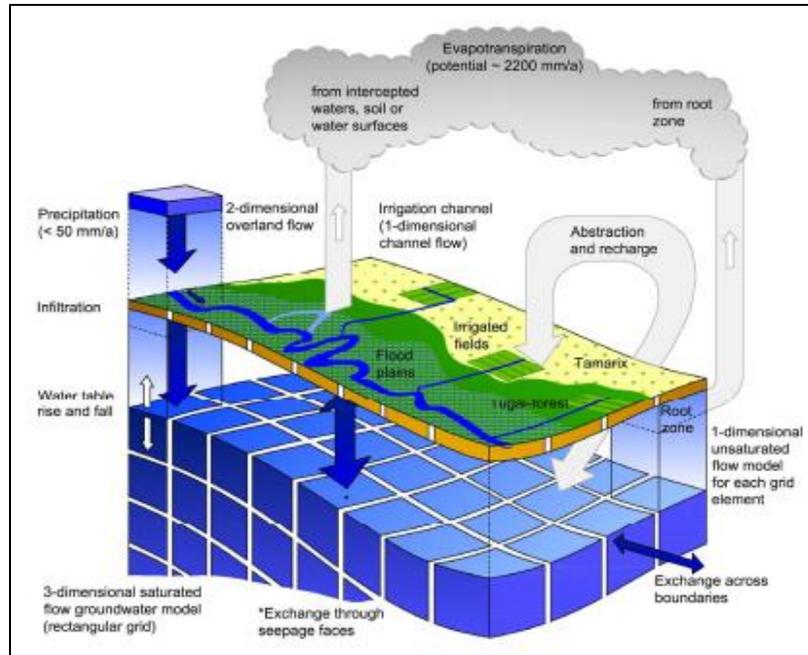
Figure 2.10: Schematic representation of components of a physically based distributed hydrologic model MIKE SHE (Refsgaard, 1995).

The meteorological input data used in the MIKE SHE model include precipitation, air temperature and solar radiation at daily time steps. MIKE SHE has three modules, for precipitation, evapotranspiration and flow. Precipitation and evapotranspiration are represented by processes like interception, drainage, evaporation and uptake of water from the canopy, ponded water and soil. The flow component is divided into unsaturated zone, surface, channel and saturated zone flows. Infiltration, moisture distribution and moisture deficit contribute to the unsaturated zone, whereas the saturated zone is facilitated by the groundwater flow and exchange. The surface and channel flow are accounted for by detention storage, surface runoff and flow routing in rivers and flooding respectively. Each of these processes can be solved at its own relevant spatial and temporal scale (Ma et al., 2016; Refsgaard et al., 1996). MIKE SHE can also combine conceptual and physics-based methods based on data availability (Graham and Butts, 2005).

MIKE SHE has been widely employed in research on hydrology and water resources and modelling the ecohydrological processes that are subject to human activities and nonstationary climate such as land-use change (Im *et al.*, 2009; Viney *et al.*, 2009), wetland management (Thompson *et al.*, 2004; Graham and Butts, 2005), groundwater extraction (Demetriou and Punthankey,1998; Foster and Allen, 2015) 2015) , irrigation-drainage (Jayatilaka et al., 1998; Singh. et al., 1999) (Jayatikala *et al*., 1998; Singh *et al*., 1999), and climate change (Karlsson *et al.*, 2016), as well as the simulation of water quality and soil erosion (Refsgaard *et al*., 1996). These studies found some comparable results in model simulations and the ability of MIKE SHE to effectively account for water balance.

2.4.10. The Variable Infiltration Capacity (VIC) Model

Variable Infiltration Capacity (VIC) is a single layer land surface hydrological model implemented in the Geophysical Fluid Dynamics Laboratory general circulation model (GCM) (Liang *et al.*, 1994). It is a hybrid of physically based and conceptual components, as illustrated in **Figure 2.11**.
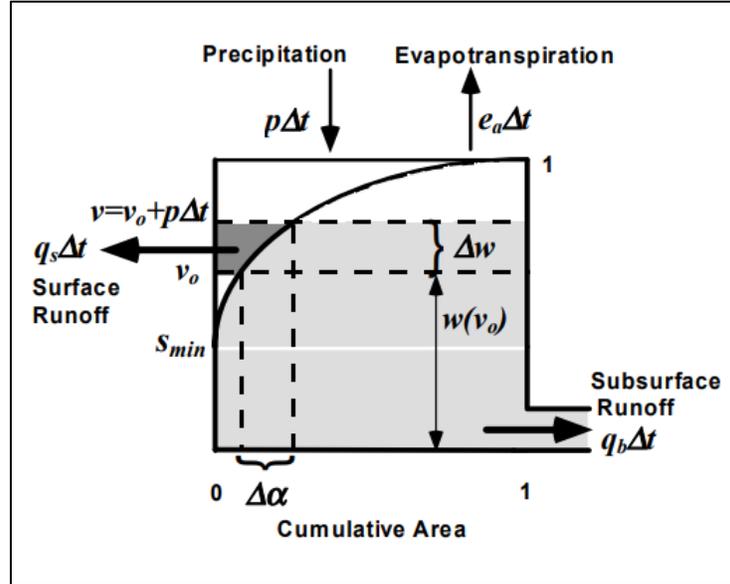
Figure 2.11: Schematic representation of the three-layer structure of VIC (Liang *et al.*, 1994).

The meteorological inputs for the model are daily precipitation and temperature and total evaporation consisting of three components: evaporation from canopy, bare soils and transpiration at daily and hourly timesteps. The latter is represented using an architectural and a canopy resistance formulation. The surface energy balance is iterated to solve for the land surface temperature at each time step from latent heat flux computations (Liang *et al.*, 1994; Cherkauer, Bowling and Lettenmaier, 2003).

The VIC model focusses on runoff processes within each vegetation class characterized by the variable infiltration curve and a representation of nonlinear baseflow. The infiltration curve is a function of parameterization of the effects of sub-grid variability in soil moisture holding capacity, whereas the baseflow is represented using the empirically based Arno baseflow curve. These two distinguishable processes makes it different from other Soil–Vegetation–Atmosphere Transfer schemes (SVATS) (Liang *et al.*, 1994; Cherkauer, Bowling and Lettenmaier, 2003). The routing model facilitates the explicit representation of reservoirs. VIC has numerous model parameter which could not summarised in Table 4 but can be accessed at :-
https://vic.readthedocs.io/en/master/Documentation/Drivers/Image/Params/.

VIC has been applied and tested in a range of catchments across the world, such as the Mekong river, to assess the effects of land management practices (e.g., irrigation and water abstraction (Tatsumi and Yamashiki, 2015), land cover and energy balance (Hengade and Eldho, 2016; Chen *et al.*, 2018) and the effects of climate change on landcover (Lu *et al.*, 2018). These studies found that VIC has difficulties in reproducing observed stream flow in the arid basins attributed to

groundwater-surface water interactions which are not modelled by VIC (it does not include a mechanism to account for deep groundwater recharge and drainage to streams). The model does not have an explicit mechanism to produce infiltration excess flow and it does not represent capillary rise in the soil zone.

2.4.11. LISFLOOD

LISFLOOD is a GIS-based hydrological model designed to simulate all the hydrological processes in a given catchment (van der Knijff, Younis and de Roo, 2010). LISFLOOD is a spatially distributed grid-based rainfall runoff and channel routing model, run using any desired time interval and on any spatial grid size. To simulate the longer-term catchment water balance and individual rapid flood events (e.g., flash floods), LISFLOOD can be run using, respectively, daily and sub-daily time steps (e.g., hours). The schematic representation of the LISFLOOD model is shown in Figure 2.12 and a detailed description of the parameters, variables and key processes can be accessed at https://ec-jrc.github.io/lisflood-model/Lisflood_Model.pdf.
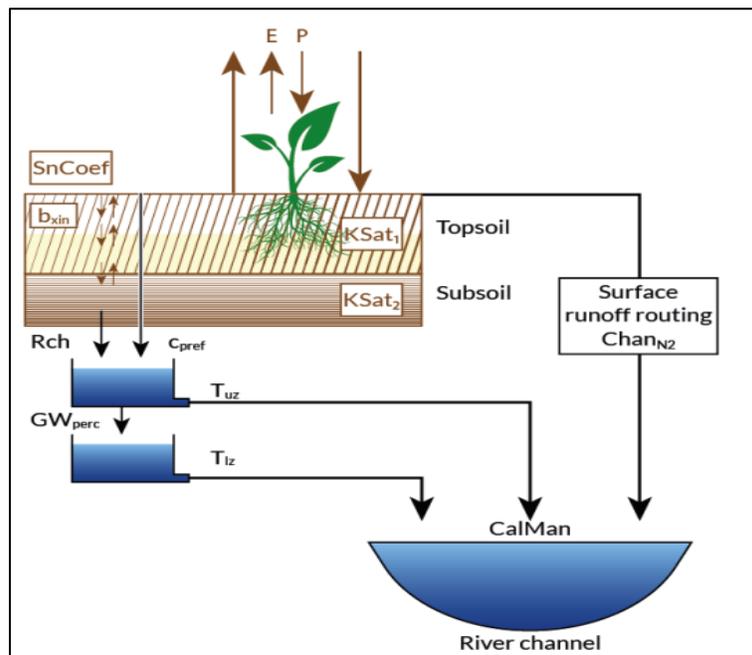
Figure 2.12: Schematic representation of Flowchart of the LISFLOOD model, showing the key processes included (van der Knijff *et al.*, 2010). P: precipitation; E: evaporation & evapotranspiration; S_nC_{oef}: snow melt; b_{xin}: infiltration; $Chan_{N2}$: surface runoff; GW_{perc}: drainage from upper- to lower groundwater zone; T_{uz}: outflow from upper groundwater zone; T_{lz}: outflow from lower groundwater zone; R_{ch}: drainage from subsoil to upper groundwater zone; drainage from top- to subsoil; C_{pref}: preferential flow to upper groundwater zone.

LISFLOOD is driven by the following meteorological variables: precipitation rate, P (mm day^{-1}), potential evapotranspiration rate with a closed canopy, ET_o (mm day^{-1}), potential evapotranspiration rate from the bare soil, ES_o (mm day^{-1}), potential evapotranspiration rate from an open surface, EW_o (mm day^{-1}) and average 24-hour temperature, T_{avg} (ºC). ET_o, ES_o, and EW_o are calculated before running the model. A separate set of equations for pre-processing standard

meteorological observations (e.g., temperature and radiation) and calculating the different potential evapotranspiration rates has been proposed as a companion part to the model.

LISFLOOD has been used in several studies globally. Examples include flash flood forecasting within the Global Flood Awareness System (GloFAS) (Alfieri *et al.*, 2013), flood forecasting for Europe (Bartholmes *et al.*, 2009; de Roo *et al.*, 2011; Pappenberger, Thielen and Del Medico, 2011), assessing water resources (Sepulcre-Canto *et al.*, 2012; Mubareka *et al.*, 2013), climate change impact assessment (Rojas et al., 2012) and recently inr Africa (Thiemig *et al.*, 2015; Bisselink *et al.*, 2016). The study by Bisselink *et al.,* (2016) for example, considered the performance of a differential split-sample test to calibrate the LISFLOOD hydrological model using different precipitation sources, to show differences in model parameters and to ensure a minimum standard for operational validation of this simulation model in the head waters of South Africa. The results indicate large discrepancies in terms of the linear correlation, bias, and variability between the observed and simulated stream flows when using different precipitation estimates as model input. However, the best model performance was obtained with products which were blended with gauge data for bias correction. The study by Thiemig *et al.* (2015) examined the forecasting capacity of the Pan-African Flood Forecasting System (AFFS) which used the LISFLOOD hydrological model to assess its ability to detect and predict flood events and its overall performance to predict streamflow. Results indicate that the system has a good capability of predicting riverine flood events well in advance, whereas it shows limitations for small-scale and short duration flood events.

2.4.12. The European Hydrological Predictions for the Environment (HYPE) model

HYPE is a conceptual semi-distributed process-based model (Lindström et al.,2010 http://hypecode.smhi.se/). The model domain may be divided into sub-basins (Figure 2.13, left), which are either independent or connected by rivers and a regional groundwater flow. Each sub-basin can in turn be divided into classes, which are the smallest computational spatial unit, defined by HRUs.
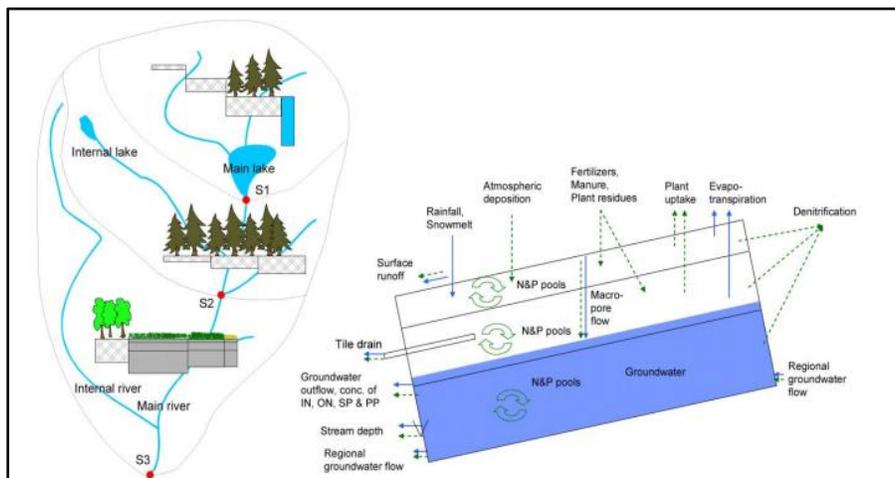
Figure 2.13: Left: schematic division into sub-basins and land classes according to elevation, soil type, vegetation and lake classes. Right: schematic model structure within a land class (i.e. a combination of a soil type and a

crop), simulated using three soil layers. Solid and dashed arrows show fluxes of water and elements, respectively (Lindstrom *et al*., 2010).

The meteorological forcing requirements for the HYPE model are daily precipitation and temperature at daily or sub-daily time steps. The processes represented include snow melt, evapotranspiration, surface runoff, infiltration, percolation, macropore flow, tile drainage, and lateral outflow to the stream from soil layers. Calculations of the flow path based on the above processes for each HRU start in each sub-basin (further detailed in Section 2.3). Each sub-catchment may consist of any combination of HRU types and each HRU can be divided vertically into three distinct soil layers (e.g., top-soil, root zone and the remaining soil). HRUs are connected directly to the stream and act in parallel. Water is routed between sub-basins through the river network and through lakes, reservoirs, and wetlands on its path to the river outlets. The HYPE model has several parameters which may be general across the domain (e.g., river routing parameters) or may depend on the soil type (e.g., field capacity), or land cover (e.g., evapotranspiration coefficient). The values of the parameters are contained in a stepwise manner and are determined in the model set-up procedure. Parameters for a HRU have no connection to locality or catchment but are specific for a given soil type/land cover combination and are then applied anywhere in the domain where that combination of soil type/land cover exists. Detailed process description of the HYPE model can be found in Donnelly *et al*. (2016).

HYPE is used operationally, and different versions have been implemented in Europe (E-HYPE) and in West Africa (Niger-HYPE). For example, the E-HYPE version was applied to European catchments and the results show good model performance of long-term means and seasonality and less good representation of short-term daily variability, especially for Mediterranean and mountainous areas (Donnelly *et al*., 2016).. HYPE was also used in 6,000 sub-basins in India to evaluate flow signatures and performance metrics, using both multiple criteria and multiple variables, and independent gauges for "blind tests" (Pechlivanidis and Arheimer, 2015). The results reveal that despite the strong physiographical gradient over the subcontinent, a single model can describe the spatial variability in dominant hydrological processes at the catchment scale. Exploration of the process improvement in the developed Niger-HYPE revealed that the original model concept could simulate the annual cycle of discharge, but not processes such as precipitation, evaporation, surface runoff, infiltration, soil storage, reservoir regulations, aquifer recharge, and flooding and river-atmosphere exchange, which all required improvements if accurate flood forecasts were to be achieved in the Inner Niger Delta. In Anderson *et al*. (2015), the application of Niger-HYPE for flow signature statistics reveals the forecasting capacity of the model over the basin with reasonable deviations of about 17%.

2.4.13. Summary classification of the reviewed Models

The models under consideration are summarized in Figure 2.14, a four quadrant Cartesian plane with colour coded distributed drops representing different categories of models outlined in Section 2.2 in each quadrant. Blue coded drops represent fully lumped conceptual models (e.g., GR4J, PDM, and MIKE NAM). The green represents semi lumped conceptual models (e.g., SMAR). Yellow represents semi-distributed, physics-based models such as GeoSFM, whereas the grey drops are for the semi-distributed conceptual models such as HBV-96. The black category is for

the distributed conceptual, models such as the SWAT, VIC and HYPE, whereas the pink shades are for the fully distributed physics-based models such as LISFLOOD, MIKE SHE and TOPMODEL.
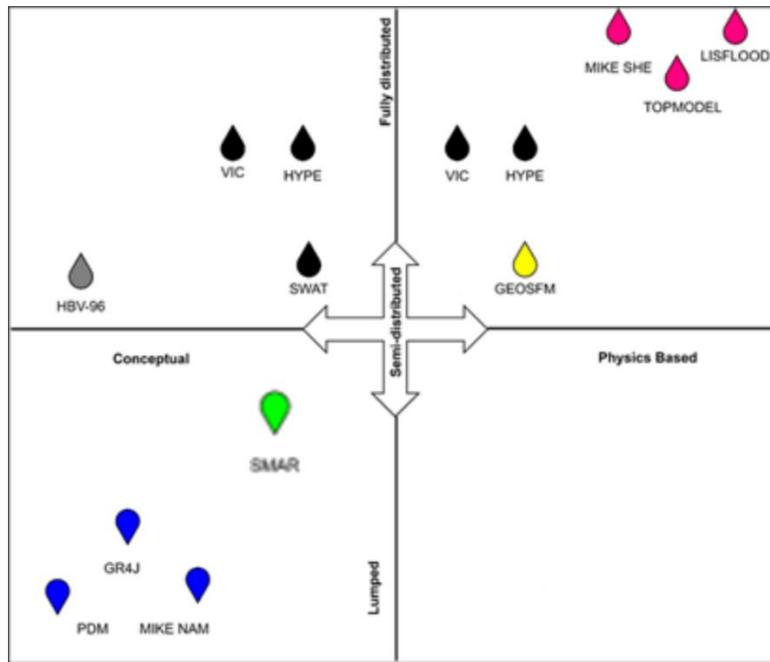
Figure 2.14: Categorical summary of the reviewed models following the model classification in Section 2.2.The colour coded distributed drops representing different categories of models outlined in each quadrant: fully lumped conceptual (blue), semi-lumped conceptual (green), semi-distributed, physics-based (yellow), semi-distributed conceptual (grey), distributed, conceptual (black), and fully distributed physics-based (pink).

The models reviewed have been applied in hydrological studies over several catchments in different parts of the world and the results show variations in model performance in different catchments of different characteristics. This points to the complexity of modelling the physical processes operating in different environments and the model parameters required to capture geographical heterogeneity. It is therefore necessary to compromise when using models for operational applications in different geographical areas. It can be inferred that settling on an appropriate tool for hydrological modelling and flood forecasting in Kenya is a complex and challenging task, hence the need for a decision framework, which is further discussed in Chapter 4 of this thesis.

Some studies suggest that most models may fail to provide good representations of processes such as runoff, soil moisture and other hydrogeological features, which are important process in flood modelling and forecasting (Andersson *et al.*, 2015). Furthermore, , ground water flow and surface and sub-surface flows are relevant to water balance in catchments, buthowever, some studies have shown that several models fail to represent them (e.g., Nijssen *et al.*, 1997; Trambauer *et al.*, 2013). For example, VIC has difficulties in reproducing observed stream flow in the arid catchments, attributed to groundwater-surface water interactions which are not modelled by VIC (it does not include a mechanism to account for deep groundwater recharge and drainage to streams). The model does not have a specific mechanism to produce infiltration excess flow and it

does not represent capillary rise in the soil zone. Moreover, the processes responsible for channel losses are not represented by the routing component. Its distributed physical nature allows independence of long time period gauge data, and complete integration of overland and ground water. However, as noted by Abbot *et al*. (1986), Beven (2010) and Ma *et a*l. (2016), 1986; Beven, 2010; Ma et al., 2016) distributed models have several shortcomings. For example, numerous specific hydraulic properties, complex spatial distribution of environmental factors, complicated physically based algorithms for the hydrological processes and a strong nonlinear relationship between parameters and simulation results are common.

At the catchment scale, significant factors affect hydrological processes, including latitude, temperature, topography, geology and land use (Thompson, 2017). Therefore, modelling is an essentially probabilistic application, with uncertainty intensified at each stage of the process, from scenario generation to issues of scale, simulation of hydrological processes, and management impacts (Jayathilake, 2019). The four main sources of uncertainty include input uncertainty, parameter estimation uncertainty, structural uncertainty and initial conditions uncertainty (Renard *et al.*, 2010). Uncertainty induced by model structures can be more significant than uncertainty from parameter estimation and input data uncertainty, but such uncertainties are difficult to assess explicitly or to separate from other uncertainties during the calibration process (Beven et al., 1992), hence the need for a comprehensive uncertainty and sensitivity analysis in modelling. A detailed discussion of the concept of uncertainty and sensitivity analysis in hydrological modelling appears in Section 2.6.1 of this thesis.

Additionally, the uncertainty in all the forcing data (mainly precipitation) is also an important issue that cannot be overlooked. Even a perfect model, if forced with biased precipitation, will fail to produce accurate representations of runoff, soil moisture and other hydrological fluxes. Data scarcity issues, discussed in the later sections, impose massive limitations on the proper identification of the limitations of each model. In this respect, many remotely sensed and reanalysis products suffice as the best alternatives to observations and many studies have evaluated their potential for hydrological modelling and applications in many parts of the world, including ungauged Kenyan catchments (Le and Pricope, 2017). Details of input data into hydrological models and their evaluation are discussed in Section 2.5 of this thesis.

Table 4:- Description of the some of the model parameters, variables and the physical meaning of symbols mentioned in in the intext in overview of the individual models. Further model references and institutions (developers) can be found in Table 8.

Model	Parameter	Parameter description	Variable description	Variable
GR4J (Perrin *et al.*, 2003)	X_1 X_2 X_3 X_4	Capacity of the production store (mm). Groundwater exchange coefficient (mm) Capacity of the nonlinear routing store (mm). Hydrograph time base (days).	Areal catchment rainfall Percolation leakage Net rainfall Total quantity of water reaching the routing function Amount of net rainfall that goes directly to the production store Amount of net rainfall that goes directly to the routing store Potential areal evapotranspirationNet evapotranspiration capacity Actual evaporation rate Groundwater exchange term Water content in the production store Water content in the routing store Total stream flow Output of UH2 Output of UH1 Routed flow component Direct flow component Unit hydrographs	*P* (mm) *Perc* () Pn Pr Ps Pn – Ps *E* (mm) En Es F*(X$_2$)* S (mm) EP (mm) R Q_1 Q_9 Q Q_r Q_d *UH1, UH2* (days)
NAM Nielsen & Hansesn, 1973)	U $_{max}$(mm) L $_{max}$(mm) CQ $_{OF}$ (-1) CK$_{IF}$ (h) CK$_{1,2}$ (h) TIF (-) TOF (-) CKBF(h) TG (-)	Max. water content in the surface storage Max. water content in the lower zone storage Overland flow runoff coefficient Time constant in the interflow Time constant of the overland interflow flow Threshold value for interflow Threshold value for overland flow Time constant of baseflow Threshold value for ground water storage	Surface storage water content Lower storage soil water content Actual evapotranspiration Potential evapotranspiration Infiltration root zone Groundwater recharge Excess rainfall	U(mm) L(mm) E $_{act}$ (mm) E $_{pot}$(mm) ΔL (mm) G (mm) P_N(mm)
PDM Goswami *et al.*,201o)	*cmax* *b* *mfl* *rtq* *rts* *%(q)* *bypass*	Maximum storage capacity of the catchment (mm) Degree of spatial variability of storage capacity in the catchment (-) Snow retention factor (-) Residence time quick flow reservoir (dt) Residence time slow flow reservoir (dt) Percentage flow through quick flow (-) Bypass of the rainfall to routing components (-)	Precipitation Potential evaporation Surface runoff Baseflow Moisture storage Surface storage Groundwater storage	*P* (mm) *E* (mm) q$_s$ q$_b$ S_1 S_2 S_3
HBV- 96 (Lindström *et al.*, 1997)	F$_C$ L$_P$	Maximum soil moisture (mm) Soil moisture threshold for reduction of evapotranspiration (-)	Snow Rain Evapotranspiration	*SF* *RF* (mm) *EA*

Model	Param.	Parameter description	Flux/variable	Symbol
	K$_0$	Shape coefficient (-)	Snow cover	*SP*
	K$_1$	Near surface flow coefficient (1/dt)	Infiltration	*IN*
	K$_2$	Recession coefficient for lower tank (1/dt)	Recharge	*R*
	UZL	Near surface flow threshold (mm)	Soil moisture	*SM*
	PERC	Maximum flow from upper to lower groundwater storage (mm/dt)	Capillary transport	*CFLUX*
	MAXBAS	Transfer function parameter (day)	Storage in the lower reservoir	*UZ*
				PERC
			Fast runoff component	*LZ*
			Slow runoff component	*Q$_0$*
			Total runoff	*Q$_1$*
SWAT (Arnold *et al*., 1998) (Only streamflow parameters listed)	ALPHA_BF Ch _K2 Ch _N2 CN2 ESCO GW_DELAY GWQMN GW_REVAP REVAPMN SOL_AWC* SOL_K* SOL_Z* OV_N* HRU_SLP	Baseflow alpha factor (d^{-1}) Effective channel hydraulic conductivity (mm h^{-1}) Manning coefficient of main channel SCS runoff curve number for moisture condition II Soil evaporation compensation factor Groundwater delay (d) Threshold depth of water in the shallow aquifer required for return flow to occur (mm) Groundwater 'revap' coefficient Threshold depth of water in the shallow aquifer for revap to occur (mm) Available water capacity of the soil layer (mm mm^{-1}) Soil saturated hydraulic conductivity (mmh^{-1}) Depth from soil surface to the bottom of layer (mm) Overland Manning roughness Average slope steepness (m^{m-1})	Precipitation Surface runoff Lateral flow Percolation Revap Groundwater flow Recharge	*p*
SMAR (O'Connor, 2005)	T H Y C G N NK K*g* Z	Conversion to potential rate Direct runoff coefficient Rainfall in excess of infiltration capacity Evaporation decay Moisture in excess of soil capacity Nash gamma shape parameter (linear routing) Nash gamma scale parameter (linear routing) Linear reservoir routing coefficient Soil moisture storage	Precipitation Evapotranspiration Evaporation Excess rainfall Direct runoff Lateral flow from in excess infiltration moisture Lateral flow from in excess soil moisture Total generated surface runoff	*R* (mm) *E* (mm) *C* (mm) *X* (mm) r$_1$ r$_2$ r$_3$ r$_4$
TOPMODEL Beven & Kirby, 1979)	M Ln (T$_o$, (m)) SR$_{max}$ SR$_{int}$ ChVel	Exponential transmissivity function Natural logarithm of effective soil transmissivity Soil profile storage Initial storage deficit in root zone Routing velocity (network width function)	Precipitation Evapotranspiration Root zone storage Storage in the unsaturated zone Drainage flux from the unsaturated zone to groundwater Groundwater flow Base flow	R (mm) E (mm) S$_{rz}$ (m) S$_{uz}$ (m) q$_v$(mg^{-1}) q$_{of}$ q$_b$
VIC (Gao *et al*., 2010)	*b* *S$_{min}$* *h* *y* *k$_c$*	Parameter controlling the curvature of the storage distribution Minimum storage required for saturated area formation. Evaporation exponent; property of soil and vegetation types Capillary fringe thickness Baseflow recession coefficient	Precipitation Evapotranspiration Surface runoff Subsurface runoff Scaled soil moisture Total soil moisture	*p* (mm) *e$_a$* *q$_s$* *q$_b$* *v* *w*

Model	Parameters	Processes	Variables	Symbols
LISFLOOD (van der Knijff et al., 2010)	S_nC_{oef} b_{xin} $Chan_{N2}$ GW_{perc} T_{uz} T_{lz} R_{ch} C_{pref}	Snowmelt Infiltration Surface runoff Drainage from upper to lower groundwater zone Outflow from upper groundwater zone Outflow from lower groundwater zone Drainage from subsoil to upper groundwater zone Preferential flow to upper groundwater zone	Precipitation Evaporation Potential evapotranspiration rate with closed canopy Potential evapotranspiration rate from bare soil. Potential evapotranspiration from open surface	P (mm d-1) E
GeoSFM (Artan et al., 2001; 2004;2007)	Multi-parameter, detailed in Artan et al., (2001)	Detailed model parameters available at https://pubs.usgs.gov/of/2007/1441/pdf/ofr2008-1441.pdf	Evapotranspiration Shallow and deep ground water exchange Snow cover Sub-soil interchange Linear lake reservoir Runoff Routing Energy balance	
HYPE (Lindstrom et al., 2010)	Multi-parameter, detailed in Lindstrom et al., (2010)	Detailed model prameters available at http://hypecode.smhi.se/	Precipitation Surface runoff Interception Evapotranspiration Ground water exchange Routing Groundwater flow	*No symbols used in the text*
MIKE SHE (Abbott et al 1986)	Multi-parameter, detailed in Keilholz et al.,(2015	Multi-parameter, detailed in Keilholz et al.,(2015	Precipitation Evapotranspiration Interception Runoff Infiltration Exchange through seepage Ground water exchange Routing Snow cover Sub-soil interchange Energy balance	*No symbols used in the text*

2.5. Input datasets for hydrological modelling

Data play an important role in hydrological modelling irrespective of the processes represented in the model. These datasets can take different forms, so that some variables are considered to vary with time (e.g., rainfall, air temperature) and some are considered to be static in time (e.g., soil characteristics, topography, vegetation, hydrogeology and other physical parameters) (Devia, Ganasri and Dwarakish, 2015). However, like other components of the modelling chain, data has its own inherent unquantified uncertainties and errors which may cause problems in modelling. Moreover, data scarcity poses challenges in modelling, an issue which has been highlighted in some studies (e.g., Fuka et al., 2014; Beck et al., 2017; Wanzala et al., 2022). As a result, there are several global or near-global remotely sensed satellite precipitation datasets at different spatial-temporal resolutions (Fortin et al., 2015; Sun et al., 2018), which have been used as substitutes

for *in situ* observations in ungauged catchments. This has in turn made precipitation evaluation studies an important topic of discussions within research communities.

Ground validation of precipitation rainfall products is a challenging task, both in terms of assessment and ground observation methodologies. Of primary concern is the quality of the ground truthing methodologies and their dependency on the spatial and temporal resoulutions of data (Gribbon and Bailey, 2004; Vischel *et al.*, 2011; Hu *et al.*, 2016). Several studies attempt to quantify and account for the sampling uncertainties when satellite and reanalyses are compared to observations (e.g., Tarek, Brissette and Arsenault, 2021). Measurement and sampling of rainfall has been hampered by its high spatiotemporal variability and its intermittent nature, especially in the Tropics, where most of the rainfall is controlled by the mesoscale convective systems. This is also compounded by the general scarcity of the operational gauge networks, except for a few dense networks operated by dedicated organizations or programmes (Nicholson et al., 2019), e.g., the African Hydromet Programme at
(https://www.worldbank.org/en/programs/africa_hydromet_program) and the African Group on Earth Observations (AfriGEO, https://earthobservations.org/afrigeo.php).

Like many other places in the world, Kenya suffers from several severe problems with observational data scarcity. The available precipitation gauging stations are sparse and/or not available at all (Figure 2.16). Most of the catchments studied in Chapters 4, 5 and 6 have either one or two stations, which leads to very sparse coverage. Not only that, but the data series is short and includes massive gaps. For example, the Tana River catchment, one of our study catchments is 96,000 km^2 but has only four precipitation gauging stations. These stations are unreliable and have large data gaps (Figure 2.15).
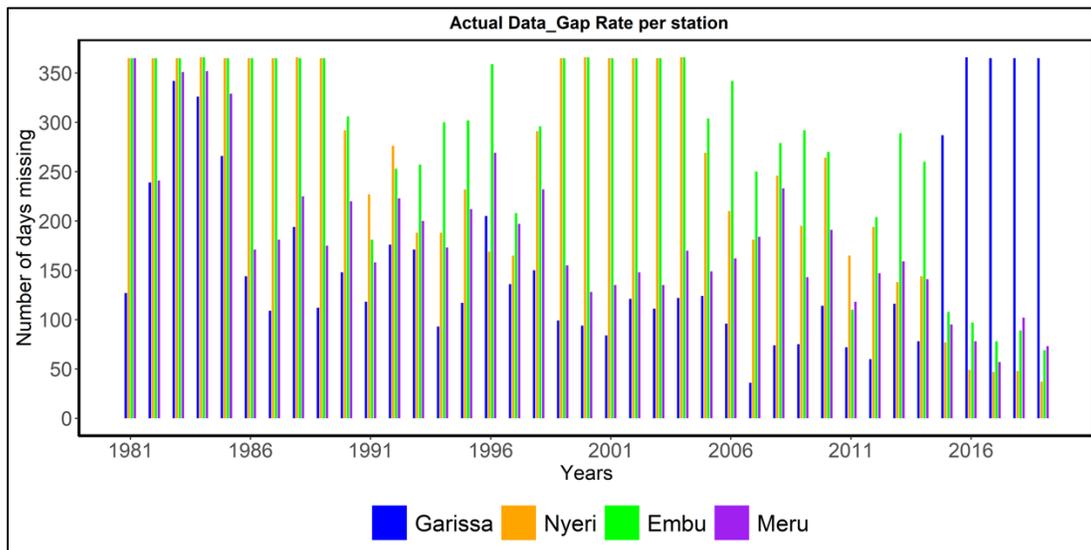
Figure 2.15: Unreliability of the Tana precipitation datasets: missing data in days for the four rainfall gauging stations in Tana catchment.
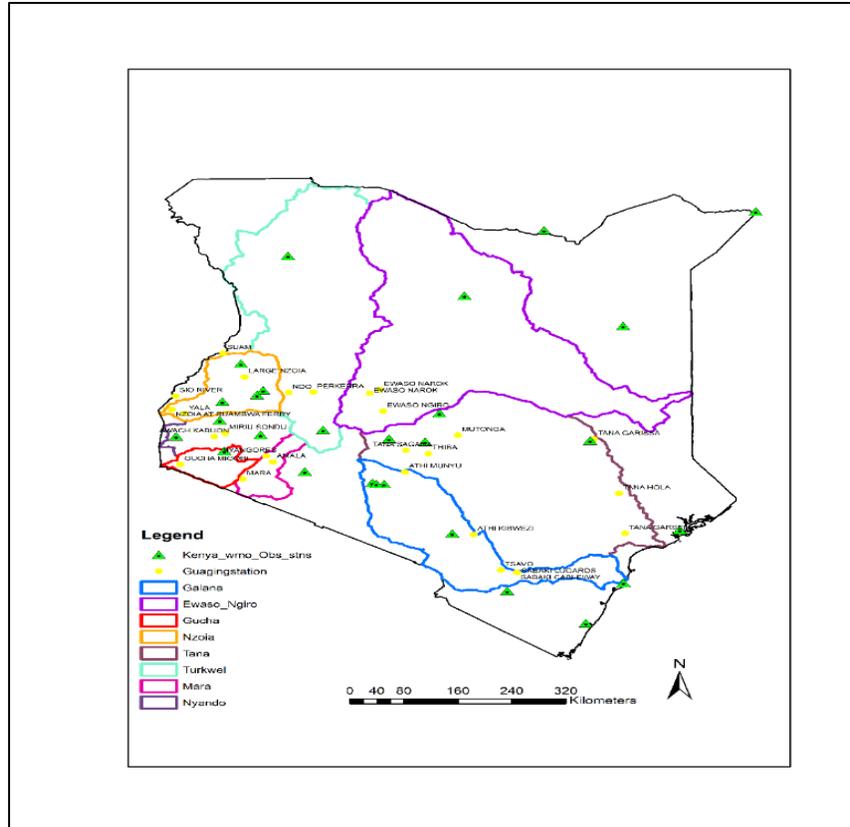
Figure 2.16: Locations of the rainfall gauging stations in each of the catchments of study.

2.5.1. Evaluation of precipitation datasets at global scale

Several precipitation evaluations for hydrological applications have been carried out globally and regionally at basin scale level. Most of these studies considered only a few precipitation datasets, did not consider reanalysis-based precipitation datasets, or did not recalibrate the hydrological model for each precipitation dataset (e.g., Bitew *et al.*, 2012; Tang *et al.*, 2020). However, some studies have used global reanalysis weather data for various hydrological applications and yielded meaningful results (Lavers *et al.*, 2012; Najafi, Moradkhani and Piechota, 2012; Quadro *et al.*, 2013; Smith and Kummerow, 2013; Fuka *et al.*, 2014). For example, Lavers *et al*. (2012) used five atmospheric reanalysis products — CFSR, ERA-Interim, 20th Century Reanalysis (20CR), MERRA, and NCEP-NCAR (National Center for Atmospheric Research) — to detect atmospheric rivers and their links to British winter floods and large-scale climatic circulation. Their study provided valuable evidence of generally good agreement on atmospheric river occurrences between the products. Quadro *et al.* (2013) evaluated the hydrological cycle over South America using CFSR, MERRA, and the NCEP Reanalysis II (NCEP-2). They observed general agreement in precipitation patterns among the three products and the observed precipitation over much of South America. They reported that the CFSR precipitation showed the smallest biases. Wu, Chen and Huang (2013) used the CFSR dataset to study the water budgets of three tropical cyclones that passed through the Taiwan Strait. They assessed the quality of CFSR for tropical cyclone studies by comparing CFSR precipitation data with TRMM precipitation data. They concluded that the CFSR data were reliable for studying tropical cyclones in this area.

43

2.5.2. Evaluation of precipitation datasets at regional scale: - Eastern Africa

In Eastern Africa, several studies have evaluated precipitation datasets at different spatial and temporal scales. These include Dinku *et al.* (2007; 2008;2011); Hirpa *et al.* (2010); Romilly and Gebremichael, (2011); Worqlul *et al.* (2014); Young *et al.* (2014); Maidment *et al.* (2013; 2014); Diem *et al.* (2014); Awange *et al.* (2016); Maidment *et al.* (2017) and Dinku *et al.* (2019). However, none of these studies has evaluated the potential of these products for hydrological applications or incorporated a comparison of reanalysis datasets. For example, Dinku *et al.* (2018) carried out an evaluation of CHIRPS at daily, dekadal (10-day) and monthly time-scales in East Africa for Ethiopia, Kenya and Tanzania by comparing the satellite products with rain-gauge data from about 1,200 stations but did not evaluate the reanalysis precipitation datasets and did not consider catchment scale evaluation for hydrological applications.

For hydrological applications, studies have been carried out on several catchments in East Africa. Most of these studies just compared one reanalysis dataset with the gauge observations in simulating catchment runoff. For example, a study by Artan *et al.* (2007) used GeoSFM, a physically-based semi-distributed hydrologic model, to show the significance of satellite precipitation datasets in simulating streamflow over four basins in East Africa. Results show the usefulness of remotely sensed precipitation data for hydrologic modeling when the hydrologic model is calibrated with such data. However, they note that the remotely sensed rainfall estimates cannot be used confidently with hydrologic models that are calibrated with rainfall measured by rain gauges, unless the model is recalibrated. Dile and Srinivasan (2014) compared the performance of Climate Forecast System Reanalysis (CFSR) with that of gauge observations in simulating observed streamflow at four river gauging stations in the Lake Tana basin, the upper part of the Upper Blue Nile basin, using the SWAT model. They found out that the gauge observations perform better than the CFSR data, but recommend the use of satellite data in data-scarce regions where there are few gauge observations. Worqlul *et al.* (2017) evaluated CFSR and TRMM Multisatellite Precipitation Analysis (TMPA) 3B42 version 7 as input to HBV and Parameter Efficient Distributed (PED) hydrological models. They found that both the gauged and the CFSR reanalysis data were able to reproduce the streamflow well for Gilgel Abay and Main Beles in the Ethiopian highlands.

2.5.3. Evaluation of precipitation datasets at country scale - Kenya

Kenyan basins are poorly gauged, according to the quantity, spatial distribution and quality of precipitation data (Le and Pricope, 2017). A low quantity and spatial distribution of rainfall gauge stations can cause overgeneralization and inaccurate quantification of water availability, while unreliable or incomplete datasets can be unable to identify seasonal or larger range temporal patterns correctly, or at all; they can alsor miss key events entirely. In such cases satellite and or reanalysis precipitation datasets are used in data-scarce areas, which proves the usefulness of evaluating these precipitation datasets. The role of reanalyses in climate and land monitoring applications is now widely recognized. For instance, ECMWF's reanalysis ERA5 is periodically used, together with other datasets, as input to the WMO annual assessment of the State of the Climate, routinely presented at the Conference of the Parties of the United Nations Framework Convention on Climate Change (UNFCCC, WMO, 2021). Reanalyses are also a resource used to

produce Essential Climate Variables (ECVs) and Climate Indicators recommended by the Global Climate Observing System (GCOS, Lavergne *et al.*, 2022). By optimally combining both satellite and conventional observations with a high-quality forecast model, reanalyses indeed provide consistent "maps without gaps" of ECVs and strive to ensure integrity and coherence in the representation of the main Earth system cycles (e.g., water, energy and carbon) (Muñoz-Sabater *et al.*, 2021; Rabier *et al.*, 2021). The consistency shown by reanalysis over several decades provides justification for using both tools to support the monitoring of heatwaves, droughts, or other extreme events such as floods. They are used to create observational and reanalysis-based climatologies that serve as a reference to detect anomalies in time series, either in the study of past events or in the prediction of future extremes (Hoffmann *et al.*, 2019).

Khan *et al.* (2011) evaluated TRMM-3B42 V6 using a distributed hydrologic model, Coupled Routing and Excess STorage (CREST), to simulate the spatiotemporal variation of water fluxes and storage. Results showed the potential of TRMM-3B42 V6 not only for the investigation of water balance but also for addressing issues pertaining to sustainability of water resources at the catchment scale. Le and Pricope (2017) evaluated the performance of the CHIRPS dataset in simulating streamflow using SWAT (see Section 2.4.3) over the Nzoia Basin in western Kenya. Comparisons of results between estimation of streamflow using *in situ* rainfall gauge station data, the CFSR and the CHIRPS dataset showed simulated streamflow estimates based on rainfall gauge station data were poor but improved significantly with the CFSR and CHIRPS datasets.

Most of the findings in the highlighted studies show variations in the performance of the different precipitation datasets in relation to geographic location. It is therefore important to assess the performance of precipitation datasets at catchment scale, due to inherent stark differences in catchment characteristics, whilst communicating uncertainties associated with each of the datasets and how they influence model parameters. To provide background information on model parameterization and uncertainty, Section 2.6 outlines the concept of hydrological model parameters, delving into uncertainty (2.6.1) and sensitivity analysis (2.6.2) as well as non-stationarity and equifinality in hydrological models (2.6.4, & 2.6.5).

2.6. Hydrological model parameterization

A hydrological model parameter is an internal model configuration variable which plays an important role in model simulation (Bárdossy and Singh, 2008; Kavetski, 2019). Beven (2001) defines model parameters as typically integral coefficients built into the structure of models to define the characteristics of the catchment area or flow domain. Model parameters are usually assumed to be constant in time, although the influence of human activities and temporal variations in climatic conditions may mean model simulations would benefit from continuously varying parameters (although this makes heavy demands on data and computational resources) (Patil and Stieglitz, 2015). For example, urbanization and dam construction may result in underlying surface changes, leading to changes in model parameters which represent the transient catchment characteristics (Legesse, Vallet-Coulomb and Gasse, 2003). Also, while calibrated parameters are expected to correct for model structure and observational data issues, model parameters may possibly vary with climatic temporal fluctuations (e.g., timing and frequency) (Merz, Parajka and Blöschl, 2011). Thus, it is important to consider the effects of such outcomes when estimating

hydrological model parameter values and carefully communicate the uncertainties: this is further discussed in Section 2.6.1.

Physically based model parameter values are usually derived, whenever feasible, from accessible field data (Refsgaard and Knudsen, 1996) and parameter estimation values are then typically evaluated in relation to physically acceptable ranges. In general, parameter variables within a certain range follow some specified probability distributions, whose shape can significantly be significantly affected by the parameter range (Xing *et al.*, 2018), but also the nature of the parameter (Cloke, Pappenberger and Renaud, 2008). For example, parameter values of a physically based catchment scale model can be fully determined but may not be right, owing to experimental restrictions and scaling issues, such as discrepancies in measurement and model grid scale, creating the need for time-to-time adjustments (Madsen, 2003).

2.6.1. Concept of uncertainty

Uncertainty arises from the intrinsic structural complexity of the modelling system and errors and assumptions within the entire modelling system, and is thus an inevitable element of hydrological processes (Boelee *et al.*, 2017). Uncertainties can be characterized as Natural uncertainty (also known as aleatory uncertainty, inherent variability and type-A uncertainty) and Epistemic uncertainty (also known as lack-of-knowledge uncertainty, ignorance and type-b uncertainty (Merz and Thieken, 2005). In catchment runoff estimation, epistemic uncertainties may relate to spatial heterogeneity in rainfall, transformation of variables (e.g., discharge rating curves) or lack of knowledge of boundary conditions (e.g., losses to deep water groundwater) (Beven, 2011, 2012). Natural uncertainty can be corrected for by statistical probabilistic methods, but in most cases it may be hard to identify a suitable model to apply, whereas for epistemic uncertainties, where a formal source of error is unknown, statistical methods can be used only to provide an approximation (McMillan, Krueger and Freer, 2012).

Baldassarre and Montanari (2009) summarize the four main sources of uncertainty in hydrologic modelling as: (i) structural uncertainty, (ii) input uncertainty, (iii) parametric uncertainty, and (iv) output uncertainty (Renard *et al.*, 2010). Structural uncertainty may arise because of the inability of hydrological models to perfectly schematize the physical processes (Gupta, Beven and Wagener, 2006; Refsgaard *et al.*, 2006), which include unknown defined hydrological processes (perceptual model) and description of processes (conceptual model) (Zhang *et al.*, 2011) and their mathematical implementation. Uncertainty induced by model structures can be more significant than parameter and input data uncertainty, but such uncertainties are difficult to assess explicitly or to separate from other uncertainties during the calibration process (Beven *et al.*, 1992). The identification of the most appropriate model and model structure, with its associated uncertainty, to be implemented in a flood modelling system is crucial, since the acceptable reproduction of hydrological processes builds reliability into the hydrological model. Input uncertainty arises from approximation in the observed hydrological variables used as input or calibration/validation data (e.g., rainfall, temperature, and river). Parameter uncertainty is induced by imperfect calibration of hydrological models. The schematic representation of common sources of uncertainty and errors is illustrated in (Figure 2.17).
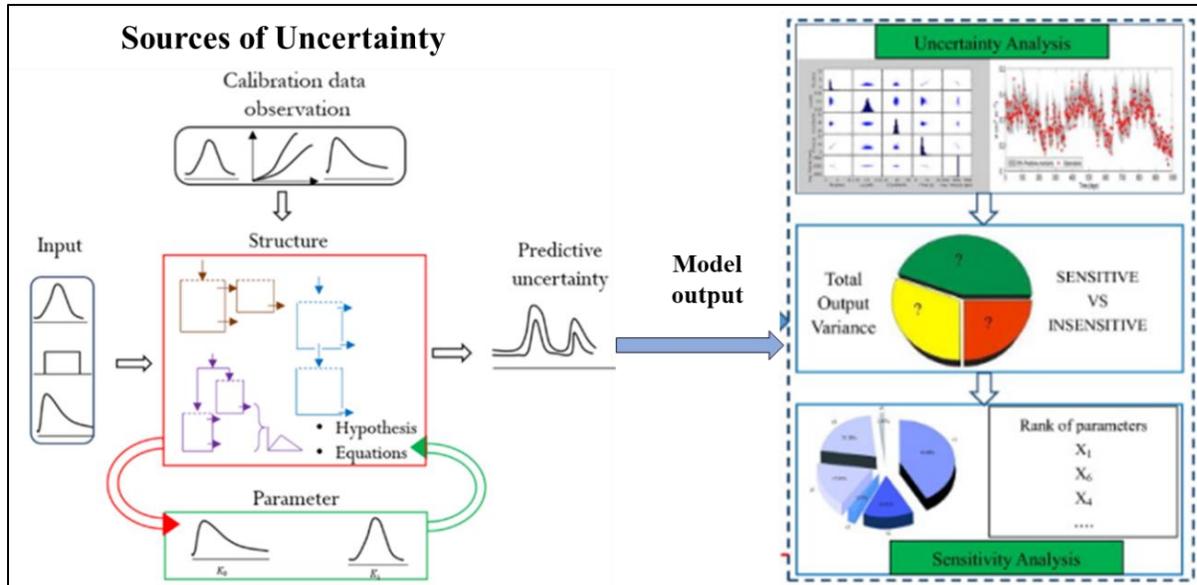
Figure 2.17:- Sketch for the relationship between sources of uncertainty (Source: Moges *at al*., 2020) and the process of uncertainty and sensitivity analysis (Source: Song *et al.,* 2015) of model parameters.

2.6.2. Sensitivity analysis of model parameters

Models often include substantial uncertainties with respect to the input data, forcing data, initial and boundary conditions, model structure, and parameters due to a lack of data and poor knowledge of hydrological response mechanisms (Gupta, Beven and Wagener, 2006). However, more complex models could have a serious interaction between parameters, causing uncertainty in the results (Herrera, Marazuela and Hofmann, 2022). As a result, a wide range of optimization algorithms have been developed to resolve this problem (Beven and Binley, 1992; Zhang *et al.*, 2009; Qi *et al.*, 2016). Obtaining an efficient optimization by including all model parameters in the calibration process may, however, be unfeasible (Nossent, Elsen and Bauwens, 2011). This problem can be solved by sensitivity analysis (SA).

A sensitivity analysis (SA) allows a reduction of the number of parameters incorporated in the optimization by determining the most influential parameters of a model (ranking) (Saltelli, Tarantola and Campolongo, 2000a). For example, this may be done by by setting non-influential parameters to a fixed value, i.e. factor fixing (FF), or focusing on the parameters that have the potential to maximally reduce the output uncertainty, i.e. factor prioritization (FP) (Sobol', 1990; Saltelli *et al.*, 2004). Besides, a SA also facilitates the understanding and interpretation of models by providing information such as the use of the parameters, the influence of specific parameter values (also known as mapping) and the associated model processes and model outcomes (also known as screening) (Saltelli, Tarantola and Campolongo, 2000a). As a result, models are said to be structurally sound in events where most or all of the parameters are sensitive: therefore, modellers can trust a model result (Shin and Kim, 2017). This, however, varies depending on the study areas (Shin 2013), hence the need for a SA in studies incorporating hydrological models and new data.

SA is classified into two categories:- global and local methods (Song *et al.*, 2015; Pianosi *et al.*, 2016). Global sensitivity analysis (GSA) involves evaluating the sensitivity of a model for a large variation of input, *i*, (Saltelli *et al.*, 2008). GSA is useful in identifying the most significant parameters and regions in the parameter space where the model produces extreme values and can be implemented prior to model calibration (van Griensven *et al.*, 2006). Several GSA methods exist whose details can be found in Saltelli *et al.* (2008). However, most are computationally intensive, especially in cases where there are many input parameters. GSA methods dependent on random or probabilistic sampling of the parameter space can be applied to a given mathematical model with multiple inputs to produce multiple outputs (Pianosi *et al.*, 2016). This can be achieved by defining the locations of the parameter space which maximize the amount of information that can be extracted from an optimal model input, e.g., the maximum and the minimum values of the model outputs. On the other hand, local sensitivity analysis (LSA) is based on computation of the sensitivity matrix, also known as the Jacobian, J. Model sensitivity with respect to each of the parameters corresponds to the coefficient J: a detailed explanation can be obtained from Devak and Dhanya (2017). Where there is interaction between model parameters, global SA methods are appropriate as this may result in a non-linear output (Saltelli *et al.*, 2010). In recent years, this broad range of possibilities has made sensitivity analysis a main area of research in hydrological modelling (e.g., Arlimasita and Lasminto, 2020; Rafiei Emam *et al.*, 2018).

2.6.3. Model parameter estimation

Kavetski (2019) outlines two main model parameter estimation strategies, namely: a priori estimation and calibration. A priori estimation refers to establishing parameter values from measured physical system properties, presupposing that the model parameters have a sufficiently reliable representation (Beven and Pappenberger, 2003). Parameter estimation in models of natural systems may therefore require measurements and tests. For example, in assessing the hydraulic conductivity of soils using MODFLOW (Harbaugh, 2005), a physically based groundwater model, parameter estimation may be done using laboratory analysis of core samples in *in situ* tests and/or geology maps (Fetter, 2018). The a priori estimation method is effective in modelling well instrumented locations using environmental physical equations (Clark *et al*., 2015). However, when this method is applied to conceptual models, it becomes hard to relate their parameters to available information (Duan *et al*., 2006), though in some cases, it is possible to deduce useful relationships (Samaniego, Kumar and Attinger, 2010). The spatial variability of vegetation and soils within and across catchments (Miller and White, 1998) and the frequent non-commensurability of modelled and observed quantities (Kuczera, 2002) have made it difficult to estimate model parameters through catchment characteristics. With advances in physical process representation, it has been suggested that models can only be described as "truly" physical if their parameters are specified independently from observed responses (Grayson, Moore and McMahon, 1992). Thus, any hydrological models typically require calibration of model parameters to some degree (Beven, 1989; Duan *et al*., 1992), due to the heterogeneity of hydrologic processes and the scale-dependence of parameters. Parameter variables within a certain range follow some specified probability distributions, whose shape can significantly be affected by the parameter range (Xing *et al.,* 2018) but also the nature of the parameter (Cloke *et al.,* 2007). Model calibration aims at

parameter estimation values which are typically evaluated in relation to physically acceptable ranges.

Model calibration is carried out by adjusting model parameters manually or automatically until the observed and simulated catchment responses match as closely as possible (Moriasi *et al.*, 2012). Manual calibration visually compares the observed and simulated outputs, whereas automatic calibration is commonly performed using optimization techniques/algorithms that ensure the best possible match between model outputs and observations to measure the goodness-of-fit of model simulations (Arnold *et al.*, 2012; Moriasi *et al.*, 2012). A good automatic parameter estimation methodology has four elements: i) an objective function; ii) an optimization algorithm; iii) termination criteria and iv) calibration data. An objective function (goodness-of-fit) is a quantity which shows how well the model reproduces the calibration data (Sorooshian & Gupta, 1995). For example, the most commonly used are: the sum of squared errors (SSE) (Merriman, 1877), root mean squared error (RMSE) (Gupta *et al.*, 2009) and Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970). However, single-objective approaches are often inadequate to match different aspects of a model response (Savic, 2002). Because of this, multi-objective approaches have gained in popularity by focusing on representing additional process representations such as peak flows, general and low flows (Gupta *et al*., 1998; Madsen, 2000; Moussa and Chahinian, 2009).

The model parameters may be restricted to physically plausible parameter range. The upper and lower bounds of the model space are defined as hypercube functions of the parameter space (Madsen, 2003).

Model calibration criteria seek to optimize the numerical measures (objective functions) that compare observations of the state of the system with corresponding simulated values, using optimization algorithms (Madsen, 2003). An optimization algorithm is a mathematical technique which analytically or numerically finds parameter optima within a model space. The optimization algorithm searches the response surface for the parameter values that optimize (minimize or maximize) the numerical value of the objective function, constrained to the pre-defined allowable ranges of the parameters (Pechlivanidis *et al*., 2011).. For example, Sorooshian and Gupta (1995) classify optimization algorithms as "local" and "global". Local search algorithms may be further divided into "direct" and "gradient-based" methods. Direct search methods use only information on the objective function value, whereas gradient-based methods also use information about the gradient of the objective function (Madsen, 2003). For multi-model objective functions, global search methods have been developed that are especially designed to locate the global optimum and avoid being trapped in local optima. Global population-evolution-based algorithms are more effective than multi-start local search procedures, which in turn perform better than pure local search methods (Duan *et al*., 2006; 1992; Madsen, 2003). The most widely applied is MOCOM-UA (Khu and Madsen, 2005; Yapo *et al*., 1998). This algorithm is an effective and efficient methodology for solving the multiple-objective global optimization problem (Mostafaie *et al*., 2018), and is an extension of the successful SCE-UA single-objective global optimization algorithm Shuffled Complex Evolution (SCE) (Duan *et al*., 1992).

The choice of optimization algorithm depends solely on the hydrological model and its objective function. It is therefore important to emphasize that model structures and observations are not error-free. Therefore, the optimum parameter set is model specific and may not remain optimum if change occurs in the model structure or the calibration data. While one optimum parameter set may be found, there will usually be many other parameter sets in place that are very nearly as good. Hence, Beven and Freer (2001) suggest that equifinality of parameters may give many representations of a catchment that may be equally valid in terms of their ability to produce acceptable simulations of the available data, given the limitations of both the model structure and data.

2.6.4. Concept of equifinality in modelling

Optimization of model parameters may not be considered a good strategy, mainly because the obtained optimal model is highly dependent on input data and the structural errors, so acceptance of the possibility of multiple behavioural models (the equifinality concept) is highly recommended (Beven, 2006;2010). Equifinality in environmental modelling is the concept that there are many different model structures, and many different parameter sets within a chosen model structure that may be behavioural, or capable of reproducing the observed behaviour of that system to an acceptable standard (Beven 2010). For example, if there are enough interactions among the components of a system (e.g., a model), unless the detailed characteristics of these components are specified independently, many representations may be equally acceptable. This has been demonstrated in studies incorporating rainfall-runoff models (Beven and Freer, 2001; Beven, 2011; Duan *et al*., 1992) and in flood frequency and inundation models (Cameron, Beven and Naden, 2000). This should not be enough to reject the idea of an optimal model parameter, but a search for the feasible model structure and parameter space is usually a good way to reveal many behavioural models with similar levels of performance in reproducing observational data. Also, where an incorrect model of errors is used, or where uncertainty in the observations with which the model is being compared is neglected, it will generally lead to over-conditioning of the parameter values.

In such cases other forms of model conditioning approaches which allow the possibility of learning from model rejection, such as Generalised Likelihood Uncertainty Estimation (GLUE) , (Beven and Binley, 1992), are used. These are more subjective but allow for easy handling and updating of the model likelihood distributions as new calibration data become available within the Bayesian framework (Beven and Freer, 2001). In the GLUE methodology, some prior information about feasible ranges of parameter values is used to control the generation of independent random parameter sets for use in each model. An input sequence is used to drive each model and the results are compared with the available calibration data. The model simulations may have either a deterministic or a stochastic dependence on the parameters and input data, but the methodology has to date been primarily used with deterministic models. A quantitative measure of performance or likelihood measure is used to assess the acceptability of each model based on the modelling residuals (Beven and Freer, 2001). Other performance criteria have also been suggested and proposed, such as fuzzy possibility measures (Franks *et al.*, 1998) and likelihood functions based on specific error models (Romanowicz and Beven, 1998). In addition, the non-stationarity of a

system plays a key role in reproducing a range of plausible model parameters, since the characteristics are expected to change with time (Beven, 2010, pg 131).

2.6.5. Concept of non-stationarity in hydrological models

Non-stationarity is a system whose characteristics are expected to change over time. For example, in cases of model systems, the parameters are expected to change over time (Beven 2010). It is thus important to distinguish non-stationarity in models because it has an impact on what we should expect in testing a model as a hypothesis of how a catchment functions, and whether it would be fit for purpose (Beven, 2016). Non-stationarity in modelling arises in stochastic processes, catchment characteristics, boundary conditions, and model residual characteristics (Beven, 2016). Koutsoyiannis and Montanari (2015) discuss the concept of non-stationarity in the context of stochastic process theory by assuming that once any deterministic structure has been considered, all forms of epistemic error can be represented by a stationary stochastic model. Change over time can be described by a deterministic function, including structure in model residuals that might compensate for consistent model or boundary condition error. All other variability will be stochastic in nature.

Non-stationarity in catchment characteristics arises where model parameters and possibly catchment characteristics are expected to change over time or space in a way that will induce model prediction error if parameters are considered stationary (Beven, 2016). Non-stationarity in boundary conditions occurs where the model boundary conditions are expected to change over time or space in a way that will induce model prediction error if boundary conditions are poorly estimated. Non-stationarity in model residual characteristics is the expectation that the statistical characteristics of the model residuals will vary significantly in time and space. This may be due to the unpredictable model error which may result from arbitrary epistemic uncertainties in boundary conditions, long-term stochastic variability, or error from the calibration data (Beven, 2016).

2.6.6. Model validation

Calibration is followed by validation to check the model's predictive power and appropriateness for real world applications. Klemeš (1986) discusses various validation tests and proposes a hierarchical testing framework which seeks to assess the transferability of a model (i.e., its ability to make predictions outside the calibration period). This framework can be applied under stationary and non-stationary climate assumptions. The proposed validation tests are: (i) the split-sample test, where the model is calibrated using a historical data period and then the same calibrated parameters are applied for simulating an independent period; (ii) the differential split-sample test, where the same approach is followed but the data period is divided into two segments with varying climate characteristics (e.g., high/low average precipitation); (iii) the proxy-basin test, where the model is calibrated and validated in various sites, which may be in the same catchment or not but with similar edapho-climatic characteristics; and (iv) the proxy-basin differential split-sample test which is applied in cases where the model is supposed to be both geographically and climatically (or land-use-wise) transposable (Klemeš, 1986; Santos *et al.*, 2018). In modelling, if the errors for validation runs on two data record segments (i.e., calibration

and validation data sets) are within acceptable ranges, the model is acceptable (Beven, 2012; Gupta et al., 2006). Basically, the modelling approach adopted depends on the required spatio-temporal scale of the problem), the type of catchment, and the modelling task (e.g., establishing the suitable data input for a hydrological model). Many model applications pose major challenges. For example, in data-scarce regions, there is insufficient time series data for model calibration and this may require the use of freely available satellite and reanalysis products which have inherent errors and uncertainty due to coarse resolutions. In addition, not all models are applicable to diverse catchments with varied characteristics, which makes the model parameter estimation difficult due to water balance problems and/or influence of water reservoirs on the flow, affecting the model simulations of catchment stream flows. It is therefore useful to consider a basic classification of model types, according to Wheater *et al*., (2007), and their strengths and weaknesses (see Table **2,** Table 3), and select the model according to a specific application.

2.7.   Trends in flood series and shifts in flood seasonality

The frequency of floods is rising and constitutes one of the main causes of detrimental consequences arising from natural disasters, not only in Kenya, but across the globe. In the 21st century, the first two decades have been characterised by many major flood events (see Hannaford 2015), which have prompted a wider discussion about flood risk management and the possibilities that the increase in flooding is due to human induced climate change (Hannaford et al 2021). Accounting for possible changes in flooding is important for various hydrological applications such as risk assessment and management (Mangini *et al*., 2019), and the design of flood protection facilities (Parry *et al.*,2007). Also, flood frequency estimation may best be understood by considering temporal trends in flood series (Harrigan *et al*., 2018).

2.7.1. Trends in river flow series

Trend analysis can be used to investigate whether there is any evidence of an increase in river floods in the observational river discharge data. Such analysis requires long records (i.e.., more than 30 years) to distinguish climate variability decisively from climate change induced trends (Svensson *et al*., 2005) and human induced trends due to deforestation and water management practices (Degefu *et al*., 2019).

To detect trends in river flow series, trend analysis studies have been performed in different parts of the world, at global scale (e.g., Berghuijs *et al*., 2017), regional scale (e.g., Paprotny *et al*., 2018; Mangini *et al*., 2019; Burn *et al*., 2016) and at country scale (e.g., Nka *et al*., 2015; Avial *et al*., 2019). Most of the studies found large-scale patterns with a similar sign of changes in flood magnitude, but also concluded that the studies are not fully comparable, due to the different time periods analysed and differences in the methodology applied to derive flood series and to detect flood changes. This points to the need for site-specific trend analysis to reveal changes and shifts in flood events at local scale. This is the aim of the research presented in Chapter 6 of this thesis.

Flood trend analysis looks at trends in the annual maximum river discharge (AMAX), i.e., a one value per year flood series (Kundzewicz *et al*. 2004; 2005) and the Peak Over Threshold (POT) technique (Burn et al 2017; Mangini et al 2019), which selects all floods over a specific threshold that occur throughout a flow record. POT allows for a trend in the frequency (counts) of floods

rather than merely their magnitude to be estimated (Svenson et al 2006). Detailed discussion on the AMAX and POT techniques used to derive times series of flood series are discussed in Section 6.5.2 of Chapter 6 of this thesis.

2.7.2. Flood timing and seasonality

Flood timing, magnitude and frequency are useful sources of information about trends in river flooding (Berghuijs *et al.*, 2017, 2019). For instance, a recent study by al., (2015) Stephens *et al.*, (2015) has shown that the timing of floods in East Africa has shifted in recent decades: the drier months have seen more rains than in the past, and the onset and cessation of the rain is much earlier or later as demonstrated in Gudoshava *et al.*, 2021). These changes in the timing of the yearly flood have far-reaching consequences for flood-based farming systems, especially for the livelihoods of people who adjust their floodplain management and agricultural activities to the rise and fall of the flood wave (Ficchì and Stephens, 2019). Thus, it should be recognized as an important contribution research for future planning and adaptation purposes, especially for Kenya.

Seasonality measures (Brun et al, 2010) are used to characterize the timing and variability of the extreme flood events. These are defined by directional statistics (Madia, 1972). The date of occurrence of a flood event is defined as a directional statistic through conversion of the Julian date of the occurrence of an event to an angular value (Ficchi and Stephens,2019). A detailed description of seasonality statistics and changes in flood timing is provided in Section 6.5.5 of Chapter 6 of this thesis.

Chapter 3

3. Study Area, Data and Methods

This Chapter describes the study catchments, data used in this research and their sources, the methods employed, and the hydrological modelling experiment. Each dataset described in the following sub-sections was applied in deploying the methods outlined to achieve the objectives of this research, as illustrated in Figure 3.2.

3.1. Study area

The study is undertaken across Kenyan catchments at 19 river gauging stations (Figure 6.1) with varying characteristics summarised in (Table 14) in Chapter 6 of the thesis. These were selected due to the frequency and magnitude of the impacts of floods (Table 1), as well as the availability of river flow observations as summarised in Table 14 of Chapter 6.
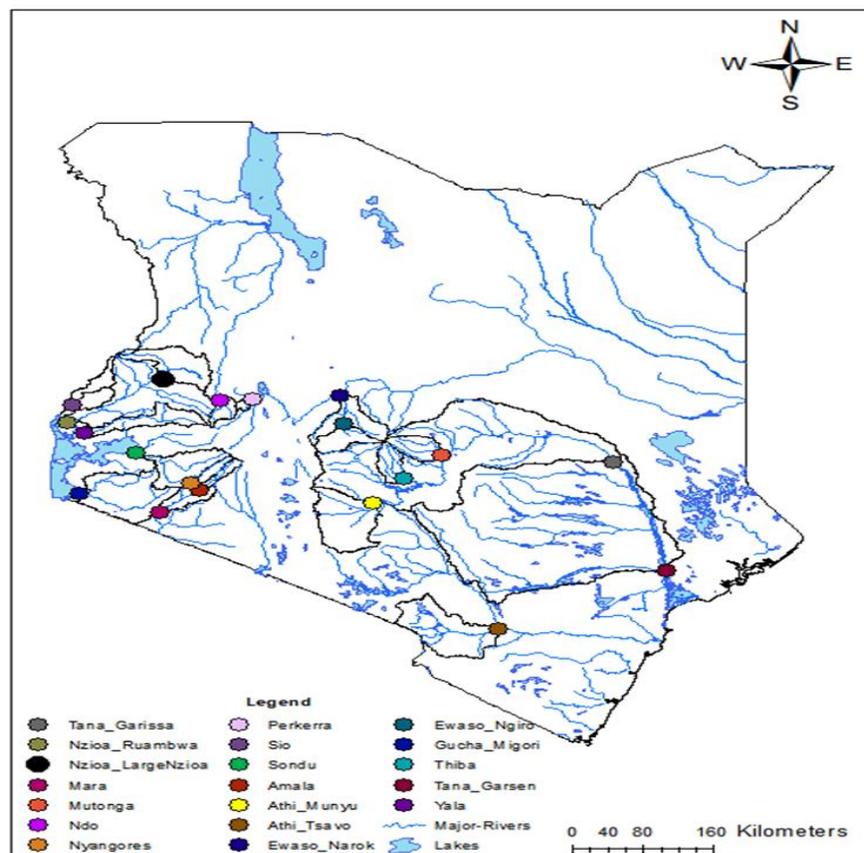
Figure 3.1: Study catchments, with the location of the outlet river gauges (as shown by circled dots) used in this study and the main irrigation schemes and major dams (Source: WRA -K) across Kenya.

3.2. Data

The data used in the study included reanalysis precipitation data, gridded rainfall observations, river discharge data and potential evaporation. The detailed descriptions of the individual datasets are in subsequent sections. All the datasets used in the study spanned the period from 1981 to 2016.

3.2.1. Reanalysis precipitation datasets

Reanalysis products merge available observations with a state-of-the-art atmospheric and/or coupled ocean–atmosphere–sea ice model to derive the best estimate of the state of the atmosphere and land surface (Decker *et al.*, 2012). Some of the applications of reanalysis products are to study the climate system, drive land surface models, and provide boundary conditions for regional modelling. Reanalysis precipitation products can be used to determine variations in precipitation, both at global and regional scales (Kim *et al.*, 2019).

Several weather centres, such as the National Centers for Environment Prediction/National Center for Atmospheric Research (NCEP/NCAR; Kistler *et al.*, 2001), the European Centre for Medium Range Weather Forecasts (ECMWF; Huang *et al.*, 2016), and the Japanese Meteorological Agency (JMA; (Kobayeshi *et al.*, 2015) have released reanalysis products over the past decades. The continuous development of three- and four-dimensional variational (4D-Var) techniques has enabled the updating of reanalysis products; consequently, the new fifth-generation atmospheric reanalysis of ECMWF (i.e., ERA5) was released in 2017. The overwhelming majority of observation data, and most of its increase over time originates from satellites. Such data include clear-sky radiance measurements from polar-orbiting and geostationary sounders and imagers, atmospheric motion vectors derived from geostationary satellites, scatterometer wind data, and ozone retrievals from various satellite-borne sensors. The total precipitable vapour estimates are also derived from satellite observations. Although manual and automatic ground observations of precipitation were also considered, the number of stations in the QDM involved in data assimilation is unclear.

While they are the best approximation of the state of the atmosphere based on both data and dynamic models, the various reanalysis products from the different centres have been found to have many deficiencies at various time scales, and especially at smaller (i.e., regional) spatial scales, mainly due to the errors of models and observations in the data assimilation system (Jiang *et al.,* 2020). The errors of the reanalysis products may lead to biased precipitation estimates for different regions and precipitation events (e.g., Beck *et al.*, 2017), which may introduce challenges in initial parameter calibration for the model simulation (see Wanzala *et al.*, 2022).

In this thesis (see Table 11), ERA 5, ERA-Interim, CFSR and JRA-55 reanalyses, which have been used in several studies due to their availability to the public (Beck *et al.,* 2017; Dinku *et al.,* 2018; Le and Pricope, 2017), were used in Chapter 5. These four are used as inputs into the hydrological model and the results compared with observations to inform their future applications over Kenyan catchments.

ERA5, the latest (fifth) generation global atmospheric reanalysis product by the ECMWF, spans the modern observing period from 1979 onward and has become available in 2018 with daily updates continuing forward in time (Hersbach, 2018). The preliminary ERA5 dataset from 1950 to 1978, which was not available at the time of undertaking this research, is now available on the Climate Data Store. Key features, not present in the previous version, ERA-Interim (Dee *et al.,* 2011), include 10 years' worth of advances in NWP, higher spatial resolution of 31 km on 139 levels, from the surface up to 0.01 hPa (around 80 km), improvements in the ingested observations,

and a near-real time updating service (2 to 5 days latency). ERA5 data products also include information about uncertainties, provided at 3-hourly intervals and at a horizontal resolution of 62 km. In addition, a dedicated ERA5 land component delivers a land-surface reanalysis product at an enhanced resolution combined with several new parameters, such as 100-metre wind speed and direction (Hersbach, 2018). In this study, 3-hourly ERA-5 was obtained from ECMWF on a fixed grid of $0.31° × 0.31°$ (https://www.ecmwf.int/en/research/climate-reanalysis/era-5).

ERA-Interim is a global reanalysis product created by ECMWF (Dee *et al.,* 2011) which was initiated in 1979 reanalysis and covers the period from 1989 to present. Unlike ERA5, ERA - Interim uses a forecasting model of [version cycle 31r1 (CY31r1)] with a horizontal resolution of T213 (~80 km) and incorporates full four-dimensional variational data assimilation (4DVar) as in ERA5. In this study, monthly ERA-Interim was obtained from ECMWF on a fixed grid of $0.75° × 0.75°$ (http://apps.ecmwf.int/) .

JRA-55 is a global reanalysis dataset constructed by the Japan Meteorological Agency (JMA) (Kobayashi *et al.,* 2015) and was adopted for this study. It is freely accessible online at (http://jra.kishou.go.jp/). JRA-55 employs a 4D-VAR with variational bias correction for satellite radiances. It aims to provide a comprehensive atmospheric dataset that is suitable for studies on climate change and related issues (Ebita *et al.*, 2011). The observation data primarily included conventional data (such as tropical cyclone wind retrievals, pilot balloons, wind profilers, etc.), wind data retrieved from geostationary TOVS (TIROS Operational Vertical Sounder), ATOVS (advanced TOVS), AMV (Atmospheric Motion Vector), CSR (Clear-Sky Radiance) data, and other remote sensing data.
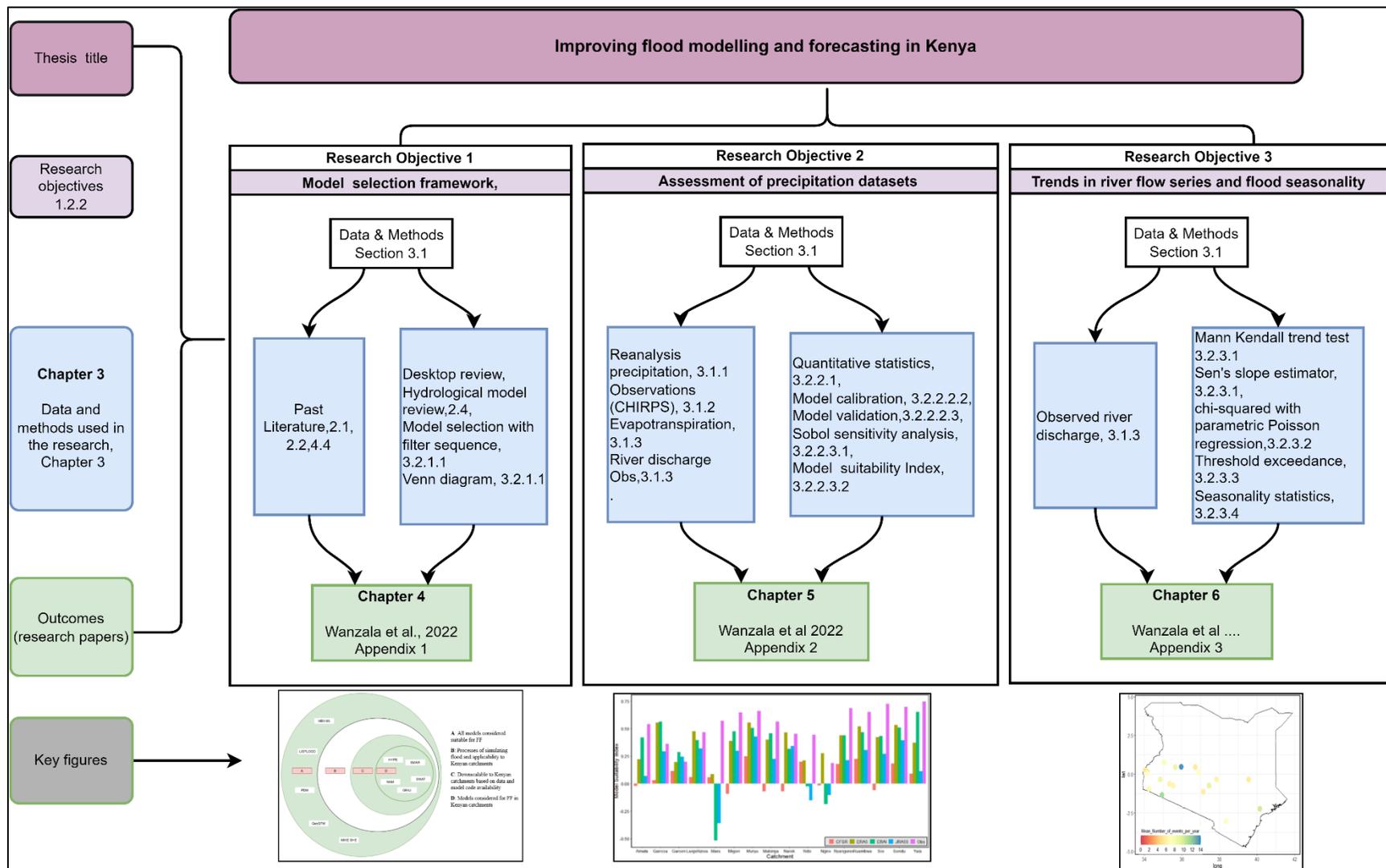
Figure 3.2:- Illustration of the data and methods implemented in the research to achieve the three research objectives.

CFSR is a global reanalysis dataset of atmosphere fields produced by the National Centers for Environmental Prediction and National Center for Atmospheric Research to meet the needs of research and climate monitoring communities (Saha *et al*., 2010). A 3D-VAR (three-dimensional variational analysis) is used in the assimilation system of CFSR. The horizontal resolution is a T62 Gaussian grid with 192×94 grids of the overall dataset (Zhao, He and Jiang, 2018). This model includes parametrizations of all major physical processes, such as convection, clouds, and an interactive surface hydrological model on a global scale. The CFSR precipitation field is constructed from short-range model forecast accumulations, but observed precipitation is not used in the assimilation phase of the model. Monthly data from CFSR were used in this work, which can be downloaded online (https://www.esrl.noaa.gov/). Key characteristics (e.g., spatial grid, temporal accumulations, etc.) of each of the reanalysis datasets used in this thesis are summarised in Table 11.

3.2.2. Gridded precipitation observations

The Climate Hazards Group Infrared Precipitation with Station (CHIRPS) is a relatively new quasi-global, high resolution, daily, pentadal and monthly precipitation dataset (Funk *et al.,* 2015). The dataset provides low latency, long recorded high-resolution gridded data and allows scientists to both analyze current trends and compare them to historic trends (Le and Pricope, 2017). The CHIRPS algorithm combines three main data sources (Dinku *et al.,* 2018): (a) the Climate Hazards group Precipitation climatology (CHPclim), a global precipitation climatology at 0.05° latitude/longitude resolution estimated for each month based on station data, averaged satellite observations, elevation, latitude and longitude (Funk *et al.,* 2015); (b) TIR-based satellite precipitation estimates (IRP); and (c) *in situ* rain-gauge measurements using a modified inverse distance weighted algorithm (Funk *et al.,* 2015). The CHPclim is distinct from other precipitation climatologies in that it uses long-term average satellite rainfall fields as a guide to deriving climatological surfaces. The incorporation of station data also helps to correct for underestimation of the intensity of precipitation events**.**

Based on infrared Cold Cloud Duration (CCD) data, CHIRPS has a long enough (useful for depicting climate variability) history of precipitation data. The algorithm is based on (i) a 5 km climatology that uses satellite data to represent sparsely gauged locations; ii) daily, pentadal, and monthly 5 km CCD-based precipitation estimates from 1981 to the present; iii) a combination of station data to generate a tentative information product with a latency of about 2 days and a final product with an average latency of about 3 weeks; and iv) interpolation weights assigned according to a novel blending method which uses the spatial correlation structure of CCD estimates. This makes it a better alternative than observations (see Funk *et al.,* 2015). CHIRPS was developed to deliver reliable, up-to-date, and more complete datasets. Both monthly and daily time series of the datasets are available. A daily time series spanning the period from 1981- 2016 was used in this study.

We acknowledge the availability of other satellite blended precipitation datasets with a high spatial resolution (≤10 km) such as the Multi-Source Weighted-Ensemble Precipitation (MSWEP;

Beck *et al.,* 2019), CPC morphing technique (CMORPH; 7 km Joyce *et al.,* 2004), Global Satellite Mapping of Precipitation (GSMaP; 10 km; Mega *at al.,* 2014), Integrated Multisatellite Retrievals for Global Precipitation Measurement (IMERG; 10 km ; Huffman *et al.,* 2015), and Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS; 4 km; Hong *et al.,* 2004). However, all these datasets, except MSWEP, have a data record of ≤20 years and neither take advantage of river discharge observations for bias correction, nor incorporate reanalysis-based precipitation estimates.

A time series was plotted, and the correlation coefficients calculated for all the satellite datasets and observations to demonstrate that CHIRPS provides a closer estimate to observations compared to other available datasets (Figure 3.3). Additionally, precipitation evaluation studies particular to Eastern Africa (including Kenya), further detailed in Section 2.5 have paid less attention to the potential of MSWEP as an alternative to observations and shown better performance for CHIRPS relative to other observations (see Dink *et al.,* 2018). Even though the daily temporal resolution of CHIRPS renders it less suitable in highly dynamic precipitation analysis and includes spurious drizzle and underestimation of peak magnitudes of the most extreme rainfall (Beck *et al.,* 2017), since the focus for this thesis was not necessarily on high temporal timescales such as hourly rainfall events, it should therefore be insensitive to these biases. Consequently, the CHIRPS dataset was used as a control dataset since it had been used in several studies (see Table 11, and was applied in Chapter 5 of this thesis, as illustrated in Figure 3.2.
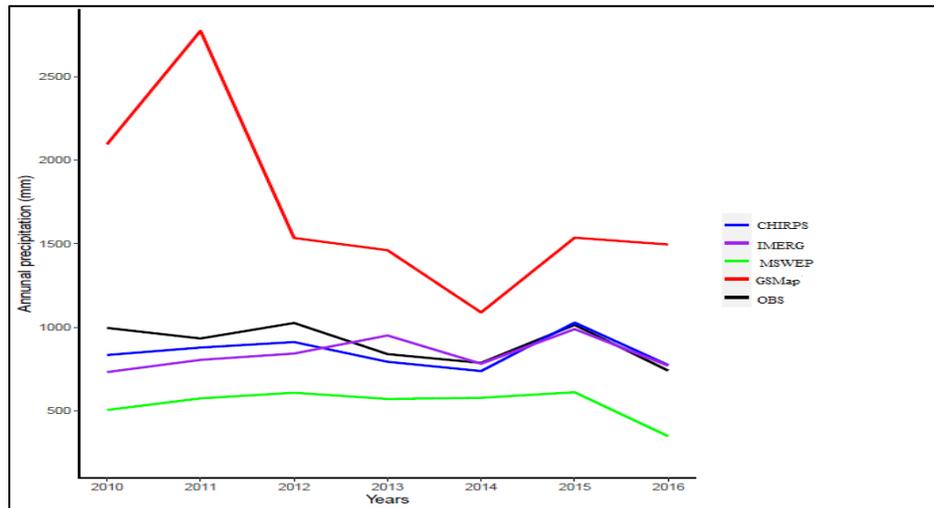
Figure 3.3:- Time series for the annual precipitation of different satellite observation datasets relative to observations over Kenya for the period 2010 – 2016. Black (gauge observations), blue (CHIRPS), red (GSMap), green (MSWEP) and purple (IMERG).

3.2.3. Observed river discharge and potential evapotranspiration

As illustrated in Figure 3.2, river discharge datasets at daily time steps for the period 1981- 2016 were used in Chapters 5 and 6 in the evaluation of the GR4J hydrological model and for extracting high river flow series for trend analysis respectively. Data for the select catchments across the country were purchased from the Kenya Water Resource Authority (WRA). This is because, like

other observational data, river discharge was also not freely available for research and could be obtained only through negotiations leading to purchase from the relevant authority (WRA).

The potential evapotranspiration (PET) for the datasets outlined in Section 3.2.1 and 3.2.2 was estimated from the average temperature. Different methods of estimating PET exist, based on factors ranging from energy balance and mass transfer to temperature. For this study, temperature-based methods were considered, as temperature readings are often the most readily available meteorological data relating to PET (Hargreaves and Samani, 1985). Also, under limited climate data conditions, many researchers propose the use of temperature-based methods (Hargreaves and Samani, 1985; Gardelin and Lindström, 1997) and a study by Oudin *et al.,* (2005), tested 27 PET for catchment models and concluded very simple models relying only on extra-terrestrial radiation and mean daily temperature are as efficient as more complex models such as the Penman model and its variants.

The Hargreaves-Samani (H-S) and Hamon methods are well known. Although the H-S method uses only a daily measurement of maximum and minimum temperatures as inputs, it effectively incorporates radiant energy indirectly. Relative humidity and cloudiness are not explicitly included in the equation, but the difference between maximum and minimum air temperature is related to relative humidity and cloudiness (Samani and Pessarakli, 1986). Hamon's method is a simple, empirical approach that uses temperature as the major driving force for evapotranspiration, but also includes other variables, such as daytime length and saturated vapour pressure (Allen *et al.*, 1998; Lu *et al.*, 2005). In this method, the daytime length is used as an index of the maximum possible incoming radiant energy and the saturated vapour pressure is the moisture holding capacity of the air at the prevailing air temperature. Hamon's method has proved to produce good results and has displayed great resilience in diverse climates around the world (Shahidian *et al.*, 2012), so was considered for this study. Hamon's equation is given by Equation 3.1:

Equation 3.1:- Hamon's method of computing potential evapotranspiration.

$$E_T = k * 0.165 * 216.7 * N * \left(\frac{e_s}{T+273.3}\right)$$

where E_T is potential evapotranspiration (PET) (mm/day^{-1}), k is the proportionality coefficient = 1^1 (unit less), N is daytime length (X/12 hours), and e_s is saturation vapor pressure (mb), which represents the moisture holding capacity of the air at the prevailing air temperature.3.3. Methods

The various methods employed to achieve the three specific objectives are illustrated in Figure 3.2. The methods applied included: model preselection with a filter sequence, a Venn diagram tool for model selection (3.3.1.1), statistical performance metrics to evaluate the reanalyses with respect to observations (correlation coefficient (CC), root mean square error (RMSE), mean absolute error (MAE) and percentage bias (BIAS) (3.3.2.1), the Michel method for model calibration (3.3.2.2.2), Split sample testing for model validation for hydrological evaluation of the reanalysis data using the GR4J model (3.3.2.2.3), Sobol' Sensitivity Analysis (SA) of model parameters to different inputs (3.3.2.3.1) and combined performance and uncertainty statistics using the Model Suitability Index (MSI, 3.3.2.3.2)). For model validation and sensitivity analysis,

model performance evaluation was undertaken using the KGE. The trends in different derived flood series from the observed and simulated river discharge were evaluated using a statistical method (Mann-Kendall and Thielsen Slope Estimator, 3.3.3.1). The derived flood series were then subjected to sensitivity analysis to assess the sensitivity of trends to different flood magnitudes and frequencies (3.3.3.2). For threshold selection, Mean Residual Life plots were used. Finally, the changes in seasonality and shift in flood timing across Kenya was assessed using seasonality statistics (3.3.3.3). The sub-sections below outline each of the methods used to achieve the three objectives of this research.

3.3.1. Model pre-selection for a Kenya national flood forecasting centre

This section outlines the methods used to achieve the first objective of the research: to design and propose an objective model pre-selection criterion with a filter sequence for a Kenyan national flood centre. The model pre-selection process following a filter sequence is outlined and a Venn diagram tool used to select the model candidates for flood applications in Kenya from several models reviewed in Section 2.4.

3.3.1.1. Filter sequence and Venn diagram

A filter sequence can be used to arrive at a special/desired subset of a partially ordered set. In most cases a weighting factor is assigned at each stage of a particular subset to assess multiple alternatives arising from a mixture of quantitative and mostly qualitative information from multiple sources. For example, in this thesis, a decision is presented by yes/no decision outcomes, which has potential implicit weighting factors of '0' or '1' according to whether the model meets a certain criterion or not. The weighting factor is applied to a filter sequence and 12 hydrological models evaluated. A Venn diagram is then used to select the model candidates suitable for flood applications. A Venn diagram is a diagrammatic illustration of a union of sets, where common elements of the sets are represented by intersections of circles or a common circle. Detailed examples and illustrations of the application of a filter sequence and Venn diagram are found in Section 4.6 of Chapter 4 of this thesis.

The first step in any model selection process is to assess the potential criteria to be considered (e.g., aim, resolution and scope of the model system) (Bennett. *et al.*, 2013; Jakeman *et al.,* 2006; Kauffeldt *et al.,* 2016). The aim of the model defines the sole purpose which the model is developed to serve. This may include different applications, such as forecasting, catchment modelling and water balance. The scope of the model development system is important in creating a balance between the users' wishes and the scientific progress involved in the models' technical advances (Kauffeldt *et al.,* 2016). However, the selection criteria are, in the end, subjective, since they are a consequence of the application in question, and the scope of model operations. For example, selection criteria may depend on identification of the user community, and demands on the model's structure, complexity and flexibility (Bennett *et al.,* 2013; Kauffeldt *et al*., 2016). We applied a filter sequence and a Venn diagram tool for objective model pre-selection for operational national flood forecasting. Detailed criteria for a filter sequence are outlined in Chapter 4.

61

3.3.2. Assessment of reanalysis precipitations for hydrological modelling

Objective II explores the performance of different reanalysis precipitation datasets for hydrological modelling across Kenyan catchments. This section outlines the statistical methods used to assess the performance relative to observations and evaluation of the simulated discharge, by a hydrological model when forced with different reanalysis datasets. The sensitivity of model parameters to different inputs (i.e., precipitation from different reanalysis products) is also explored, and the model's performance is assessed through a combination of performance statistics and sensitivity indices at catchment scale.

3.3.2.1. Performance statistics

The performance statistics of the reanalysis datasets were quantitatively evaluated in terms of temporal dynamics and biases with respect to observations (CHIRPS), considering the following metrics: the Pearson Linear Correlation Coefficient (CC), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Error (ME), and long-term relative percentage bias (BIAS).

The Pearson Correlation Coefficient (CC) is unitless and used to evaluate the goodness of fit of the relationship between the observation and the reanalysis. The value varies from -1 to 1, where positive and negative values indicate positive and negative correlation respectively and a value of 1 is the perfect score. The equation is expressed by Equation **3.2**.

Equation 3.2:- Pearson Correlation Coefficient.

$$CC = \frac{\sum_{i=1}^{n}\left(O_i - \bar{\bar{O}}\right)\left(P_i - \bar{\bar{P}}\right)}{\sqrt{\sum_{i=1}^{n}\left(O_i - \bar{\bar{O}}\right)^2 \left(P_i - \bar{\bar{P}}\right)^2}}$$

The root mean square error (RMSE) measures the absolute mean difference between two datasets (observation and reanalysis). A value of 0 is the perfect score. The equation is expressed by Equation **3.3**.

Equation 3.3:- Root Mean Square Error.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}\left(O_i - P_i\right)^2}$$

Mean Absolute Error (MAE) demonstrates the magnitude of mean error and is given by Equation **3.4**.

Equation 3.4:- Mean Absolute Error.

$$MAE = \frac{1}{n}\sum_{i-1}^{n}\left|O_i - P_i\right|$$

Bias is a measure of how the average precipitation magnitude compares to the observed precipitation. It provides information on the magnitude of underestimation or overestimation between two datasets, in which the closer to 0 the BIAS, the better the performance of the precipitation dataset. Bias can be expressed as a percentage and takes the form of Equation **3.5**.

Equation 3.5:- Computation of precipitation BIAS.

$$BIAS = \frac{\sum_{i=1}^{n}(P_i - O_i)}{\sum_{i=1}^{n} P_i} \times 100\%$$

where n is the total number of monthly samples; O_i and P_i represent the observed and reanalysis precipitation for the ith month respectively; $\overline{\overline{O}}_i$ and $\overline{\overline{P}}_i$ are the average values over n months of O_i and P_i respectively.

3.3.2.2. Hydrological evaluation

The hydrological modelling experiment, further detailed in Chapter 5 was set up in the study to address objective (II). The performance of the hydrological model for Kenyan catchments was evaluated against observed discharge for different reanalysis products to determine the sensitivity of results to these different reanalysis products. The experiment involves hydrological model calibration, and validation using different reanalyses and observations as inputs. From the models reviewed in Section 2.4, the GR4J model was chosen for the modelling experiment because it has fewer parameters and is easy to calibrate and apply to diversified environments. The following subsections outline the detailed description of the GR4J model, calibration and validation methods applied.

3.3.2.2.1. GR4J a daily four-parameter hydrological model

There are four free main parameters to be calibrated in the GR4J model (Figure 3.4), namely: maximum capacity of production store (*X1*, mm), groundwater exchange coefficient (*X2*, mm), maximum capacity of non-linear routing store (*X3*, mm), and time base of the unit hydrograph (*X4*, days). There are also a number of fixed parameters, whose values were set by Perrin, Michel and Andréassian (2003). All four free parameters are real numbers. *X1* and *X3* are positive, *X4* is greater than 0.5 and *X2* can be either positive, zero or negative. The description of the free and fixed parameters are summarized in

Table **5.**

In GR4J model, all the water quantities including input, output and internal variables are expressed in mm and the operations described in the following equations are relative to a given time-step corresponding to a discrete model simulation. GR4J model physical meaning of the internal variables are summarized in Table 4 of Chapter 2 in the thesis..

Table 5:- Free and fixed parameters of the GR4J hydrological model (from Perrin *et al*.,2003, except for *Ut* and *nres*).

Parameter type	Parameter name	Description	Typical value	Unit
Free	*X1*	Capacity of the production store	[0, 1000]	[mm]
	X2	Groundwater exchange coefficient	[-5, 5]	[mm]
	X3	Capacity of the nonlinear routing store	[0, 300]	[mm]
	X4	Unit hydrograph time base	[0.5, 5]	[days]
Fixed	α	Production precipitation exponent	2	[-]
	β	Percolation exponent	5	[-]
	γ	Routing outflow exponent	5	[-]
	ω	Exchange exponent	3.5	[-]
	ϵ	Unit hydrograph coefficient	1.5	[-]
	ϕ	Partition between routing store and direct flow	0.9	[-]
	ν	Percolation coefficient	4/9	[-]
	U_t	One time step length	1	[days]
	nres	Number of stores in Nash cascade	11	[-]

Determination of the net rainfall and PE: The initial step is to determine the net rainfall P_n or net evapotranspiration capacity E_n, by subtracting E from P, which is computed by assuming an interception storage of zero, thus P_n and E_n are computed following Equation 3.6 and

Equation3.7.

Equation 3.6:- Net precipitation when precipitation is greater or equal to evaporation

$$if\left(P \geq E\right)then\left(P_n = P - E\right)and\left(E_n = 0\right)$$

Equation3.7:-Net evapotranspiration capacity when net precipitation zero

Otherwise:

$$\left(P_n = 0\right)and\left(E_n = E - P\right)$$

Figure 3.4: Schematic representation of the GR4J rainfall-runoff model (Source: Perrin et al., 2003). P is rainfall depth; E is potential evapotranspiration estimate; Q is total streamflow; X_i are the model parameters; all other letters are model variables or fluxes summarized in Table 4.

Production store can be considered as the soil moisture accounting (SMA). If P_n is not equal to zero, then a part P_s of P_n fills the production store. In this case, the production (SMA) store is determined as a function of the level S in the store by Equation **3.8**.

Equation 3.8:- Soil Moisture Accounting store when net precipitation is not equal to zero.

$$P_s = \frac{x_1 \cdot \left(1 - \left(\dfrac{S}{x_1}\right)^2\right) \cdot \tanh\left(\dfrac{P_n}{x_1}\right)}{1 + \dfrac{S}{x_1} \cdot \tanh\left(\dfrac{P_n}{x_1}\right)}$$

where the terms are defined in Table 4 and 5. When $E_n \neq 0$ an actual evaporation rate is determined as a function of the level in the production store to calculate the quantity of E_s, i.e., water likely to evaporate from the store, which is obtained by Equation **3.9**,

Equation 3.9:-Water likely to evaporate from the SMA store when net evaporation is net equal to zero.

$$E_s = \frac{S \cdot \left(2 - \dfrac{S}{x_1}\right) \cdot \tanh\left(\dfrac{E_n}{x_1}\right)}{1 + \left(1 - \dfrac{S}{x_1}\right) \cdot \tanh\left(\dfrac{E_n}{x_1}\right)}$$

Equations 3.8 and 3.9 are as a result of integration over the time-step of the differential equations which follow a parabolic form with terms in $\left(\dfrac{S}{x_1}\right)^2$, as detailed in Edijaton and Michel, (1989).

The water content in the production store is updated with Equation **3.10**,

Equation 3.10:- Update function of the water content in the production store.

$$S = S - E_s + P$$

S can never exceed x_1

GR4J has a **percolation function** computed from the production store, thus making it distinguishable from the previous GR3J version (Edijaton *et al.*, 1999). The percolation leakage, *Perc* , is always smaller than S and is calculated as a power function of the reservoir water content using

Equation 3.11.

Equation 3.11:- Computation of the percolation leakage function.

$$Perc = S\left\{1 - \left[1 + \left(\frac{4}{9}\frac{S}{x_1}\right)^4\right]^{-\frac{1}{4}}\right\};$$

$Perc$, is always smaller than S , thus the reservoir content is given by Equation 3.12:

Equation 3.12:- Resultant reservoir water content after percolation

$$S = S - Perc$$

However, percolation does not contribute much to the streamflow and is interesting mainly for low flow simulation.

A **linear routing component with unit hydrograph** uses unit hydrographs to determine the total quantity P_r of water that reaches the routing function, which is given by Equation 3.13

Equation 3.13: Total quantity of water reaching the routing function.

$$P_r = Perc + \left(P_n - P_s \right)$$

P_r is divided into two components following a fixed split where 90% of P_r is routed by a unit hydrograph $UH1$ and then a non-linear routing store, and the remaining 10% of P_r is routed by a single unit hydrograph $UH2$ UH1 and UH2 are useful in simulating the time lag between the rainfall event and the resulting streamflow peak. $UH1$ and $UH2$ ordinates are used in the model to spread effective rainfall over several successive time-steps. Both unit hydrographs depend on the same time parameter x_4 expressed in days. However, $UH1$ has a time base of x_4 days whereas $UH2$ has a time base of $2 \bullet x_4$ days. x_4 can take real values and is greater than 0.5 days.

Unit hydrographs $UH1$ and $UH2$, in their given discrete form have n and m ordinates respectively, n and m represent smallest integers exceeding x_4 and $2 \bullet x_4$ respectively, which physically means that water is staggered into n unit hydrograph inputs for $UH1$ and m inputs for $UH2$ (see Equation 3.24). The corresponding S-curves (cumulative proportion of the input with time, denoted by $SH1$ and $SH2$) are used to derive the ordinates of each of the unit hydrographs.

$SH1$ is defined along time t and is derived by Equations 3.14; 3.15 and 3.16:

Equation 3.14:- Cumulative proportion of the input in $UH1$ when $t \leq 0$.
$$For\, t \leq 0, SH1(t) = 0$$

Equation 3.15: Cumulative proportion of the input in $UH1$ when $0 < t < x_4$.

$$For\, 0 < t < x_4, SH1(t) = \left(\frac{t}{x_4} \right)^{\frac{5}{2}}$$

Equation 3.16:- Cumulative proportion of the input in $UH1$ when $t \geq x_4$.
$$For\, t \geq x_4, SH1(t) = 1$$

$SH2$ is similarly defined along time t and is derived by Equations 3.17; 3.18; 3.19 and 3.20:

Equation 3.17: Cumulative proportion of the input in $UH2$ when $t \leq 0$.

$$For\, t \leq 0, SH2(t) = 0$$

Equation 3.18: Cumulative proportion of the input in $UH2$ when $0 < t < x_4$.

$$For\, 0 < t < x_4, SH2(t) = \left(\frac{t}{x_4} \right)^{\frac{5}{2}}$$

Equation 3.19: Cumulative proportion of the input in $UH2$ when $x_4 < t < 2 \cdot x_4$

$$For x_4 < t < 2 \cdot x_4, SH2(t) = 1 - \frac{1}{2} \left(2 - \frac{t}{x_4} \right)^{\frac{5}{2}}$$

Equation 3.20: Cumulative proportion of the input in $UH2$ when $t \geq 2 \bullet x_4$

$$For t \geq 2 \bullet x_4, SH2(t) = 1$$

The ordinates of the unit hydrographs $UH1$ and $UH2$ are then calculated by Equations 3.21 and 3.22 respectively.

Equation 3.21: Ordinate of the unit hydrograph $UH1$.

$$UH1(j) = SH1(j) - SH1(j-1)$$

Equation 3.22: Ordinate of the unit hydrograph $UH2$.

$$UH2(j) = SH2(j) - SH2(j-1)$$

Where j is an integer and for any given $0.5 \leq x_4 \leq 1$, $UH1$ has a single ordinate which is equal to one and $UH2$ the has two ordinates (see Figure 3.4).

The outputs $Q9$ and $Q1$ of the two unit hydrographs correspond to the discrete convolution products at each time-step, and are computed by Equations 3.23 and 3.24 respectively.

Equation 3.23: Output of the unit hydrograph $UH1$.

$$Q9(k) = 0.9 \sum_{j=1}^{m} UH1(j) \bullet P_r(k - j + 1)$$

Equation 3.24: Output of the unit hydrograph $UH2$.

$$Q1(k) = 0.1 \sum_{j=1}^{m} UH2(j) \bullet P_r(k - j + 1)$$

Where $n = \text{int}(x_4) + 1$ and $m = \text{int}(2 \bullet x_4) + 1$

The **groundwater exchange** term F acts on both flow components and the higher the level in the routing store, the larger the exchange, which is given by Equation **3.25**.

Equation 3.25:- Groundwater exchange term which acts on both flow components.

$$F = x_2 \left(\frac{R}{x_3} \right)^{\frac{7}{2}}$$

where R is the level and not the water quantity in the routing store, x_3 is 'reference' capacity and x_2 the water exchange coefficient. x_2 can be either positive in the case of water imports, negative for water exports or zero when there is no water exchange. The higher the level in the routing store, the larger the exchange. In absolute value, F cannot be greater than x_2 : x_2 represents the maximum quantity of water that can be added (or released) to (from) each model flow component when the routing store level equals x_3.

The **non-linear routing store** is made up of the cumulative proportion of the input with time (t ≤ 0) of the $UH1$ and the ground water exchange term F, therefore, first the level of the water content in the routing store R is updated by adding the output $Q9$ of $UH1$ and F by Equation 3.26:

Equation 3.26; An update function of the level of water content in the routing store.

$$R = \max\left(0; R + Q9 + F\right)$$

The outflow Q_r of the reservoir is then calculated by Equation 3.27

Equation 3.27: Reservoir outflow of the non-linear routing store.

$$Q_r = R \cdot \left\{ 1 - \left[1 + \left(\frac{R}{x_3} \right) 4 \right] - \frac{1}{4} \right\}$$

Q_r is always lower than R, so the formulation of the output of the store is the same as the percolation from the SMA store and level in the reservoir becomes:

Equation 3.28: Direct computation of the reservoir level.

$$R = R - Q_r$$

Total streamflow follows the same approach as the content of the routing store; the output $Q1$ of $UH2$ is subject to the same water exchange F to give the flow component Q_d as in Equation 3.29:

Equation 3.29: Outflow component of the ground water exchange component.

$$Q_d = \max\left(0; Q1 + F\right)$$

Thus, total stream flow is Q is obtained by Equation 3.30:

Equation 3.30: Total stream flow as a function of the outflow from ground water exchange and routing store

$$Q = Q_r + Q_d$$

3.3.2.2.2. Model calibration

The four free parameters of the GR4J model were calibrated using the default optimisation algorithm provided in the airGR package (Coron *et al.*, 2019; Delaigue *et al.*, 2019). This simple optimization algorithm, mainly based on a local optimisation, proved to be equally efficient in locating a robust optimum compared to more complex global search algorithms (Coron *et al.,* 2019) and proved effective in terms of the number of model runs required for convergence (Mathevet *et al.*, 2006). The Michel method (Michel, 1983) is based on two steps:

(i) A systematic inspection of the global parameter space is performed to determine the most likely zone of convergence. In our study, this is done by direct grid-screening.

(ii) A steepest descent local search procedure is carried out to find an estimate of the optimum parameter set starting from the best parameter set from step 1.

GR4J model parameters were calibrated by applying the Kling Gupta Efficiency (KGE) as the objective function and the daily observed data of the selected catchments as reference. The KGE is a dimensionless statistic and represents a weighting of three components that correspond to bias, correlation, and flow variability (Gupta *et al.,* 2009) ensuring that KGE is sensitive to errors in the overall distribution of streamflow (Kling *et al.,* 2012). The KGE can be expressed by Equation **3.31**,

Equation 3.31:- Kling Gupta Efficiency

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

Equation 3.32:- The bias ratio component of the KGE

$$\beta = \frac{\mu_s}{\mu_o}$$

Equation 3.33:- Correlation coefficient component of the KGE

$$\gamma = \frac{CV_s}{CV_O}$$

where r is the correlation coefficient between simulated and observed runoff, β is the bias ration, γ is the variability ratio, μ is the mean runoff in m³/s, CV is the coefficient of variation, and the indices s and o represent simulated and observed runoff values respectively.

The KGE, r, β γ have their optimum value at unity (Kling *et al.,* 2012) and the variability ratio $\frac{CV_s}{CV_O}$ ensures the bias and the variability are not cross-correlated, which otherwise may occur when for instance there is a bias in the precipitation inputs. Further discussions of the strengths of KGE can be found in Knoben et al (2019).

3.2.2.2.3. Model validation

Split-sample validation testing (Klemeš, 1986) was used to test model efficiency beyond the calibration period. For this study, 36 years (1981-2016) of streamflow data for each catchment were used, split into two equal 18-year Split-Sample Testing (SST) periods referred to as SST1 and SST2.The model was validated using direct validation (i.e., calibrated and validated using each set of reanalysis forcing data) (Figure 3.5). The value of the objective function (*KGE*) for the calibration parameter was used as the model performance statistic in the calibration as well.

1981	1982	1983 -2000	2001	2002	2002- 2016
Calibration to reanalysis & observation input			Validation to reanalysis & observation input		
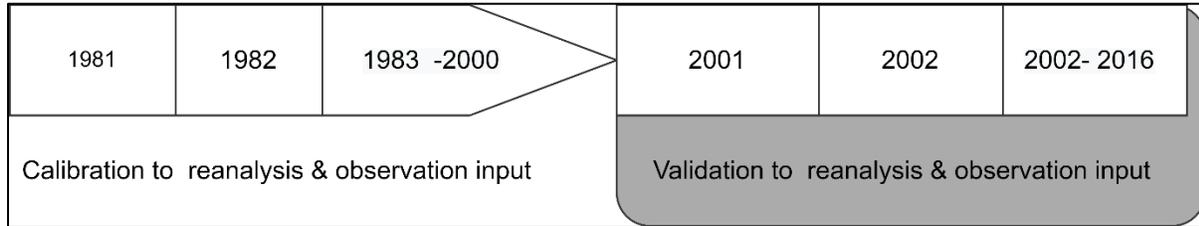
Figure 3.5: Schematic representation of direct-validation method applied in model validation experiment.

3.3.2.3. Sensitivity Analysis

Sensitivity analysis (SA) was implemented in this research to diagnostically establish how the uncertainty in the output of the GR4J model can be apportioned to different sources of uncertainty in the model input, in this case reanalyses of precipitation and evapotranspiration. Sobol' sensitivity analysis detailed in the sub-section below was applied in the sensitivity modelling experiment.

3.3.2.3.1. Sobol' Sensitivity Analysis

Sobol' Sensitivity Analysis (Sobol', 1990) is a global sensitivity method based on variance decomposition and can handle non-monotonic and non-linear models and functions, as those are usually found in hydrological models. The method estimates the relative contribution of individual model parameters and their interactions through the decomposition of model output variance. The quantification of sensitivity of a given parameter is expressed as a ratio of the partial variance to the total variance, i.e., the Sobol' sensitivity indices (SI, with a variation range of [0, 1]; Sobol', 2001), expressed as in Equation **3.34**, Equation **3.35** and Equation **3.36** respectively.

Equation 3.34:- Sobol' first order sensitivity index.

First order SI $\quad S_i = \dfrac{V_i}{V}$

Equation 3.35:- Sobol' second order sensitivity index.

Second order SI $\quad S_{ij} = \dfrac{V_{ij}}{V}$

Equation 3.36:- Sobol' total sensitivity index.

Total SI $\quad S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + ...$

71

The first order index, S_i , is a measure for the variance contribution of the individual parameter X_i to the total model variance. The partial variance V_i is given by the variance of the conditional expectation expressed by Equation 3.37.

Equation 3.37:- Partial variance of the conditional expectation of individual model parameter.

$$V_i = V\left[E\left(Y/X_i\right)\right]$$

and is also called the 'main effect' of X_i on Y . S_{Ti} measures the effect of each parameter and its interaction with the other parameters on the output (Saltelli, Tarantola and Campolongo, 2000). More details on the Sobol' indices can be found in Nossent, Elsen and Bauwens (2011).

In the research, S_{Ti} was adopted and applied as it provides reliable results in terms of the overall effect on the output for each model parameter. Following the definition in Saltelli *et al.*, (2010), S_{Ti} can be written as in Equation **3.38**.

Equation 3.38:- Sobol' total sensitivity index expresses as the ratio expected to total (unconditional) variance

$$S_{Ti} = \frac{E_{x-i}\left(V_{xi}\left(Y/X_{x-i}\right)\right)}{V(Y)}$$

where $E_{x-i}\left(V_{xi}\left(Y/X_{x-i}\right)\right)$ is the expected variance that would be left if all factors but X_i could be fixed and X_{x-i} denotes the matrix of all factors but X_i. The variance $V(Y)$ in the denominator is the total (unconditioned) variance.

Monte Carlo integrals were applied to estimate the total unconditional variance using a large number (n) of samples (i.e., model evaluations) because of the complexity and non-linearity of the environmental model. Following the Saltelli *et al.*, (2001) scheme, the R "sensitivity" package (Rosenbaum, 2015) was used to compute the S_{Ti} using a reduced sample size from $n(2k+2)$ to $n(k+2)$,where n represents the sample size, and k is the number of model parameters (in this study, $k = 4$ and $n= 14{,}000$). The choice of the sample size is based on the literature (e.g., Shin *et al.*, 2015 & Nossent *et al.,* 2011), to ensure the convergence of parameters. Therefore, in this study the total number of samples generated for the GR4J model (4 parameter model) was 84,000. A detailed description of Monte Carlo Sampling can be found in Nossent, Elsen and Bauwens, (2011).

Bootstrapping (with a large resampling size of 14,000), was applied to assess the confidence intervals of the S_{Ti} following (Efron and Tibshirani, 1994; Saltelli, 2002; Nossent, Elsen and

Bauwens, 2011). The percentile method was then applied to ensure reliable estimates of the 95% confidence intervals (Archer, Saltelli and Sobol', 1997).

The *KGE* was used as the target function for the SA, to assess the parameters related to flow simulations across the Kenyan catchments.

3.3.2.3.2. Model Suitability Index
The Model Suitability Index (MSI) (Shin and Kim, 2017) was used to assess the suitability of reanalysis precipitation datasets for hydrological modelling at catchment scale. MSI couples the model performance statistics and the SA results, unveiling the strengths and weaknesses of each of the inputs into the GR4J model. A detailed description of the index is outlined in Section 5.5.2 of Chapter 5.

3.3.3. Trend detection in flood series and shifts in flood timing

Objective (III) of this research explores the trend detection in flood series and changes in flood timing across Kenyan catchments. Three flow indices were used: maximum daily mean river flow (AMAX) and the magnitude and frequency of peak over threshold (POT) and are described in Section 6.5.1 of Chapter 6. The statistical trend analysis method, sensitivity of the trends to different flood series and the flood seasonality methodologies are outlined in the following sub-sections.

3.3.3.1. Trend analysis

The Mann-Kendall (MK) trend test (Mann 1945; Kendall, 1948) was used to test for trends in flood series described in section 6.5.3 of Chapter 6. We adopted the modified MK (m-MK), which incorporates the variance correction approach of Yue and Wang (2004). This was preferred because most of the hydrological time-series exhibit serial correlation, likely to affect the results of the MK test for significance trends (Von Storch and Navarra, 1999).The variance correction method, used to eliminate the influence of serial correlation, is preferred to the commonly used pre-whitening method (Von Storch and Navarra, 1999), because the latter is suitable only when no trend exists (Yue and Wang, 2004; 2002).

In m-MK, ranks from the de-trended series of equivalent normal variants are obtained by Equation **3.39**:

Equation 3.39:- Modified Mann Kendall ranks from the de-trended series of equivalent normal variants.

$$Z_i = \Phi^{-1}\left(\frac{R_i}{n+1}\right) \quad \text{for i = 1: n,}$$

where R_i is the rank of the de-trended series x_i^l , n is the length of the time series, and Φ^{-1} is the inverse standard normal distribution function (mean = 0, standard deviation = 1).

Maximizing the log likelihood function outlined in Hipel and McLeod (1978), gives the scaling coefficient, also known as the Hurst coefficient (H). The estimated H follows a normal distribution for the uncorrelated data, given that H = 0.5. For a given H value, the correlation matrix is obtained by the Equation **3.40** and Equation **3.41**:

Equation 3.40:- Hurst coefficient (normal distribution scaling coefficient).

$$C_n(H) = \left[\rho_{|j-i|} \right], \quad \text{for i = 1: n, j = 1:n,}$$

Equation 3.41:- Autocorrelation function of the Hurst coefficient.

$$\rho_l = \frac{1}{2}(|l+1|^{2H} - 2|l|^{2H} + |l-1|^{2H})$$

where ρ_l is the autocorrelation function at lag l for a given H, and is independent of the timescale of aggregation for the time series (Sa'adi *et al.*, 2019). The value of H is obtained by maximizing the log likelihood function of H, shown by Equation **3.42**:

Equation 3.42:- Maximizing the log likelihood function of the Hurst coefficient.

$$\log L(H) = -\frac{1}{2}\log|C_n(H)| - \frac{Z^\tau [C_n(H)]^{-1} Z}{2\gamma_o}$$

where $|C_n(H)|$ is the correlation determinant matrix $|C_n(H)|$, Z^τ is the transpose vector of the equivalent normal variates Z, $[C_n(H)]^{-1}$ is the inverse matrix, and γ_o is the variance of z_i. For different values of H, Equation 3.29 can numerically be solved and the value for which $\log L(H)$ is maximum, is taken as the H value for a given timeseries x_i. In this case, the value of H was solved following Sa'adi *et al.*, (2019), where H is taken to be between 0.5 and 0.97 with an incremental step of 0.01.

When $H = 0.5$ (for normal distribution), the mean (μ_H) and the standard deviation σ_H are used to determine a significance level of H as illustrated in Equation **3.43** and Equation **3.44** following Hamed (2008);

Equation 3.43:- The mean for the significance test for a normally distributed Hust coefficient.

$$\mu_H = 0.5 - 2.87n^{-0.9067}$$

Equation 3.44:- The standard devaition for the significance test for a normally distributed Hust coefficient.

$$\sigma_n = 0.7765n^{-0.5} - 0.0062$$

In this research, 5% significance level was used for flood series to determine the significance of H using μ_H and σ_n in Eq. (29) and (30). It is important to determine the persistence or long-term memory in a time series, which can be measured using the scaling coefficient (H). Hamed (2008) outlines this method of testing for significance of the H value. If H is significant, the variance of S ($V(S)$) can be computed following Equation **3.45** for a given H .

Equation 3.45:- The variance of the significance test of the Hurst coefficient when determining persistence in time series.

$$V(S)^{H^1} = \sum_{i<j}\cdot\sum_{k<l} \frac{2}{\pi} \sin^{-1}\left(\frac{\rho|j-i|-\rho|i-l|-\rho|j-k|+\rho|i-k|}{\sqrt{(2-2\rho|i-j|)(2-2\rho|k-l|)}} \right)$$

in which ρ_l is computed using Eq. (27) for given H and $V(S)^{H^1}$ is the estimated bias. $V(S)^H$ which is the unbiased estimate is computed by multiplying by a bias correcting factor B as shown in Equation **3.46**

Equation 3.46:- Bias correction estimate of the variance of significance test of the Hurst coefficient.

$$V(S)^H = V(S)^{H^1} \times B$$

where B is a function of H shown in Equation **3.47**.

Equation 3.47:- Correcting factor of the bias in Hurst coefficient

$$B = a_0 + a_1 H + a_2 H^2 + a_3 H^3 + a_4 H^4$$

The values of the coefficients a_0, a_1, a_2, a_3 and a_4 can be found in Hamed (2008) and are taken as functions of the sample size n . The significance of the m-MK is calculated using $V(S)^H$ in the place of $V(S)$. More detailed explanation of the m-MK method applied in this research can be found in Hamed (2008) and Saadi *et al.,* (2019).

Sen's slope estimator (Sen 1968) was used to estimate the slope or rate of change in flood series. This was used because it has been used in many hydrological time series and meteorological studies (Blöschl *et al.,* 2017). First the data pairs are calculated to derive an estimate of the slope Q following Equation **3.48,**

Equation 3.48:- Sen's slope estimator.

$$Q_i = \frac{x_j - x_k}{j - k}, i = 1, 2, ...N, j > k$$

If there are n values in the time series, many $N = n(n-1)/2$ slope estimates as Q_i are obtained. The median of these N values of Q_i gives the Sen's slope estimator. The ranking of the N values is from the smallest to the largest and the median of all slopes is obtained from the Sen's estimator as shown in Equation **3.49**:

Equation 3.49:- Median of all Sen slope estimates.

$$Q_{med} = \begin{cases} Q|(N+1)/2| & \text{If } N \text{ is odd} \\ \dfrac{Q|N/2| + Q|(N+2)/2|}{2} & \text{If } N \text{ is even} \end{cases}$$

The chi-squared with Poisson regression method was used to detect trends in flood frequency of the POT series following Mangini *et al.,* 2019. Unlike the MK test, in which the rank correlation procedure may fail in finding a hierarchy in count series containing several paired values (Vormoor *et al.,* 2016), the Poisson regression is a generalized linear regression model able to fit count series. The model assumes the counts to be Poisson distributed with the logarithm of their expected value varying linearly with time. The Chi-squared significance test assesses whether the slope parameter of the regression is significantly different from zero, which, in this case, means that a significant trend in flood frequency is detected.

3.3.3.2. Sensitivity analysis of trends to different thresholds

Using an exceedance threshold was found to affect the number in each of the flood series used. As a result, different thresholds were applied to derive POT series and a sensitivity analysis used to test for effects of threshold selection on the trends. Threshold selection was aided by the creation of different plots, such as a mean residual life plot, to determine a suitable threshold level, following Burn *et al.* (2016). The detailed procedure is outlined in Section 6.5.4 of Chapter 6 of this thesis.

3.3.3.3. Flood seasonality and shifts in flood timing

To characterize the timing and variability of flood events, seasonality measures were used (Burn *et al.*, 2010), which are defined by directional statistics (Mardia, 1972). Seasonality measures are subsequently suggested as an appropriate basis for characterizing the similarity of the flooding response of catchments. Directional statistics define the date of occurrence of a flood event by converting the Julian date to angular value (Ficchì and Stephens, 2019), where January 1 is Day 1 and December 31 is Day 365 of the flood occurrence for the event *i* following Equation **3.50**.

Equation 3.50:- Conversion of Julian date to angular value.

$$\theta_i = (JulianDate)_i \left(\frac{2\pi}{365} \right)$$

where θ_i is the angular value (in radians) for the flood date for flood event i, following the interpretation that a flood date is a vector with a unit magnitude and a direction given by θ_i . For a given sample of n flood events, the mean flood date can be given by the x- and y- coordinates using Equation **3.51** and Equation **3.52**.

Equation 3.51:- The x- coordinate of the sample mean date of a given number of flood events.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i}\cos(\theta_i)$$

Equation 3.52:- The y- coordinate of the sample mean date of a given number of flood events.

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{i}\sin(\theta_i)$$

where $\bar{x}$ and $\bar{y}$ represent the x- and y-coordinates of the mean flood date and lie within, or on, the unit circle. To obtain the mean direction ($\bar{\theta}$) of the flood dates, Equation **3.53** was used,

Equation 3.53:- Mean direction of the flood dates.

$$\bar{\theta} = \tan^{-1}\left(\frac{\bar{y}}{\bar{x}}\right)$$

where $\bar{\theta}$ can be converted back to a day of the year using Equation **3.54**:

Equation 3.54:- Conversion of the mean direction to Julian day of the year.

$$MD = \bar{y}\frac{365}{2\pi}$$

The measure of the average time of occurrence of flood events for a given catchment is represented by the variable MD, which is expected to be similar for catchments with similar hydrologic characteristics (such as size and location).

To determine a measure of variability of the n flood occurrence about the mean date, a mean resultant is defined using Equation **3.55**.

Equation 3.55:- Variability measure of a given number of flood occurrences about the mean date.

$$\bar{r} = \sqrt{\bar{x}^2 + \bar{y}^2}$$

where the dimensionless measure of the data spread is defined by $\bar{r}$ and may assume values from 0 to 1. A value close to 1 is an indication that all floods in a given sample occurred on the same day of the year, while values close to zero point to a higher variability in the date of occurrence of flood events for a catchment. It follows that the higher values of $\bar{r}$ are associated with higher regularity in the timing of flood events, hence more predictable, whereas low values of $\bar{r}$ is an

indication of higher irregularity in flood occurrence, hence lower predictability. More details of directional statistics can be found in Burn (1997), and this method was adopted for this analysis because it has been used in several flood seasonality studies across different parts of the world (e.g., Ficchì and Stephens, 2019; Berghuijs *et al.,* 2019).

To characterize a shift/change in the annual flood, the time series was split into two equal parts of 18 years each. The difference in the days between the flood timing of the first part and the second part was then circulated. The significance of these differences was tested using a nonparametric bootstrapping resampling of 1,000 replicates following Ficchì and Stephens, 2019. This was important as it provided a distribution for the mean flood timing. The significance was tested at 90% significance level for each of the year sections, by calculating the overlap (or Szymkiewicz-Simpson) coefficient of the estimated confidence intervals (i.e., the intersection divided by the smaller of the two intervals).

Chapter 4

4. Hydrological model pre-selection with a filter sequence for the national flood forecasting system in Kenya

4.1. Objective addressed and publication details

This Chapter provides a model selection with a filter sequence for flood forecasting applications in data-scarce regions, using Kenya as an example building on the existing literature, concentrating on six aspects: (i) process representation, (ii) model applicability to different climatic and physiographic settings, (iii) data requirements and model resolution, (iv) ability to be downscaled to smaller scales, (v) availability of model code, and (vi) possibility of adoption of the model into an operation flood forecasting system. Based on the review of hydrological models outlined in Section 2.4, the proposed criteria are applied following a decision tree as a filter sequence to evaluate the twelve models and provide insights on the models' possible applicability to Kenyan catchments. This chapter is important because the work presented serves as an objective model pre-selection criterion to propose a modelling tool that can be adopted in development and operational flood forecasting for the end-users of an early warning system that can help mitigate the effects of floods in data-scarce regions such as Kenya.

This work has been reviewed and published in the *Journal of Flood Risk Management* and can be accessed at https://doi.org/10.1111/jfr3.12846. The exact publication copy is in Appendix **A2** material of this thesis.

4.2. Authors' contributions

Maureen A. Wanzala (65%): - Conceptualization, Data curation- Lead, Methodology- Lead, Investigation, Software, Formal analysis - Lead, Visualization, Writing – original draft, Writing - Review & Editing. Hannah L. Cloke (15%): Supervision, Conceptualization, Methodology, Writing - Review & Editing, Project administration, Elisabeth M. Stephens (10%): Supervision, Conceptualization, Methodology, Writing - Review & Editing- Support, Project administration, Funding acquisition. Andrea Ficchì (10%): Conceptualization, Supervision, Methodology- Support, Analysis - Support, Visualization, Writing - Review & Editing- Support.

4.3. Introduction

Hydrological models predict hydrological variables, particularly river flow. In some cases, where input and output data are scarce, the model can be used to estimate the runoff and river flow in ungauged catchments (Hrachowitz, 2013; Sivapalan, 2003). Therefore, models are useful in applications such as short to extended-range flood forecasting (Alfieri *et al.*, 2013; Emerton et al., 2018), climate assessment (Tamm, Luhamaa and Tamm, 2016; Hattermann *et al.*, 2017; Lu *et al.*, 2018), hazard and risk-mapping (Artan *et al.*, 2001; Ward et al., 2015), drought prediction (Van Huijgevoort *et al.*, 2014), and water resource assessment (Dessu et al., 2016; Mutie, 2019; Praskievicz & Sang, 2009; Sood & Smakhtan, 2015). However, the application's ability to extract

viable information varies across different classes of models, according to their different spatial and temporal scales and intended purposes.

The choice of model for operational flood forecasting is not simple because of different process representations, data scarcity issues and propagation of errors and uncertainty down the modelling chain (e.g., Paul *et al*., 2020; 2019a). For example, the practice of choosing a model for an application may be difficult due to several reasons highlighted in Melsen, (2019): (i) popular models are not tailored to specific climates or circumstances (unless the west European climate counts, implicitly), which makes exclusion on process presentation alone difficult; (ii) most popular models share the same main properties and the same weaknesses; (iii) the community has failed to create a generalized benchmarking system to rank models and model set-ups, so that suitability has to be ascertained on a case-by-case basis; finally, model evaluation is primarily based on streamflow, which in itself is too little to distinguish between models, especially calibrated models. It is necessary for a modeller to know the perceptual model (Wagener *et al.*, 2021), including quantitative or qualitative descriptions of the existing knowledge and understanding of the catchments (Beven, 2011; Gupta *et al*., 2008; Westerberg et al., 2017). For instance, Wegener *et al*. (2021) illustrates a generic perceptual model included in catchment hydrology functions. The processes herein are dynamic and evolve with time in response to changes in water management or land-use, climate conditions and geomorphological changes, so they need to be integrated into the model development. This implies that if such changes are not taken into consideration during and/or after model development and upgrade, then the relevant processes will not be presented adequately, thus limiting the application of a single model over the entire country.

Models are simplifications of reality and thus cannot completely represent every process and aspect of the catchment. The importance and impact of many processes can evolve with time, for example in response to changes in water management. In addition, the right approach now is not necessarily the right approach in the future. Significant buy-in is required to develop operational forecasting capacity with a specific model, and so in recognition of changes in the importance and impact of many processes because of land use change, water management, etc., it may mean it is more efficient to choose a modelling approach that can represent a larger range of processes. When there are distinct zones of hydro-climatology within a country it could be necessary to adopt different modelling approaches, but this needs to be balanced against the scaling up of the resources required to extend human and technical capacity across several different models.

Moreover, data play an important role in hydrological modelling, irrespective of the processes represented in a model (Wahren *et al.*, 2016). Many studies point to challenges in modelling due to data scarcity (e.g. Beck *et al*., 2017; Fuka *et al*., 2014; Lavers *et al*., 2012; Najafi *et al*., 2012; Quadro *et al*., 2013; Smith *et al*., 2013; Wu *et al*., 2013) which limits the applications of very detailed and complex models due to inherent unquantified uncertainties. Recognizing that data and models are not free from error, for the sake of brevity, within this paper our descriptions of the models and their characteristics will consider only those uncertainties related to model structure (Pechlivanidis et al., 2011; Smith et al., 2015).The choice of model depends on its intended purpose, and the modeller needs to select a model objectively, based on the end-user's need for

more reliable decisions (Perker, 2020; Boelee et al., 2017; Todini, 2007). Various hydrological models exist at different spatial and temporal scales with diverse levels of complexity and data requirement. Additionally, there are differences between model codes and implemented modelling systems, which may cause difficulties in the choice and application of a particular model. A Multi Criteria Analysis (Sherlock and Duffy, 2019) is recommended to evaluate and grade models, from which a small number of models would be constructed, calibrated, and tested in a real-world context and, at the end, a choice would be made of one or more models to be used in the operational FFC experiment. However, the proposed MCA relies heavily on evaluation data, and is very time consuming because of the number of models available: hence, for data-scarce regions, and/or agencies with limited resources, and in general use, an additional decision tree is helpful to trim down the number of options. There is a need for further evaluation of the limited selection with, for example, an MCA and the FFC experiment. To aid this hypothetical modeller there is a clear need for well-conceived and systematic strategies for selecting model structures and establishing data requirements, which is the innovative contribution of this research.

There is a plethora of model reviews with applications at global and continental scale, including Devia *et al* (2015 Emerton *et al* (2016 Pechlivanidis *et al* (2011 Salvadore *et al* (2015 Sood *et al* (2015 Trambaur *et al* (2013 and Kauffeldt *et al.* (2016). Most of these reviews highlight and compare existing modelling concepts and gaps but none has focused on model selection frameworks for final application except for Trambaur *et al*. (2013) and Kauffeldt *et al*. (2016). Kauffeldt *et al*. (2016) provides a technical review of large-scale hydrological models for implementation in operational flood forecasting at continental level. Trambaur *et al.* (2013) reviews continental scale hydrological models highlighting their suitability for drought forecasting in sub-Saharan Africa. The two cited works look at model review and a selection framework for flood and drought applications at continental scales respectively, and to the best of my knowledge this is the first model overview and practical objective model selection framework for flood applications at national scale taking into consideration varied catchment characteristics and data scarcity issues.

This paper proposes a practical approach building on Kauffeldt *et al*. (2016) and Trambaur *et al*. (2013) for selecting a model based on a step-by-step filter sequence following objective aspects (such as the ability to simulate relevant processes to flood applications), as well as considering more practical aspects such as model code availability and ease of use at catchment scale with varied climate characteristics. We follow the filter sequence and develop a Venn diagram to select suitable model candidates. This practical approach is applied to a case study of the development of an early warning system that can help mitigate the effects of floods in data-scarce regions within Kenya, where there is a lack of good observations of climate variables such as precipitation, temperature, etc., and this factor inhibits the proper identification of the limitations of model applications at catchment scale.

Our paper is structured as follows. Kenyan hydrology and applications of hydrological models to simulations of flood processes are discussed in Section 4.4. The decision tree is built based on deliberations about Kenyan hydrology and current forecasting experience in Kenya, which are outlined in Section 4.4.1. The presented model selection framework is outlined in Section 4.5.

Then a detailed evaluation and selection of the models based on the decision tree are outlined in Section 4.6. In Section 4.7, the focus is on specific discussions regarding model selection and the novelty of the decision tree. The paper then concludes with the key contributions of the suggested pre-selection along with recommendations for the next steps to evaluate the models objectively to improve FF in Kenya.

4.4. Kenyan hydrology and forecasting

4.4.1. Applying hydrological forecasting models to the simulations of floods in Kenya

It is important to consider the application of the hydrological model when determining which model to use, due to differences in process generation and representation (Cloke *et al.*, 2011). For example, floods are generated by a range of processes related to extreme rainfall (interception, through-flow), runoff generation (infiltration, saturation excesses and subsurface storm flow) and runoff routing (Rosbjerg *et al.*, 2013). In addition, floods in snow dominated catchments are regularly caused by snow melt, and consequently representation of this process in a hydrological model is crucial, because it requires an optimal simulation of the snow related hydrological processes such as snow accumulation and melt (Verzano, 2009).However, this case does not apply to Kenyan catchments.

Moreover, flood formation is a complex combination of extreme precipitation or temperature rise or both, the retention of the water in different storages and finally the flow through the river networks. Hydrologic responses commonly exhibit threshold behaviours in terms of process sensitivity e.g. soil storage, evaporation, evapotranspiration and interception as a function of system (model) state, both in space and time, giving rise to *hydrology uniqueness* (Beven, 2000) A flood peak caused by extreme rainfall in the upstream part of a catchment naturally reaches the downstream part of the catchment after a temporal delay (Verzano, 2009; Tallaksen and Lanen, 2004). Therefore, several effects influence the magnitude of the flood wave in the downstream area, such as tributary contributions and retention in lakes and wetlands. The lateral transport of water through the river network is a particularly important process for the routing of discharge. This applies to average flow conditions as well as low or high flows. Therefore, it is meaningful to route the water within a hydrological model with a variable flow velocity because the flow velocity varies with the actual river discharge (Verzano, 2009) among other relevant flood generating processes. In many hydrological forecasting systems, the treatment of the rainfall-runoff component (traditionally the core of what is meant by hydrological models) and the routing can be separated. Whether the routing should be built in, or should be specifically modular, could be another criterion that qualifies the models under consideration. An operational Flood Forecasting System (FFS) aims at producing accurate timely and valuable flood forecast information far in advance to reduce flood-related losses by increasing preparation time. A typical FFS requires a hydrological model and data sources, as well as main processes and an interactive friendly user interface. For example, a simplified conceptual model for a large-scale flood forecasting system, the components required, and the output generated within each component. Is shown in Figure 4.1.
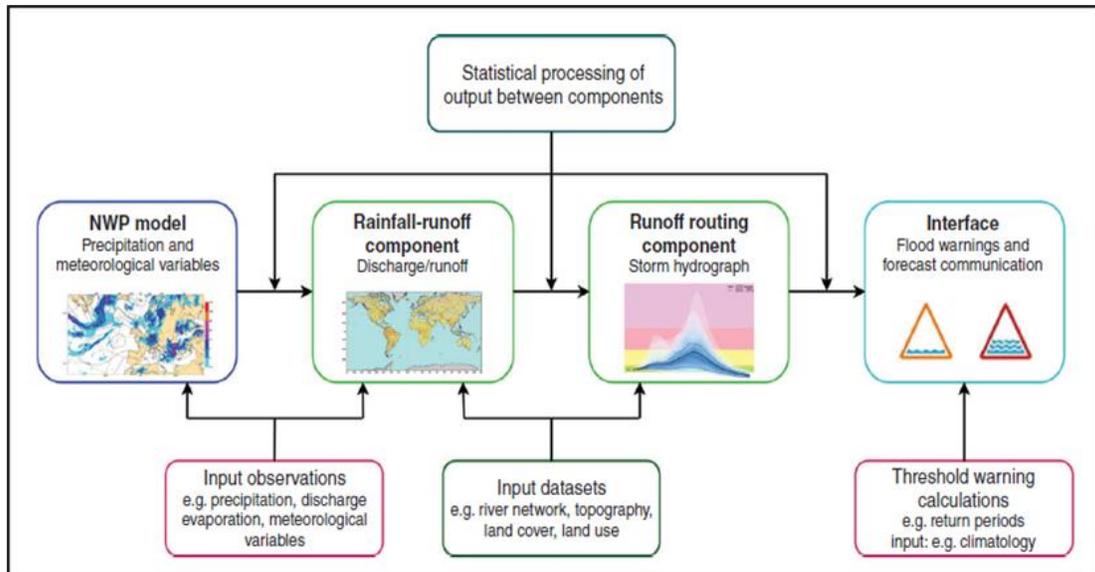
Figure 4.1:- Conceptual large-scale hydro-meteorological forecasting system (Emerton et al., 2016).

4.4.2. Subsequent argument for the need for a decision framework with a filter sequence in Kenya

Both the hydroclimate and human influences create challenges for hydrological modelling and forecasting (Bai *et al.*, 2015) because of their massive influence on the catchment processes. For example, Kenya exhibits high variability in physiographic and hydroclimatic conditions (see Figure 4.2). The highest point is at about 5000m a.s.l. (mostly areas around central highlands), while the lowest point is about 20m a.s.l. (mainly around coastal areas). The vegetation cover is mainly a mixed tree cover, grass, and sparse vegetation in most of parts of the country and shrubs and bare land in the arid and semi-arid areas of northern Kenya. As a result, Kenya experiences different climate related extremes in terms of intensity, magnitude, and timing.
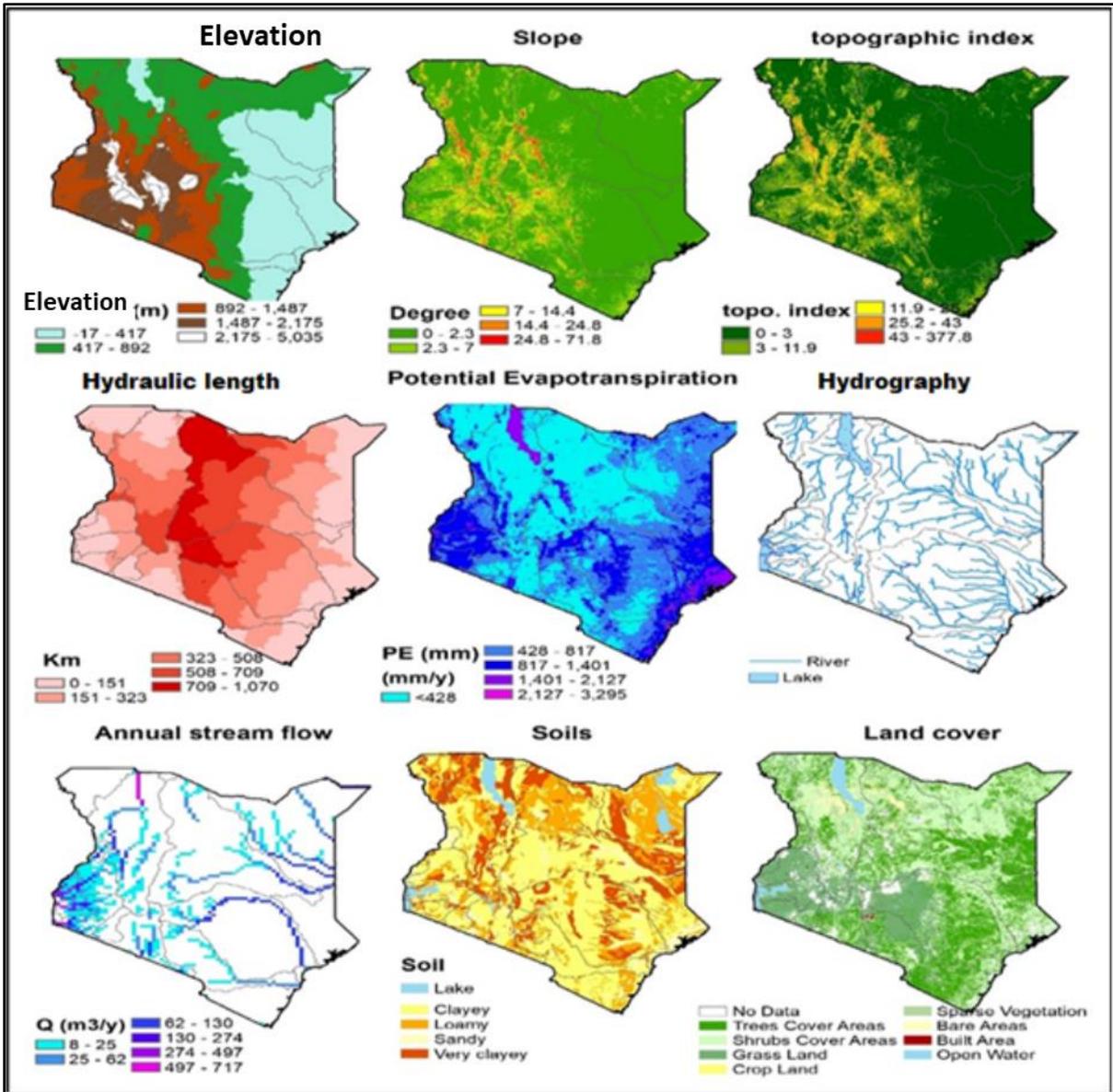
Figure 4.2:- Physiographic and hydroclimatic characteristics of Kenya.

Rainfall pattern follows a bimodal rainfall seasonality (Ongoma and Chen, 2017) with high spatiotemporal variability (Figure 4.3) (Hession & Moore, 2011). Three seasons are experienced: the 'long rains' of March - April - May (MAM), the non-rainy months of June - July - August (JJA), and the 'short rains' of October - November - December (OND) (Ogallo, 1988; Ongoma *et al.*, 2015). About 42% of the total annual rainfall is observed during the MAM rainfall season (Ongoma and Chen, 2017), with the highest intensity observed near the water bodies of the Indian Ocean, Lake Victoria, and the Kenyan highlands
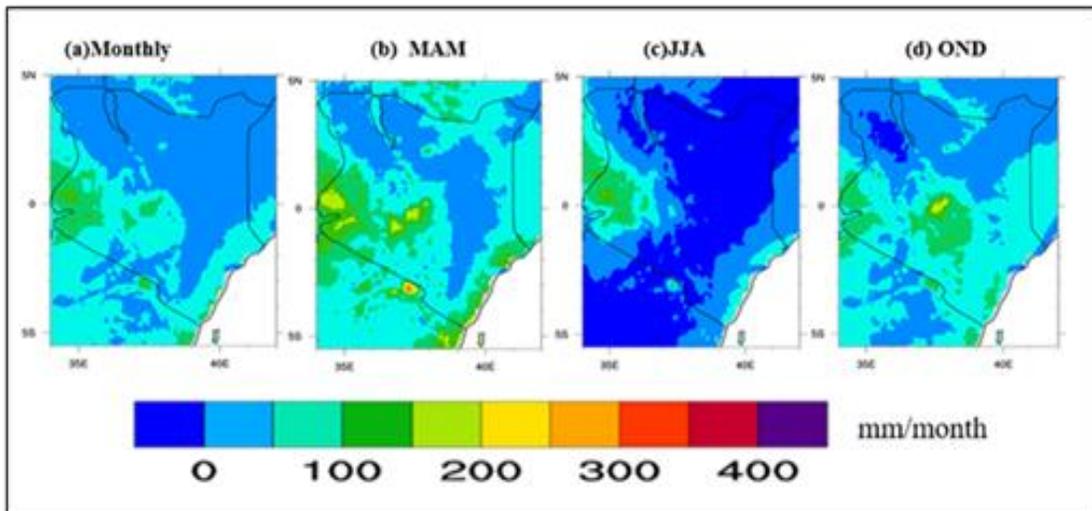
Figure 4.3:- Spatial pattern of long-term mean monthly and seasonal rainfall over Kenya

There are five major basins (Marwick *et al.*, 2014) in Kenya (see Figure 4.4 left panel). These catchments are highly influenced by settlements as well as human activities such as dam constructions and irrigation activities (Figure 4.4, right panel), which have adverse effects on the catchment response to rainfall runoff processes. At the catchment scale, there is high variability in catchment hydroclimatic characteristics such as surface area and average annual rainfall (Figure 4.5).
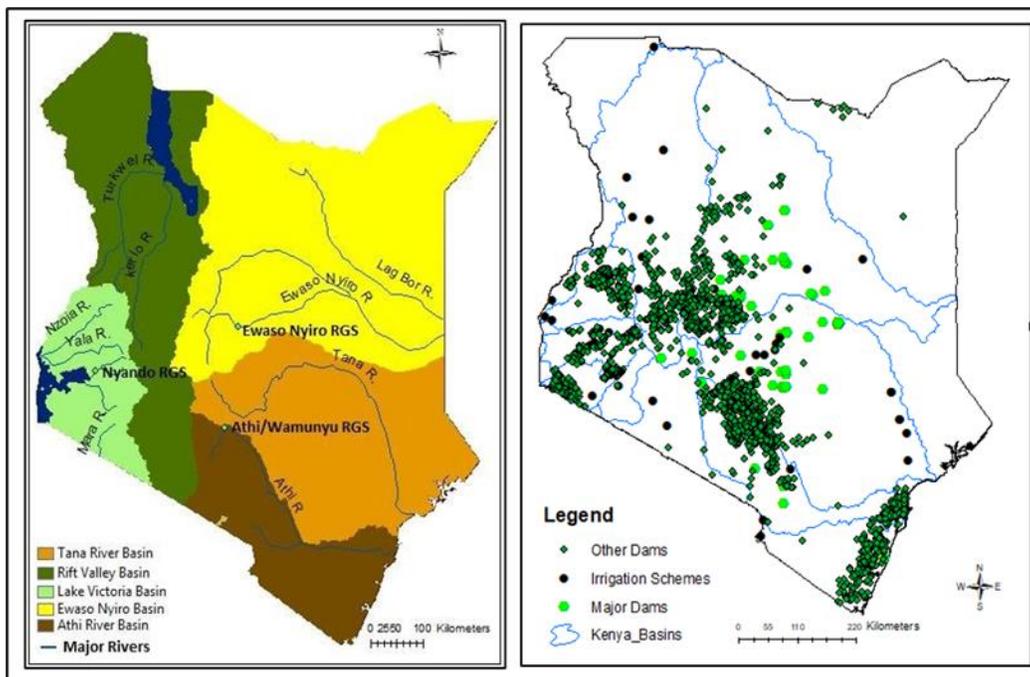
Figure 4.4:- Standard Kenya main basins (left panel) and the ongoing human activities (constructed major and small (other) dams and irrigation schemes) in the select catchments (left panel) (Data source: WRA-K).
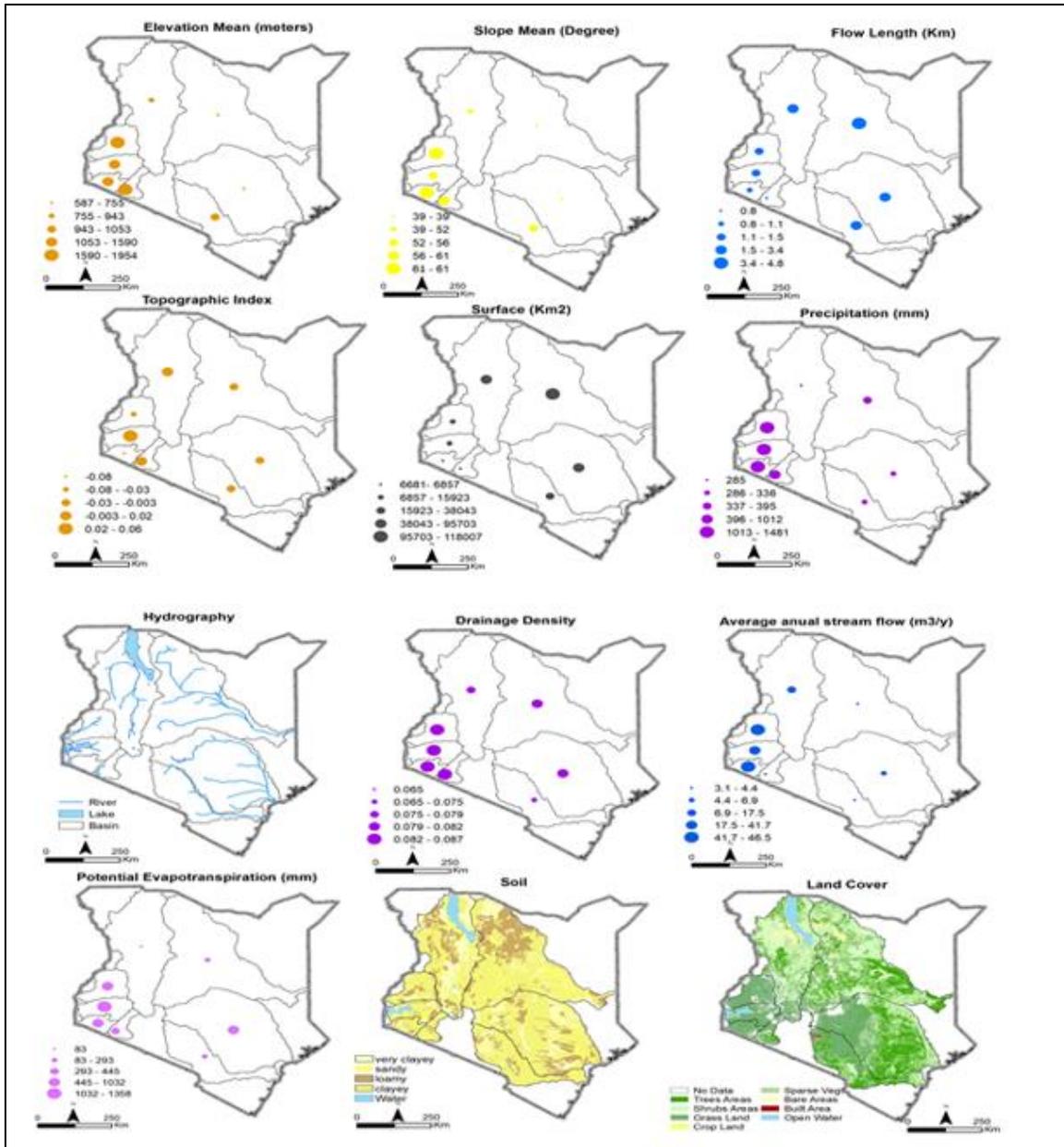
Figure 4.5:- Spatial distribution of the morphological and hydro-climatic characteristics across the catchment. The catchments were delineated based on the 19 outlet discharge points.

Therefore, it is important to consider the variability in catchment characteristics and the knowledge gaps in the perceptual model (e.g., landcover changes, human activity, data uncertainty and accounting for groundwater fluxes) when selecting a model for application as this may influence the performance of the model. The following section discusses the aspects to consider when objectively pre-selecting a model for application to Kenyan catchments. We focus on the important aspects, highlighting the problems and challenges associated with hydrological models which are specific to Kenya, and draw clear conclusions on model requirements from the basin characteristics in Section 4.4. We apply the proposed criteria to Kenyan select catchments, outlining characteristics associated with catchments in Section 4.6 as well as providing a detailed

review of the models and a subsequent evaluation. In Section 4.7, we discuss the strength and gaps in the selection framework and provide a way forward and the next steps for future research. Section 4.8 draws a clear conclusion and defines the contribution of this work to a national flood forecasting centre.

4.5. Model Selection Framework

Selection framework in this paper follows a selection criterion including the ability to represent relevant processes, the model's structure, flexibility, and complexity, the availability of the model code and the needs of the user community (Bennett *et al.*, 2013; Kauffeldt *et al.*, 2016), and consequently it is qualitative rather than quantitative. For example, a good model should be able to represent all relevant processes, such as gross precipitation (snow, rain), interception storage, evaporation, transpiration, snowpack storage, snowmelt, overland flow, soil storage, recharge to shallow aquifer, capillary rise, intermediate flow, baseflow, and leakage to deep aquifer. However, these will require relevant input datasets and more complex models (e.g., fully distributed with numerous parameters) to effectively represent the processes, but it is worth note that increasing the model's complexity by incorporating all the above processes does not necessarily improve its performance (Butts *et al.*, 2004; Birkel *et al.*, 2010). The application and performance of a model may also vary depending on the site's size and other characteristics (Van Lanen *et al.*, 2013; Bai *et al.*, 2015). Therefore, the following sections summarise the aspects that aid the objective selection of a hydrological model for flood applications in the Kenyan context, considering Kenya's hydrogeological, physiographic and climatic conditions discussed in section 1. In total six criteria were found to aid in the decision making. In the next subsections each of the six criteria is evaluated in relation to Kenya.

4.5.1. Represented processes and fluxes

A complete hydrological model would represent all the water balance components and fluxes (e.g., as illustrated in Mendoza *et al.* (2012)). Complexity in models often results in the need for many parameters to be determined, which requires more hydrogeological data (Dobler and Pappenberger, 2013; Muleta and Nicklow, 2005). There needs to be a compromise between complexity and efficiency for a model to work.

More data is needed to make more complex models more accurate. The choice of an appropriate model structure is a crucial step towards the accurate prediction of streamflow or other variables, and to comprehension of the dominant physical controls on catchments' responses to climate change (Clark *et al.*, 2008). In Kenya, this requires more data, on matters such as groundwater level, which are not readily available.

Some catchments, especially those in the arid and semi-arid regions of Kenya, have sandy and rocky riverbeds and tend to run dry most of the dry months. To such as these, the fixed velocity and river channel fields represented in some hydrological models may not apply. This is because of failure to properly represent the roughness index, which varies not only with boundary characteristics but also with flow velocity, water depth, and other hydraulic factors (Addy and Wilkinson, 2019; Zang *et al*., 2016).

In addition, the more processes are represented in a model, the more the parameterization schemes. For example, *a priori* estimation requires establishing parameter values from measured physical system properties, presupposing that the model parameters have a sufficiently reliable representation (Beven and Pappenberger, 2003). Therefore, parameter estimation in models of natural systems may require measurements and tests. It then follows that effective calibration for such model parameters requires more computational power, which may be lacking in the Kenyan operational flood forecasting centre.

4.5.2. Model applicability to Kenyan hydroclimatic conditions and physiographic settings

Processes that are most relevant for simulating flood conditions in Kenya (see Barasa *et al*., 2018; Onyando *et al*., 2003) should be represented in a model. Some extra processes, such as channel losses, evaporation from rivers, and wetlands representations, are not considered important in average conditions in some regions due to complexity or lack of interest (Rosbjerg *et al.*, 2013), so they can be discounted. This is because models incorporating such complex processes require more skilled personnel and higher budgets to install and run. This is a challenge in most operational systems in developing countries including Kenya. Temperature plays an important role in river channel and catchment evaporation. In Kenya's case, annual mean temperatures range from 15 to 35 °C which closely correlates with topography, with the lowest temperature experienced in the central highlands and the highest in the lowlands (Mutimba *et al.*, 2010) and a model incorporating this would be best suited for such a place.

Model selection in dry and wet catchments must be more careful, due to the large performance difference in dry catchments (Bai *et al.*, 2015). Wet catchments runoff simulation is significantly better than that in dry catchments (Haddeland *et al.*, 2011), because of the high non-linearity and heterogeneity of rainfall–runoff processes (Atkinson, Woods and Sivapalan, 2002). In addition, high uncertainty is introduced during model parameter estimation, resulting in significant differences in simulated runoff behaviour (Andersson *et al.*, 2015). Large river basins are often strongly influenced by human activities (e.g., irrigation, reservoirs, and groundwater use) for which information is rarely available (Döll, Fiedler and Zhang, 2009). The Kenyan case, where most basins are ungauged, may increase such uncertainty (Sivapalan *et al.*, 2003; Hrachowitz and Weiler, 2011).

When there are distinct zones of hydroclimatology within a country, it could be necessary to adopt different modelling approaches, but this needs to be balanced against the scaling up of the resources requiring human and technical capacity across several different models, which is one of the main challenges in the Kenyan case.

4.5.3. Data requirements and spatial and temporal resolution of the model

Kenya suffers from a lack of good observations of climate and hydrological data. This is a limiting factor for the proper identification of the limitations of model applications at catchment scale. For example, a detailed representation of groundwater flows and tables and soil moisture content would be very relevant to flood forecasting. However, reliable data (on ground water and

reservoirs, for example) are unavailable for research applications, thus limiting the use of models incorporating this kind of data. As a result, a compromise must be reached regarding model spatial variability, due to the ungauged status of most Kenyan catchments (Trambauer et al., 2013), to allow the use of alternative freely available remote sensing data. Applying a distributed model would require high spatial and temporal resolution data to represent each of the catchment HRUs whereas a lumped conceptual model would represent an entire river basin (Krysanova, Bronstert and Müller-Wohlfeil, 1999), but since gauging stations are sparse or absent in some of the catchments, this limits the use of most distributed models across Kenyan catchments. The issue of observation data scarcity in Kenya is huge and the reliability of the readily available satellite and reanalysis datasets has not been fully established due to low performance skills in the evaluation studies highlighted in Section 2.5. This limits application of some of the distributed models with observations due to requirement of more input data, even though this can be solved with satellite and reanalysis data. In addition, the existence of some regionalisation technique that can make models work well in data scarce areas could be an alternative, but the expertise and skill to implement this is lacking at the Kenya national level.However, limiting the models to the type that can run only when directly calibrated on an outlet would be a mistake. This is because there are plenty of ways to discretize in HRUs without individually calibrating each HRU independently. There are ways to calibrate transfer functions to enable modelling and ungauged HRUs (Samaniego, 2010). There are model set-ups that do not rely on calibration as a first principle (such as wflow-sbm, Imhoff, 2020) and are based on globally available data. The challenge here is the transferability of the model to suit Kenyan catchments and operations and represent the catchment processes adequately, because it needs to be as simple as possible.

Moreover, modelling experiments on Kenyan catchments may yield more plausible results if data at high frequency time steps are used as they contain more information (Ficchì, Perrin and Andréassian, 2016). This is because the better modelling of the rainfall–runoff relationship is severely affected by the sub-hourly dynamics of precipitation(Paschalis et al., 2014) due to the nonlinear nature of the infiltration process (Blöschl & Sivapalan, 1995), such as the peak discharge value (Gabellani et al., 2007) and runoff volume (Viglione et al., 2010). In Kenya, the temporal resolution of the available reliable data may be limited to higher time steps (such as monthly and yearly) and this may limit the application of a model on sub-daily/hourly timesteps. Models incorporating higher timesteps data, such as daily and monthly, are more easily applicable to the n Kenyan case compared to those limited to hourly or sub-daily timesteps.

4.5.4. Capability of the model to be downscaled to a river basin scale

The issue of scale in hydrological models is highlighted in Beven (1995), where the aggregation approach towards macroscale hydrologic modelling is an inadequate approach to the scale problem. For semi-distributed and distributed models, grid size selection is intricately linked to the spatial scale at which the model will be applied. Also, when lumped approaches are applied to considerably larger basins the integration of the processes will naturally occur over a greater area, and thus any differences in small scale processes within the basin will not be well considered. Due to lack of locally developed models, the continental models are applied at catchment scale, and therefore need to be downscaled to suit the grid size under application. However, for larger grids,

processes that are important only at the local scale (such as overland flow) may not be considered in the model structure unless there is an extensive change in the model grid width, and this may at times introduce structural uncertainties. Some models may not be easily downscaled to Kenyan river basins with varying spatial scales (see Table 7) without making significant changes in the structure of the model.

4.5.5. Operational model for flood early warning system at large scales with potential adoption to local scale

With the increase in flood events in Africa in the recent past, Thiemig *et al*. (2015) proposed a FFS for Africa, hereafter referred to as AFFS. Following the illustration, LISFLOOD, a physically based hydrological model, was selected to be the AFFS; it relies on historical hydrological observations, historical as well as near real-time meteorological observations, real-time meteorological forecasts, and an African GIS dataset. The four main processes AFFS runs are: the calculation of hydrological thresholds, the computation of the initial hydrological conditions, the computation of the ensemble hydrological predictions, and the identification of flood events. This was developed as a protype for Africa but never taken forward to operations and since then no literature or research on the efficiency or applicability of this system has been documented.

In addition, Princeton University has developed the African Flood and Monitor (AFDM) tool (Wood, 2015). The aim is to demonstrate the potential for tracking drought conditions across Africa using available satellite products and modeling in data-scarce regions. The system provides daily updates in near real-time (2-3 days lag) of surface hydrology, streamflow and vegetation stress, short-term hydrological forecasts for flooding, and seasonal forecasts for drought and agricultural impacts (https://platform.princetonclimate.com/platform-ng/pca/products).

The system has been installed at regional centers in Africa, most notably in West Africa (ACMAD), where it is operational for the Niger basin using the Hype-Niger model and the World-Hype applied to the whole West Africa region. A schematic illustration of the FFS for the Niger Basin in West Africa is shown in Figure 4.6.

Narrowing the focus down to Kenya, the Kenya Meteorological Department (KMD) runs an operational flood forecast system in Nzioa basin (Personal communication from Andrew Njogu) with plans underway to upscale to 9 additional flood prone areas spread across the other 7 basins (Athi, Galana, Sabaki, Nyando, Tana, Sondu, Ewaso Ngiro etc). A schematic representation of the FFS in River Nzioa Basin in Kenya and the steps involved appears in Figure 4.7. The model adopted for this system is the Soil Moisture Accounting and Routing Model (SMAR) incorporated in the Galway Flow Forecasting System (GFFS) (O'Connor, 2005). The GFFS is a suite of models developed at the Department of Engineering Hydrology at the National University of Ireland, Galway, Ireland. The five models embedded in the software comprise four system theoretic models (simple linear model (SLM), linear perturbation model (LPM), linearly varying gain factor model (LVGF) and artificial neural network (ANN)) and one conceptual model: the soil moisture accounting and routing (SMAR) model. The ordinary least square solution for SLM, LPM, and LVGF, conjugate gradient algorithm for ANN and Rosenbrock, and simple search and genetic

optimization methods for SMAR are used for calibration of the model parameters (O'Connor, 2005).
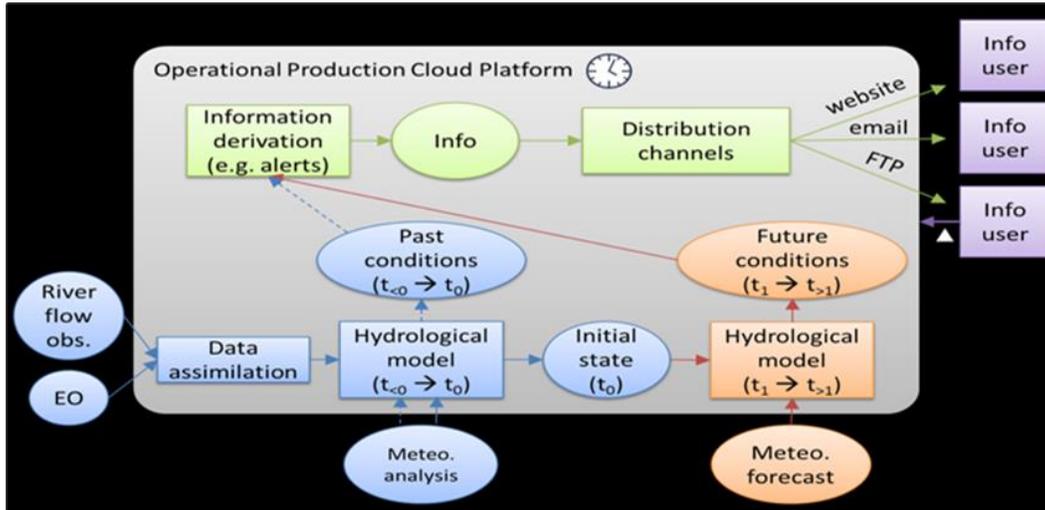
Figure 4.6:- Niger HYPE-model for Niger basin and Hype World for the rest of West Africahttps://fanfar.eu/production/

In an interview, Njongu – head of hydrological modelling section at the KMD, noted that the choice and use of the SMR model was an entirely subjective matter, mainly driven by the project funding following the push to implement a FF system in Nzioa after destructive flooding events. Additionally, he noted that there was limited documented research on efficiency assessment to inform the choice of the SMAR model adopted for this purpose; instead, it was chosen for its simplicity and comparatively low data requirements. Moreover, model choice is dependent on the project funds available, and the implementers and collaborators are likely to trial a model they chose to advance their own interests and advocate for the application's abilities, irrespective of its underlying performance measures. It then follows that the choice and application of the SMAR model in the Kenyan FFS was due to this reason.
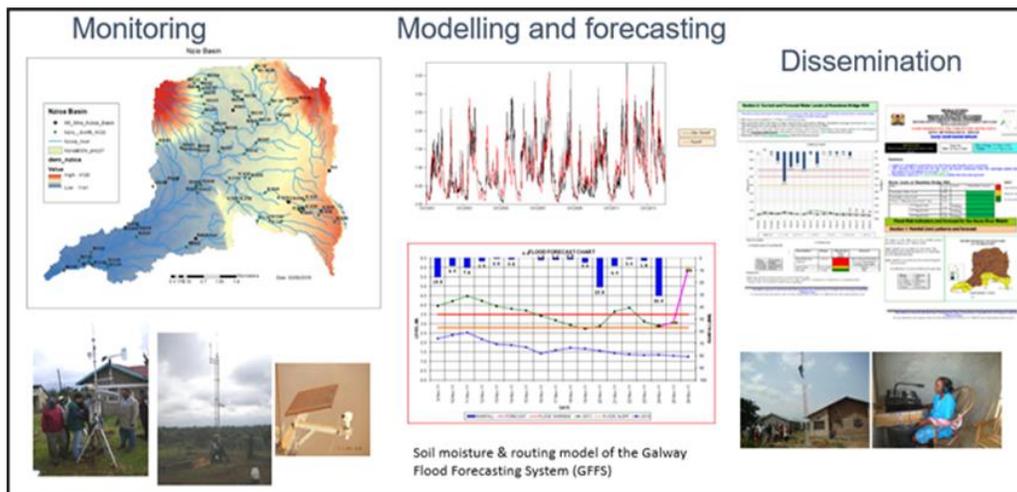
Figure 4.7:- Overview of the flood monitoring, modelling, forecasting and dissemination for the operational flood forecasting in Nzioa basin, Kenya (Source KMD).

91

Because of current developments, there have been ongoing initiatives spearheaded by the Kenya Water Resources Authority – a parastatal body mandated to set and manage the water resources rules and hydrological data. Under the ongoing Kenya Water Security and Climate Resilience Project, WRA is in the process of trialing three hydrological models (SMAR, NAM, GR4J) in Nzoia to be incorporated into the FFEWS under development (WRA reports). For example, an initial assessment for model performance in Nzioa basin has been started. Soil moisture representation in SMAR, GR4J and MIKE NAM over the basin is illustrated in Figure 4.8.
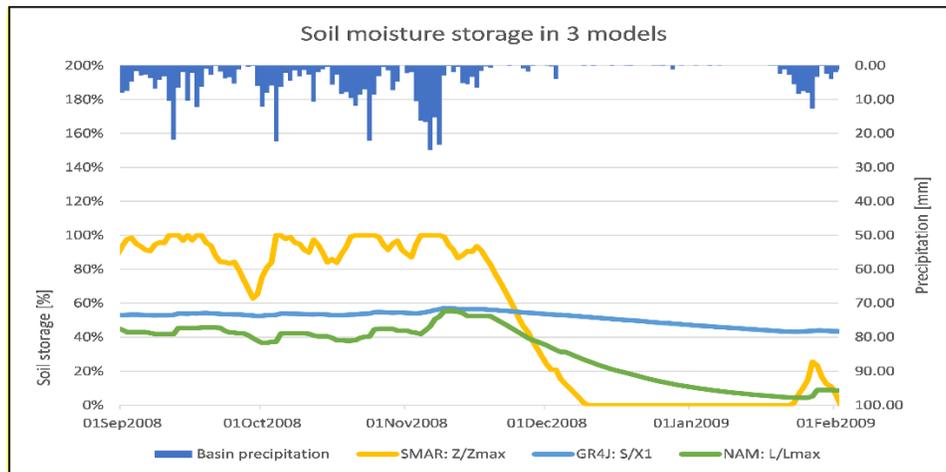
Figure 4.8:- A representation of the soil moisture evaluation in SMAR (yellow), GR4J (blue) and MIKE NAM (green) over the Nzioa basin in Kenya (Source KMD).

The above highlights point to fact that a model needs to be able to be incorporated into an operational (up and running) system, if the main aim of the model selection is to provide a tool for the end-users of an early warning system that can help to mitigate the effects of floods. In this respect, a model that can easily be implemented in a forecasting environment is preferred. Hence, the model should be stable, have reliable error and inconsistency checks, be able to flag up missing data (e.g., when input sources fail), be able to fit into an operational environment and preferably be user friendly.

4.5.6. Availability of model code and model run-time.

The code must be available for use (open source or through agreements) with the potential for adaptation to specific purposes (e.g., ability to change the represented processes, ingested time-step and/or catchment discretization). These adaptions are possible but do not actually existing most of the freely available model codes. The code must be actively used and developed with core developers identified to ensure that proper support can be given in the initial phases. Executable code is not enough, since changes, for instance reading of input data, will be necessary (Paul *et al*., 2020). Forecast deliveries run the risk of being delayed if bug fixes or updates cannot quickly be incorporated in the model. A key aspect here is the service level agreement struck between the model and the forecasting system provider, outlining a clear overview of which parts are maintained locally, and which parts are outsourced. In addition, codes available only through purchase may limit the use of models, especially for research and operational purposes, so the model should be opensource; not all open-source licenses, however, are the same.

Some modelling communities have provided accessible packages for some select models with dedicated functions, such as input data preparations, data processing and transformation, calibration, etc. These packages may have all functionalities for the application under consideration, but limited functionalities would limit it use. Some of the freely available packages, the proposed models and the package functionalities that can be executed are summarised in Table 6.

The model run-time (Central Processing Unit) - computational time to run a simulation from model spin up - varies with different models and areas of application. For example, Astagneau et al. (2021) show how different models and implementations can differ by an order of magnitude in required calculation time for the same set of catchments. There are shortages of computational power in many of the African National Meteorological and Hydrological Services (NMHS), especially if ensemble simulations, data assimilation methods and further computationally intensive uncertainty estimation methods are to be applied, and Kenya is not an exception.

Table 6:- Freely available packages, proposed models, and inbuilt model functionalities of some of the commonly applied models (Source: Smith et al 2021).

Package	Repository	Proposed Model	Package functionalities						
			Data processing	Criteria	Data transformation	Automatic calibration	Plot function	Graphical user Interface	Independent snow function
airGR	R	GR models	√	√	√	√	√	√	√
dynatopmodel	R	Dynamic TOPMODEL	√	√	×	×	√	×	×
HBV.IANIGLA	R	HBV	√	×	×	×	×	×	√
hydromad	GitHub	IHACRES AWBM GR4J Sacramento	√	√	√	√	√		√
sacsmar	GitHub	Sacramento	×	×	×	×	×	×	√
topmodel	R	TOPMODEL 1995	√	√	×	×	×	×	×
TUWmodel	R	Modified HBV	×	√×		√×	×	×	×
WALRUS	GitHub	WALRUS	√	√	√	√×	√	√×	√

√- Present ˟ - Not present, √˟ -Partially present

4.6.   Application of the selection framework to Kenya's chosen catchments

The above section outlines the aspects to consider when selecting a suitable model for national flood forecasting and application in Kenya. The application of the selection framework to Kenya, based on the above proposed criteria, is outlined in Table 7. There are marked differences from catchment to catchment, which point to the fact that a single model and single initialization with all the same parameters cannot be efficiently applicable at country level but rather at catchment scale, so it is necessary to ensure that a model is selected according to objective criteria, based on the users' needs and the catchment processes.

Table 7:- Summary of catchment-by-catchment evaluation based on the proposed framework.

Catchment by catchment evaluation based on the proposed framework									
Catchment Name	EWASO NG'IRO	TANA RIVER	ATHI	NZIOA	YALA	NYANDO	TURKWEL	GUCHA	MARA
Catchment Area (KM²)	30,000	96,000	46,600	12,800	2,777	5,110	9, 303	6,310	13, 750
Dominant Hydrological Processes fluxes present									
Precipitation	√(unimodal)	√ (bimodal)	√(bimodal)	√ (bimodal)	√(bimodal)	√(bimodal)	√ (unimodal)	√(bimodal)	√ (bimodal)
Infiltration	√	√×(good upstream, poor in delta)	√×(good upstream, poor in delta)	√×(good upstream, poor in floodplain)	√	√ ×	√	√	√
Interception	√ ×	√	√	√	√	√	√ ×	√	√ ×
Evapotranspiration	√ (low)	√ (low upstream, high)	√ (low)	√(high)	√(high)	√(high)	√(low)	√(high)	√(low)
Snow	×	×	×	×	×	×	×	×	×
Soil storage	√ ×	√ ×	√ ×	√	√	√	√ ×	√	√ ×
Ground water storage	×	√ Shallow & deep	√ Shallow & deep	×	×	×	√ Shallow & deep	×	×
Lake & reservoirs	×	√ Linear res.	√ Linear res.	√	√	√	√ Linear res.	√	×
Runoff	√ infiltration excess	√	√	√saturation excess/	√saturation excess	√infiltration excess overland	√infiltration excess	√	√
Ground water recharge	×	√	√	√	√	√	×	√	√
Hydroclimatic and physiographic setting									
Arid	√	×	×	×	×	×	√	×	×
Semi-arid	√	√ (upstream)	√(upstream)	×	×	×	√		√
Wetland	×	√(wet delta)	√(wet delta)	√	√	√	×	√	×

	Data Availability and resolution								
Observed Meteorological data (Precipitation, Temperature, wind etc)	√×Daily √monthly	√×Daily √monthly	√×Daily √monthly	√×Daily √monthly	√×Daily √monthly	√×Daily √monthly	√×Daily √monthly	√×Daily √monthly	√×Daily √monthly
Hydrological data (Discharge, Ground water Reservoir levels,)	√× discharge for some stations	× Ground water & reservoir levels √× discharge for some stations	× Ground water & reservoir levels √× discharge for some stations	×Lake reservoir levels √× discharge for Ruambwa and missing years for other stations	√× discharge for some years	√× discharge for some years	× Ground water & reservoir levels √×discharge for some stations	√ ×discharge for some years	√× discharge for some stations & years
	Catchment characteristics								
Size (Sq.km)	30,000	96,000	46,600	12,700	3.200	3,400	28,000	5,100	13,750
Human influence	√Irrigation, dams	√Irrigation, dams	√Irrigation, dams	×√ Irrigation	×√ Irrigation	×√ Irrigation	√Irrigation, dams	×√ Irrigation	×√ Ground water abstraction
Soil type	Sandy/loamy	Sandy/loamy	Sandy/loamy	Loamy/clay	Loamy/Clay	Loamy clay	Sand/loamy	Loamy/clay	Sand/loamy
Vegetation cover	Bare land/Sparse vegetation	Bare land/Sparse vegetation	Bare land/Sparse vegetation	Grassland/Trees /Croplands	Grassland/Trees / Croplands	Grassland/Trees /Croplands	Bare land/Sparse vegetation	Grassland/Trees /Croplands	Bare land/Sparse vegetation
	Some Models Calibrated over these catchments								
Model applied to the catchment	√×	×	×	√	×	√	√	√	√

√: Present ×: Not Present ×√: Partially Present

4.6.1. Application of decision tree to Kenyan catchments

This section assesses the suitability of hydrological models with reference to flood applications in Kenya, considering the aspects described. A decision tree (flow diagram) representing the filter sequence in the selection criteria is presented in Figure 4.9. At the top of the decision tree are all the processes that are deemed important in a model for effective flood applications in Kenya. Firstly, Kenya has major differences in terms of climates: some areas are Arid and Semi—Arid (ASALs), for example in the Eastern and North-eastern parts, whereas others are wetlands (e.g., Western and Central Highlands) (see Figure 4.2). Therefore, distinctions are made in the second step to accommodate processes that are important to the different climatic zones. Secondly, Kenya is currently facing data scarcity due to the ungauged nature of many catchments. This, however, should not be a setback to hydrological studies and as a result we filter the model based on the input data availability and ability to use alternative data. In the third step, we explore the availability of the model code to a wider user community. Here the concepts of code executability and online updating, accessibility and computational run time are explored. At the fourth stage, the model's ability to be downscaled to catchment local scale is considered. Fixed grid sizes and limitations of applicability to certain basin sizes are mainly considered here. Finally, we explore the preferences of the model based on their ease of implementation in the forecasting system environment. However, this piece of work does not involve the actual analysis of the models under consideration, and it is based on the elimination method, following previous studies on the performances of the models over the region. As a result, we present a yes/no decision tree which has potential implicit weighting factors of '0' or '1' based on whether the model meets a certain criterion or not from the MCA perspective.

The above aspects of the selection framework form the basis of this model overview and the discussion of the selection process. In this study, a combination of conceptual and process-based lumped and distributed hydrological models is considered for further evaluation to establish if they are compatible with the above aspects. The hydrological model should be suitable for evaluation of the spatial and temporal occurrence of floods based on a defined indicator. Therefore, the models considered (and described in the supplementary material) range from the few applied to or under consideration for the Kenyan context as well as the other widely used models employed in FF studies across the African continent that in our opinion would be applicable to the Kenyan case. A total of 12 rainfall-runoff models were initially listed as potential candidates for small scale operational flood forecasting (see Table 8 for main references). LISFLOOD and HYPE are included in this review, despite being developed for large-scale applications, because they were adopted as the prototypes in the AFFS and West Africa, respectively. The models were chosen mainly because of the existing literature reviews and application studies, particularly with reference to Africa and Kenya. We evaluated the models according to the criteria (aspects) discussed in Section 3 and outlined on the decision tree. We have applied this to all the 12 reviewed models and outline the summary of the evaluation in Table 9. This provides the summary statistics for each of the models based on process representation, data input requirements, model code availability, ease of downscaling to Kenyan catchments, and its application to operational flood forecasting.

97
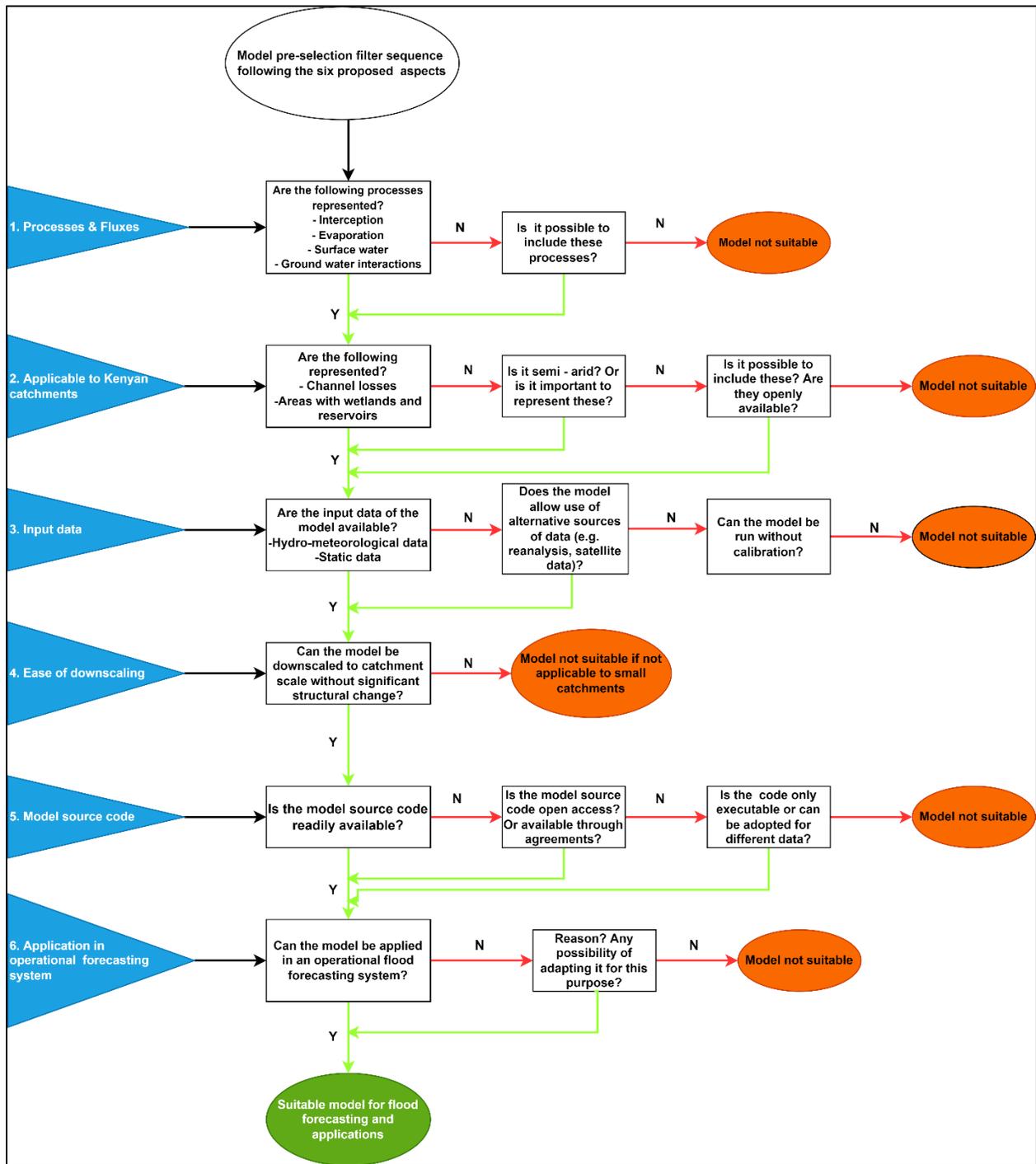
Figure 4.9:- Flow sequence to serve as a decision tree for evaluating and selecting a suitable hydrological model for flood forecasting in Kenya, based on the proposed criteria.

4.6.2. Actual model selection based on the evaluation

The Venn diagram (Figure 4.10) presents model selection following a comprehensive evaluation carried out in Table 9. All the models under consideration are described and summarised in Section

2.4. Following the filter sequence presented in Figure 4.9, each model is evaluated step by step, then potential models are summarised in the actual selection presented in Figure 4.10. Out of the twelve models, only VIC and TOPMODEL do not represent important processes for flood generation unique to Kenyan catchments. VIC and TOPmodel were eliminated because they could not represent groundwater processes and required the calibration of all the parameters, which in turn meant that the calibration data must be available, which is hardly the case in most of the Kenyan catchments. As a result, they were excluded from the final selection presented in the Venn diagram.

From the 12 models reviewed, five are considered suitable candidates for flood applications in Kenya (Figure 10). The outermost circle (A) presents the 10 models under consideration excluding VIC and TOPMODEL. [VIC and TOPMODEL were not included at this point because of their lack of representation of groundwater processes and other important factors (see Table 9). In addition, this category includes all the models which can be applied to the study catchments due to reasonable data input requirements, code availability, ease of downscaling to Kenyan catchments in drylands, semi-arid areas and wetlands, and application to operational flood forecasting.

Table 8:- 12 rainfall-runoff models listed as potential candidates for small scale flood applications with their main technical references.

Model	Main references
GR4J (modele du G ` enie Rural ´ a` 4 parametres au pas de temps Journalier	Technical: (Perrin *et al.,* 2003)
NAM (Nedbør-Afstrømnings-Model)	Technical: (Nielsen & Hansen, 1973)
SMAR (Soil Moisture Accounting and Routing)	Technical: O'Connor 2005;
PDM (Probability Distribution Model)	Technical: (Goswami & O'Connor 2010; Moore, 2007)
SWAT (Soil Water Assessment Tool)	Technical: (Arnold *et al*., 1998; Neitsch *et al.,* 2005)
MIKE SHE (MIKE Système Hydrologique Européen	Technical: (AbbottM B *et al.,* 1986; Ma *et al.,* 2016)
HBV-96 (Hydrologiska Byråns Vattenbalansavdelning)	Technical: (Lindström *et al.,* 1997)
TOPMODEL (TOPography based hydrological)	Technical (Beven and Kirby 1979; *Beven et al.,* 1984)
GeoSFM (Geospatial Streamflow Simulation Model)	(Artan. *et al.,* 2004; 2001; Asante *et al.,* 2008)
VIC (Variable Infiltration Capacity)	Technical: (Gao *et al.,* 2010; Lohmann *et al.,* 1996)
LISFLOOD	Technical: (Burek, 2013; van der Knijff *et al.,* 2010)
HYPE (European Hydrological Predictions for the Environment	Technical: (Lindström*et al.,*2010) http://hypecode.smhi.se

Table 9: Evaluation of the 12 models based on the selection criteria.

Model Name/ criteria	VIC	SWAT	GeoSFM	HBV-96	MIKE-SHE	TOPM ODEL	PDM	SMAR	NAM	HYPE	GR4J	LISFL OOD
Represented processes and fluxes												
Interception	√ f(LAI)	√ f(LAI)	√f(forest/land	×√ (Modified)	√	×√(Modified)	√ f(canopy)	√ f(LAI)	√ f(LAI)	√ f(LAI)	√ f(LAI)	√ f(LAI)
Evaporation	√ Penman	√ P-M/P-T/	√ Penman-Monteith	√	√	√	√	√	√	√	√	√input
Snow	√ Energy	√Degree day	√ Degree day	√ Degree day	√	√Degree day	√Degree day	√ Degree	√ Degree day	√ Degree day	√Degre e day	√Degre e day
Soil	√ 2 or 3 layers	√ ≤ 10 layers	√ 2 layers	√ 2 layers	√ 3 layers	√ 2layers	√ 2layers	√ 2layers	√ 2layers	√ 2 layers	√ 2layers	√ 2layers
Ground water	×	√ Shallow	√ Shallow & deep		√ Shallow	√ subsurfa	√ subsurface	√	√Shallow & deep	√ subsurface	√	√ 2 parallel
Lake, reservoir	×	√ Linear res.	√ Linear res.	√	√	√	√ Linear res.	√	√	√ Linear res	√	√linear reservoi
Runoff	√ saturati	√ (SCS-CN)	√ (SCS-CN)	√ saturation	√	√ infiltrati	√	√	√	√	√	√ infiltrati
Routing	√ Linear transfer	√Muskin gum-	√Muskingu m- Cunge/	√Muskin gum-	√	√	√ cubic non-linear	√	√	√	√	√Kinem atic
Calibration	×√ Several	√ Several	×	√ Several	√	√ Several	√ Several	√9 paramet	×√9 parameters	√ Several	√4 paramet	×
Energy balance	√	×	√	√	√	√	×		×	×	×	×
Water use	×	√	√	√	×	√	√		×	√	×	×
Data requirement and resolution of the model												
Input Meteorol ogical data	Daily or sub-daily precipit ation, air tempera	Daily precipita tion, minimu m and maximu m	Daily precipitation , potential evaporation	Daily precipitat ion, temperatu re & estimates of	precipit ation, air tempera ture and solar	Precipita tion	Daily rainfall and potential evapotrans piration	Daily rainfall and Temper ature	Daily rainfall, potential evapotrans piration and	Daily precipitatio n, estimates of potentialvap oration	Daily precipit ation, estimate s of potentia l	Daily rainfall, potentia l vaporati on & daily

	ture &wind speed	temperature		potential evaporation	radiation			temperature		vaporation	mean air temperature
Model spatial resolution	0.5°	Subbasins	Semidistributed	Semidistributed	Subgrids	Distributed	0.5°	Lumped	Lumped	Sub-basins	Lumped	100m and larger
Model temporal resolution	Daily	Daily	Daily	Daily	Daily	Hourly/daily	Hourly/Daily	Daily	Daily	Daily	Daily	Hourly/Daily
Model Code availability												
Open source	×	√	×	√ Executable	×	√ (as an R code called topmodel	√	√	√	√open source	√ R-package called airGR	×
Only Executable			√		√							√
Model applicability to Kenyan catchments												
Geographically	√	√	√	√	√	×√ sensitive to grid size (≤50 recommended	√	√	√	√	√	√

Climatic conditions	× in semi-arid catchments √ in humid catchments	√	× in semi-arid catchments √ in humid catchments	× in semi-arid catchments √ in humid catchments	× in semi-arid catchments √ in humid catchments	√in semi-arid catchments √ in humid catchments	× in semi-arid catchments √ in humid catchments	√in semi-arid catchments √ in humid catchments	√in semi-arid catchments √ in humid catchments	√in semi-arid catchments √ in humid catchments	√in semi-arid catchments √ in humid catchments	√in semi-arid catchments √ in humid catchments
The ease of downscaling model to river basin scale												
Ease of downscaling without model structure modification	√	√	√	√	√	×√	√	√	√	√ can be run on subgrid	√	× √
Models have been calibrated over some Kenyan catchments and applied to operational FF												
Model applied to Kenyan catchment	×	× √	× √	×	× √	×	×	√	× √	× √	× √	×

√: Considered ×: Not considered ×√: Partially considere
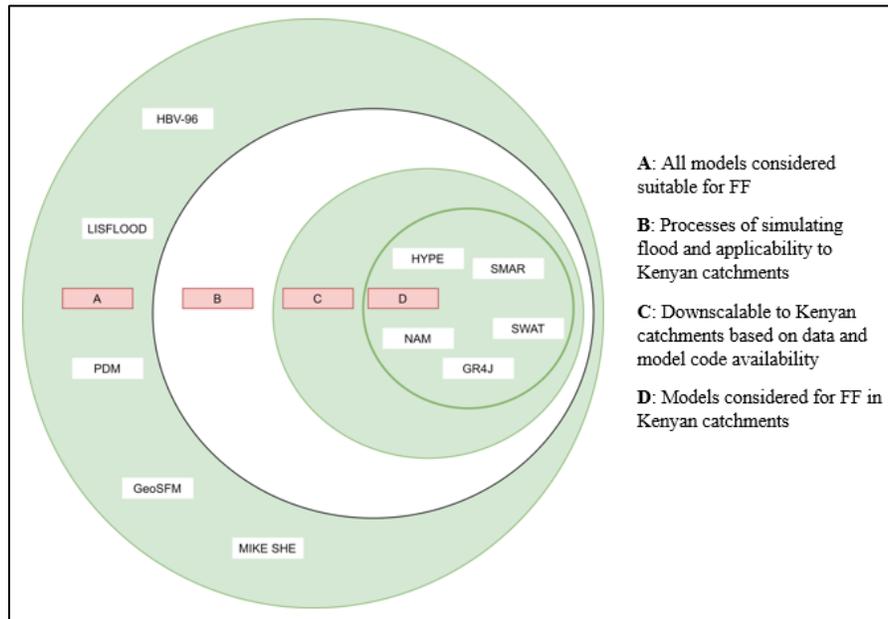
Figure 4.10:- Venn diagram following the model selection procedure, starting with all the models under consideration in circle A resulting with the selected models in innermost circle D.

Circle B represents model selection based on data input requirements and the number of calibrated parameters. At this stage, we eliminate LISFLOOD, HBV-96, PDM, GeoSFM and MIKE SHE. LISFLOOD, MIKE SHE, and HBV 96 and GeoSFM are fully and semi-distributed models respectively, with very many parameters to be calibrated (van der Knijff *et al*., 2010; Ma *et al*., 2016; Berglöv *et al*., 2009). In addition, they are run on hourly timesteps with very many data input requirements. The calibration of many parameters will also require intensive computer run time which may be a challenge in many NMHS (Vema and Sudheer, 2020). The ungauged nature of most of the operational centres in Kenya means they may not have reliable data at high frequency (e.g., at hourly or even daily timesteps). However, circle B is a white area because there is the option of alternative remotely sensed data. For these models with high data requirements in data-scarce areas, there are alternative sources of satellite and reanalysis datasets that are effectively utilized to reinforce the model, but they must be regarded with caution. This is because the datasets come with their own uncertainties, including random and systematic errors (Fortin *et al*., 2015; Sun *et al*., 2018). Inherent input uncertainties will affect the performance of models for a given catchment as a result, we eliminated LISFLOOD, HBV-96 and MIKE SHE at this stage. PDM is also eliminated at this point because the model's configuration comprises probability-distributed soil moisture storage, surface storage, and groundwater storage components (Moore, 2007). The latter is hardly available input as Kenya's NHMS has no data on reservoirs and ground water storage.

Circle C represents models whose code is easily available as a free open source. This category is meant to rule out models whose codes are available, but only in executable format, as changes, for instance reading of input data, may be necessary and are not provided for in executable model

codes. The candidate models filtered through to this step, HYPE, SWAT and SMAR, have freely available open-source codes (Paul *et al*., 2020). GR4J and NAM source codes are available through open collaborations (Humphrey *et al*., 2016). The innermost green circle represents models that can be applied easily to Kenyan catchments through simple downscaling and are suitable for flood forecasting in different Kenyan catchments. Regarding the last criterion, as to whether the model is suited for operational purposes, all models reviewed are continuous simulation models and no model is rejected at this step because we assume that, if necessary, they can be modified to be suitable for use in an operational environment.

4.7.   Discussion

We provide an insight into the need for comprehension of the quantitative or qualitative description of existing knowledge and understanding of the catchments and how this would influence the choice of the modelling tools at catchment scale, acknowledging the gaps and challenges. Models used for different applications in different parts of the world are reviewed according to six six crucial aspects, which builds on the previous works of Kauffeldt *et al*. (2016) and Trambaur *et al*. (2013), with the aim of assessing their suitability for flood applications in Kenya. These two foundational works provide a technical review of large-scale hydrological models for implementation in operational flood forecasting, highlighting their suitability for drought forecasting at continental level, specifically in sub-Saharan Africa. They are important and provide a comprehensive model review and a selection framework for flood and drought application at continental scales respectively. However, these studies are applied at a larger scale (continental), yet models simulate process differently in different hydro-climatic conditions, revealing the need to link the process at catchment scale to model specifications and applications.

It can be noted that not all models are good at capturing and/or representing the important processes relevant to flood generation (e.g., as transmission losses along the river channel, re-infiltration, and subsequent evaporation of surface) both in wetland and ASALs of Kenya as summarised in Table 7. It should be noted that, with the current data scarcity, most modelling frameworks incorporate satellite and reanalysis data. These products have a coarse resolution (with the exceptions of some newly high-resolution satellite products e.g., CHIRPS, MSWEP) and high uncertainty in their estimations at catchment scale, which in turn impacts the model's performance. Thus, the way forward for objective choice of modelling tools should ensure that the models are stable, have reliable error and inconsistency checks, are able to flag missing data errors (e.g., when input sources fail), fit into an operational environment and should preferably be user friendly. Considering the data scarcity issues, most models can be implemented as the redundancy related to missing data can be incorporated in the pre-processing. Therefore, if a model can run with missing data, it is a requirement that the run is clearly flagged as having missing data. Model stability can be tested by looking at the distributions of parameters where they became remarkably well-behaved and near elliptic when numerical error control is implemented in the model (Kavetsk *et al*., 2005). However, since the properties of parameter distributions are dependent on (i) the data, (ii) the model and (iii) the objective function, testing model stability before application may not be achieved. Model errors can be established by running a sensitivity and uncertainty analysis

of model parameters, which should be reliable (Song *et al*., 2015), but this requires more computational power, which is unavailable in Kenya.

The practical proposed and presented model pre-selection with a filter sequence for flood applications was used to filter out models to a subset considered suitable for Kenyan catchment types. Through the sequence presented, possible adaptation assumptions are considered in some cases. The filter sequence criteria to assess model suitability included the representation of important processes, availability of the model code, existing user community, input data requirements, the possibility of calibration, model resolution and data assimilation with operational implementation into a flood forecasting system. Out of the 12 models, only 5, SWAT, SMAR, GR4J, NAM, HYPE, were considered suitable candidates for catchment scale flood forecasting by local authorities in Kenya. The above pre-selection process suffices as a milestone to addressing some at least, if not all the challenges associated with choice of a modelling tool to the end-users of to be used effectively both at catchment scale modelling and potentially adopted in an operational early warning system to help mitigate the effects of floods in data-scarce regions such as Kenya.

This work does not look at direct analysis of each of the proposed models to evaluate its performance based on some past events. As a starting point, this work provides a background of hydrological models and the Kenyan context to provide criterion.] for model pre-section for flood applications at national level. The modellers and users of the models can then use the information and arrive at models to apply to some selected events. A Multi Criteria Analysis (Sherlock and Duffy, 2019) forms the basis of this initial step. The whole process of an MCA is designed to assess multiple alternatives based on a mix of quantitative and mostly qualitative information from multiple sources. However, the proposed MCA relies heavily on evaluation data, and is very time consuming for the number of models available; consequently, for data-scarce regions, and/or agencies with limited resources, and in general, an additional decision tree is helpful to trim down the number of options. It is necessary to evaluate further the limited selection with, for example, an MCA and the FFC experiment. This is mainly because, within the same catchment, inhomogeneities of the physical and hydroclimatic properties give rise to a complex issue that it is essential to consider when deciding which model to use: this demonstrates the importance of the selection criteria.

4.8.   Concluding remarks

There are some challenges that are inherent, when applying the above decision framework, not only to data-scarce regions but also to a wider global scale. For example, with the advancement in research, there is an increasing number of models and none of them is error free, mainly due to a compromise reached when considering model complexity and computational run time, which is a major challenge (Melsen, *et al.,* 2019). Also, it is difficult to balance complexity of model structure with   parameterization and input data requirements because complex models do not guarantee reliable results (Trambauer *et a*l., 2013; Paul *et al*., 2020). The use of certain models depends on the computational capabilities (skills) of the individuals as well as the NMHS in general. As a result, model selection may be biased towards ease of applications, depending on the skills of the modeller. In addition, there is no documented research, outlining the pros and cons of each of

models in a single platform, which a potential model user might easily use to identify which model is suitable (Mannschatz and Hülsmann, 2016).

To address the highlighted challenges confronting modelling communities in developing countries, Paul *et al*. (2020) and Souffront Alcantara *et al*. (2019) suggest some of the possible ways forward. For example, developing countries should consider working on developing their own models. The current models are tailored to catchment scale or specific geographical locations, developed with rivers flowing nicely all year round in relatively wet catchments, and the inclusion of a variety of hydrologists and model developers with different needs and perspectives is most welcome and necessary to produce hydrological models for a wider range of environments. This may take a long time, due to inadequate technological capacities, but will suffice as a milestone on the way to addressing some of the challenges associated with model selection. A well-prepared and comprehensive database platform with useful information pooled together, such as different input information, and the advantages and disadvantages of different models, is important in providing initial information, enabling the formation of a judgement by eye as to which model would work best. This is also likely to facilitate easy model selection alongside frequent webinars by model developers to enhance the skill of modellers in developing countries.

This research provides initial steps to inform the choice of modelling tools in data scarce regions. There is a need for further analysis of the proposed models' applicability to Kenyan catchments, to assess their abilities to simulate past events. This will provide additional and useful information on the choice and application of these models at catchment scale to areas with varied hydro-climatic characteristics. We acknowledge that it has not been proven that the criteria 'suffice' as the selection procedure because they lead to multiple models; furthermore, no follow-up strategy is presented here: these matters form the basis of future work. Additionally, the steps of the filtering process are not operationalized to the level where it can be said to be objective. For example, a model may be excluded on consideration of 'many parameters' because the pre-selection criteria presented here follow a flow chart. First, we make a pre-selection based on expert judgement and links to models that have been applied to various diverse environments, that are deemed suitable candidates for the Kenyan context.

4.9.   Acknowledgements

Chapter 5

5. Assessment of global reanalysis precipitation for hydrological modelling in data scarce regions: a case study of Kenya

5.1. Objective addressed and publication details

This chapter assesses the performance of four reanalysis datasets (ERA5, ERA-Interim, CFSR and JRA55) over Kenya for the period 1981–2016 on daily, monthly, seasonal, and annual timescales. Firstly, evaluation of the reanalysis datasets by comparing them with observations from the Climate Hazards group Infrared Precipitation with Station is explored. Secondly, evaluation of the ability of these reanalysis datasets to simulate streamflow using GR4J model, considering both model performance and parameters' sensitivity and identifiability, is presented. This work is important because it informs future applications of reanalysis products for setting up hydrological models that can be used for flood forecasting, early warning, and early action in data-scarce regions, such as Kenya.

This work has been reviewed and published in the *Journal of Hydrology- Regional Studies, Volume 41, June 2022, 101105* and can be accessed at https://doi.org/10.1016/j.ejrh.2022.101105. The exact publication copy is in Appendix **A3** material of this thesis.

5.2. Authors' contributions

Maureen A. Wanzala (60%): - Conceptualization, Data curation- Lead, Methodology- Lead, Investigation, Software, Formal analysis - Lead, Visualization, Writing – original draft, Writing - Review & Editing- Lead. Andrea Ficchì (10%): Conceptualization, Supervision, Data curation- Support, Methodology- Support, Investigation, Software, Analysis - Support, Visualization, Writing - Review & Editing- Support. Hannah L. Cloke (10%): Supervision, Conceptualization, Writing - Review & Editing, Project administration. Elisabeth M. Stephens (10%): Conceptualization, Writing - Review & Editing- Support, Project administration, Supervision, Funding acquisition. Heou M. Badjana (5%): Methodology and data curation- Support Writing – Review & Editing-Support. David A. Lavers (5%): Data curation, Investigation

5.3. Introduction

Precipitation is arguably the most important driver of catchment hydrological response (e.g., MacLeod *et al*., 2021), but it is challenging to get accurate information on the amount, duration, and intensity of rainfall events (Beck *et al*., 2017; Tapiador *et al*., 2012), due to the high spatio-temporal variability (Vischel *et al.*, 2011; Nicholson *et al.*, 2019). This is compounded by a low spatial coverage and a net decline in the number of ground gauge stations in the historical climatological observation network, especially in developing countries such as Kenya (Zaitchik *et al.*, 2011; Menne *et al.*, 2018; Tarek, Brissette and Arsenault, 2020, 2021). Unreliable or incomplete datasets are unable to correctly identify seasonal or short-range temporal patterns (Gosset *et al*., 2013; Le *et al*., 2017).

Other sources of precipitation data such as those from satellite remote sensing are now available, but they come with their own errors, including random and systematic (see Beck *et al*., 2021; Beck *et al*., 2017a; Beck *et al*., 2017b; Fortin *et al*., 2015; Sun *et al*., 2018). Another freely available source of precipitation data consists of meteorological reanalysis products, which are becoming increasingly promising due to upgrades in their spatial resolution and improved representation of atmospheric processes in global models (Hersbach, 2018). Reanalysis data combine a wide range of remotely sensed observations with a dynamical–physical coupled numerical model to produce the best estimate of the state of the atmosphere. Reanalysis is not reliant on the density of surface observational networks and can give surface variables in locations with little to no surface coverage. As a result, they can generate several variables both at the land surface and on vertical atmospheric levels, and hence have been applied in several studies both for climatological and hydrological purposes across the world (e.g., Beck, *et al*., 2017a; Chen *et al*., 2018; Emerton *et al*., 2017; Essou *et al*., 2017). Several different reanalysis products exist but they are known to vary in quality with recurrent upgrades. It is important to evaluate them carefully both to inform the users and the developers of the datasets. The developers of these products can work on improving their updates only when there is a complete feedback loop between applications and developments Therefore, ground validation of reanalysis precipitation is very important but very challenging, particularly where the rain gauge networks are sparse.

Several studies attempt to quantify and account for the sampling errors, comparing reanalysis data with observations in different parts of the world (*e.g*., Guo, 2018; Tang *et al*., 2020; Xu *et al*., 2020; Zaitchik *et a*l., 2011), at a global scale (Beck, et al., 2017a; 2017b), at regional or basin scale (e.g., Acharya *et al*., 2019; Nkiaka *et al*., 2017; Tarek *et al*., 2020) and at a national scale (e.g. Arshad *et al*., 2021; Gleixner *et a*l., 2020; Koukoula *et al*., 2020; Lakew *et al.*, 2020; Shayeghi *et al*., 2020; Tafaseye *et al*., 2017). However, the findings of these studies were mixed. Differences in approaches, regions, and time scales resulted in inconsistency in product performance, implying that site-specific performance evaluation may be required. Existing studies also aimed at analysing a single product or a few products for short periods of time, so their estimated errors may not reflect long-term behaviour.

Additionally, the temporal dynamics of rainfall are very important as they play a crucial role in the total accumulated rainfall on daily and monthly timescales, thus influencing the bimodal seasonality observed over Kenya. The highly-variable temporal dynamics are also key in explaining the nonlinear nature of infiltration processes (Blöschl & Sivapalan, 1995), such as the peak discharge value (Gabellani *et al.*, 2007) and runoff volume (Viglione *et al.*, 2010) in hydrological modelling. Thus, the above observations highlight the need to consider different temporal scales, when evaluating the reanalysis precipitation relative to observation.

In Kenya, there were 20 major floods from 1964 – 2020 which were driven by precipitation falling in the seasonal rains. More than 160,000 people were displaced countrywide by floods in October 2019 (ReliefWeb, 2019; UNICEF, 2019; Opere, 2013). Annual average economic loss from flooding is estimated to be 5.5% of gross domestic product (Njogu, 2021). Therefore, understanding the best representation of precipitation in flood models which can be used for forecasting or risk analysis is of great societal importance. Kenya has a widely varying physical

geography resulting in great variability of river catchment characteristics across the country. Thus, it is essential not only to understand the representation of precipitation at a country scale, but also on a catchment-by-catchment basis (Golian and Murphy, 2021; Meresa *et al.*, 2021). Previous evaluations of reanalysis products in capturing Kenyan rainfall show varied levels of agreement in spatio-temporal variability relative to observations (e.g., Alemayehu *et al.*, 2018; Dile and Srinivasan, 2014; Gleixner *et al.*, 2020; Khan *et al.*, 2011). Moreover, studies employing hydrological modelling generally used discharge observations from a small number of catchments (e.g., Alemayehu *et al.*, 2018; Bitew *et al.*, 2012; Langat *et al.*, 2017; Le *et al.*, 2017; Worqlul *et al.*, 2017) and did not quantify uncertainties associated with each reanalysis (e.g., Alemayehu *et al.*, 2018), leading to combined rainfall and model uncertainty that is not easily interpreted. Hence, there is a notable gap in the literature associated with evaluating the accuracy of multiple reanalysis products across different catchments, accounting for both model and input errors, especially in data scarce regions like Kenya, and this gap was an important motivation for the present study. This paper evaluates four reanalysis precipitation products with respect to observations and assesses their suitability for use in hydrological modelling in 19 Kenyan catchments. We assess their performance in reproducing the most important features of rainfall events and regimes, and in simulating catchment streamflow, through answering the following research questions:

1. How well do the precipitation datasets compare in terms of temporal dynamics at the basin scale? Which product is the most accurate compared to observations?

2. How well do precipitation datasets compare in terms of spatial patterns? Which product shows consistency in spatial heterogeneity compared to observations?

3. How does the general hydrological model performance vary with different datasets?

4. How does the sensitivity of a rainfall runoff model (GR4J) vary with alternative rainfall forcing?

We consider both model performance and parameter uncertainty and compute a Model Suitability Index (MSI) by coupling the results of model performance statistics with Global Sensitivity Analysis. We compare four reanalysis datasets using the GR4J model across 19 Kenyan catchments with varied climate and morphological characteristics, to investigate which input data are suitable or require caution, when used in the place of observation datasets in different regions. This work is a stepping stone and an essential guide for hydrological applications of global reanalysis datasets because it compares several reanalysis products to observations on daily, monthly, and seasonal scales, and unveils the propagation of uncertainty from different reanalysis when used as model inputs. All the above reviewed studies looked at the performance of at most one reanalysis dataset in simulating streamflow and only over one catchment; none, however, looked at such a country-scale performance. To our knowledge, this is the very first evaluation of the different reanalysis products over Kenya for simulating streamflow coupled with sensitivity analysis.

5.4. Study area and catchment characteristics

The study is undertaken in 19 Kenyan catchments (Figure 5.1) with varying characteristics (Table 10). These were selected due to the frequency and magnitude of the impacts of floods, as well as the availability of river flow observations (Table 1). Kenya mainly experiences a bimodal rainfall pattern, occurring in the seasons of March - April - May (MAM) and October -November - December (OND) (Ayugi et al., 2016; Yang et al., 2015), which are commonly known as the 'long' and 'short' rains respectively. The rainfall seasonality and the migration of the precipitation zone is mainly influenced by the north-south movement of the inter-tropical convergence zone (ITCZ) (Black *et al*., 2003; Ongoma *et al*., 2015). The rainfall season migrates northward at a slower rate than it migrates southward, hence the two different names – 'long rains' and 'short rains' respectively (Nyenzi, 1988). The rainfall exhibits high spatiotemporal and interannual variability (Ongoma and Chen, 2017) and is strongly influenced by perturbations in global Sea Surface Temperatures (SSTs) especially in the Pacific and Indian Oceans with the El-Niño Southern Oscillation (ENSO) (Ogallo, 1993; Black, Slingo and Sperber, 2003) and the Indian Ocean Dipole (IOD) (Blau *et al*., 2020; Owiti *et al*., 2008) being the most important modes. Other systems that influence rainfall variability include the high pressure systems (e.g., the Mascarene and the Arabian) (Ogwang *et al.*, 2015), the Quasi-Biennial Oscillation (QBO) (Collier *et al*., 2016; Indenje and Semazzi, 2000), the Madden-Julian Oscillation (MJO) (Kilavi *et al.*, 2018), Tropical cyclones (Finney *et al.*, 2020; Wainwright *et al.*, 2021) and jet streams, e.g., the Turkana jet (Kinuthia, 1992; Hartman, 2018). The country has complex topography with the lowest altitudes along the coastline and Lake Victoria basin which are particularly prone to floods, while in the highlands, frequent thunderstorms and lightning threaten life.

5.5. Data and methodology

5.5.1. Datasets

5.5.1.1. Reanalysis and observational data

Four reanalysis products, namely ERA5, ERA-Interim (hereafter ERAI), Climate Forecast System Reanalysis (CFSR), and the Japanese 55-year Reanalysis (JRA55), and a gridded observational dataset, the Climate Hazards group Infrared Precipitation with Station (CHIRPS), were used in this study (see Table 11). We used the daily precipitation and maximum and minimum temperature variables from the reanalysis products for the study.
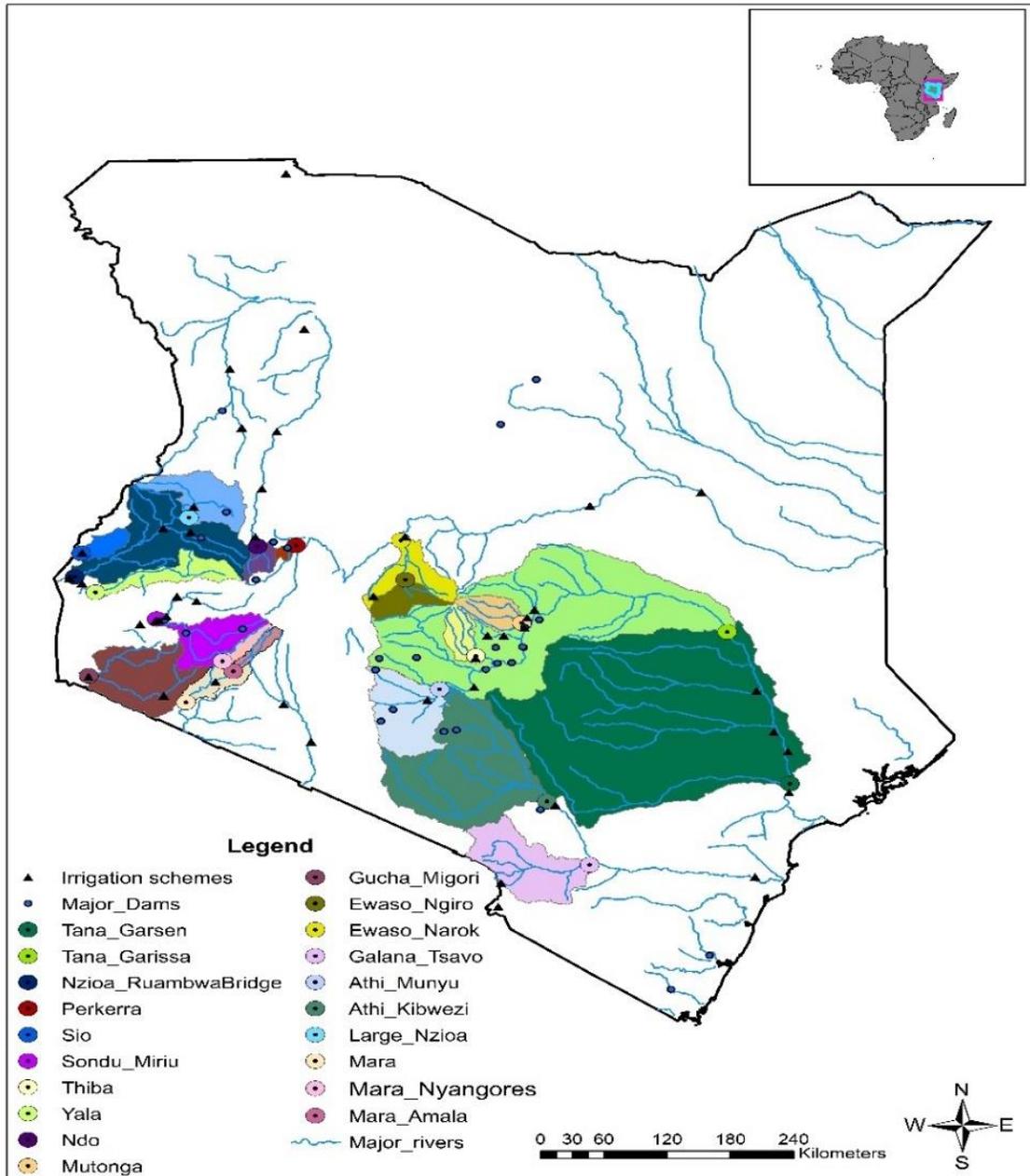
Figure 5.1:- Study catchments, with the location of the outlet river gauges (as shown by circled dots) used in this study and the main irrigation schemes and major dams (Source: WRA -K) across Kenya.

Table 10:- Summary of the catchments considered, their characteristics, and the main human influences, including number of dams and water abstraction activities (Source: WRA).

River Name	Catchment Outlet point	Station ID	Lon	Lat	Drainage Area (km²)	Mean Elevation (m.a.s.l)	Mean Annual Rainfall (mm)	Annual Discharge (m³s⁻¹)	Catchment Characteristics	Human Influence		First & Last year of record	Record length (years)	Amount missing (%)
										Dams	Irrigation schemes			
Tana	Tana Garsen	4G02	40.11	-2.28	80 760	720	672	135.8	Semi-arid plains	9	11	1981-2016	36	58.2
Tana	Tana Garissa	4G01	39.7	-0.45	32 695	870	868	169.3	Highlands on the upstream & semi-arid plains in the lowlands	8	7	1981-2018	38	14.2
Nzioa	RuambwaBridge	1EF01	34.09	0.12	12 643	1740	1488	151.2	Dense forest cover (highlands) & low trees & bushes (lower reaches)	2	4	1981-2018	38	13.6
Galana	Galana Tsavo	3G02	38.47	-2.99	6560	930	628	3.3	Semi-arid savannah plains	3	1	1981-2015	35	59.6
Gucha	Gucha Migori	1KB05	34.21	-0.95	6 310	1650	1435	45.0	Eastern lowlands with dense vegetation cover	0	2	1981-2015	35	47.8
Athi	Athi Munyu	3DA02	37.19	-1.09	5689	1730	822	18.8	Highlands and forest cover	3	1	1981-2017	37	21.6
Nzioa	Large Nzioa	1BD02	35.06	0.76	3878	1720	1267	15.3	Dense forest cover	1	1	1981-2011	31	28.8
Sondu	Sondu Miriu	1JG04	34.80	-0.33	3444	2017	1614	53.9	Low lying plains (Western) & highlands (Eastern)	2	2	1981-2018	38	64.4
Mara	Mara	1LA04	35.04	-1.23	2977	2100	1262	11.8	Low lying shrubs, semi-arid	0	1	1981-2015	35	77.7

Yala	Yala	1FG02	34.27	0.04	2700		1696	40.8	Swampy	0	0	1981-2019	39	59.6
Ewaso	Ewaso Narok	5AC10	36.73	0.43	2597	1600	880	5.3	Low lying shrubs & mainly semi-arid	0	2	1981-2018	38	26.5
Tana	Mutonga	4EA07	37.89	-0.38	1867	1830	1427	35.5	Highlands and forest cover	0	1	1981-2016	36	44.2
Ewaso	Ewaso Ngiro	5BC04	36.91	0.09	1837	1700	972	20.6	Low lying shrubs & mainly semi-arid	0	0	1981-2019	39	35.0
Sio	Sio	2EE07A	34.14	0.39	1011	1650	1822	15.5	Low trees & bushes & swampy in lower reaches	0	1	1981-2018	38	18.1
Turkwel	Ndo	2C07	35.65	0.45	897	1133	1371	9.1	Extensive palaeo-floodplain & arid conditions	0	1	1981-1993	13	47.2
Mara	Amala	1LB02	35.44	-0.89	695	2100	1377	6.8	Low lying shrubs, semi-arid	0	0	1981-2017	37	25.6
Mara	Nyangores	1LA03	35.35	-0.79	692	2008	1262	11.8	Semi-arid savannah plains, low lying shrubs, semi-arid	0	0	1981-2017	37	15.5
Turkwel	Perkerra	2EE07A	35.97	0.46	371	1023	832	5.7	Extensive palaeo-floodplain and arid conditions	1	1	1985-2005	21	50.1

ERA5 is the latest global atmospheric reanalysis product from the European Centre for Medium-Range Weather Forecasts (ECMWF), which spans the modern observing period from 1950 onward (Hersbach, 2018). In this study, 3-hourly ERA5 was obtained from ECMWF on a fixed grid of $0.31° \times 0.31°$. ERAI is the previous global reanalysis product created by ECMWF (Dee *et al.*, 2011). Daily ERAI was obtained from ECMWF on a fixed grid of $0.75° \times 0.75°$. JRA55 is a global reanalysis dataset constructed by the Japan Meteorological Agency (JMA) (Kobayashi et al., 2015). Daily JRA55 was obtained from the National Center for Atmospheric Research (NCAR) climate data guide at a fixed gird of $0.56° \times 0.56°$. CFSR is a global reanalysis dataset of atmosphere fields produced by the National Centers for Environmental Prediction and for Atmospheric Research (NCEP/NCAR) (Saha *et al.*, 2010). The CHIRPS dataset was used as a benchmark observation dataset since it has been used in several studies showing good results compared to observations over eastern Africa (Dinku *et al.*, 2018). CHIRPS is a quasi-global, high resolution, daily, pentad, and monthly precipitation dataset (Funk *et al.*, 2015). Based on infrared Cold Cloud Duration (CCD) data, CHIRPS has a long enough history of precipitation data. The algorithm (i) is based on a 5 km climatology that uses satellite data to represent sparsely gauged locations, ii) includes daily, pentadal, and monthly 5 km CCD-based precipitation estimates from 1981 to the present, iii) combines station data to generate tentative information product with a latency of about 2 days and a final product with an average latency of about 3 weeks, and iv) interpolation weights are assigned based on a novel blending method which uses the spatial correlation structure of CCD estimates. This makes it comparatively an alternative in data-scarce regions. We opted for the gridded observations as the daily observed gauge datasets were not available for the catchments of study and are known to be very sparse and present large data gaps (Dinku *et. al*., 2019; 2018; Le *et. al*., 2017).

5.5.1.2. Observed River Discharge and Potential Evapotranspiration

River discharge datasets at daily time steps for the period 1981- 2016 were provided by the Kenya Water Resource Authority (WRA) for the selected catchments across the country, summarized in (Table 10). The potential evapotranspiration (PET) required for the catchment modelling was estimated from the average daily temperature of the four reanalysis products and CHIRTS-daily data from the Climate Hazard Centre (CHC). As temperature readings were the readily available meteorological data relating to PET, for this study, temperature-based methods were used to estimate the PET (Hargreaves & Samani, 1985). For this study, the Hamon method (Hamon, 1960) was used to estimate PET daily averages for different datasets.

5.5.2. Modelling experiment methodology

To obtain the monthly and annual totals for observations and reanalysis datasets, the daily values were accumulated. The seasonal total precipitation was calculated by summing monthly precipitation for three seasons: (i) March-April-May, hereafter referred to as MAM, (ii) June-July-August, hereafter JJA, and (iii) October-November-December, hereafter OND. All datasets were converted to the same units for consistency (e.g., JRA55 and CFSR were converted from $kg/m^2/s$ to mm/d). ERA5, ERAI, JRA-55 and CFSR were regridded by first-order conservative interpolations to a horizontal grid of 0.5º x 0.5º (Schulzweida, 2019).

We first qualitatively evaluate the performance statistics of the reanalysis datasets in terms of temporal dynamics and biases with respect to precipitation observations (CHIRPS), considering the following metrics: Pearson Linear Correlation Coefficient (CC), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Error (ME), long-term relative bias (BIAS) and the annual number of dry days calculated on monthly, annual, and seasonal scales. We produce spatial maps for the standardized precipitation anomalies, bias and annual number of dry days, to assess their consistency compared to observations, and tabulate the other statistics to show the aggregate performance across the different datasets.

Secondly, we calibrate the GR4J (Perrin *et al.*, 2003) rainfall runoff model. In the study, we used five different sources of inputs (for both precipitation and PET) into the GR4J model: CHIRPS, ERA5, ERAI, CFSR and JRA55. We calibrate the model with each of the input sources at a time, compute the KGE score, and compare how this varies across the four different datasets relative to observations. The GR4J model is a simple daily lumped rainfall-runoff model belonging to the family of soil moisture accounting models. There are four main parameters (Figure 5.2) to be calibrated in the GR4J model, namely: (1) the maximum capacity of the production store (X1, mm), (2) the groundwater exchange coefficient (X2, mm), (3) the maximum capacity of the non-linear routing store (X3, mm), and (4) the time base of the unit hydrograph (X4, days). There are also a few fixed parameters, whose values were set by Perrin *et al.*, (2003)(

Table 5). All four free parameters are real numbers: *X1* and *X3* are positive, *X4* is greater than 0.5, and *X2* can be either positive, zero or negative. The typical inputs of GR4J are the areal precipitation depth (P, mm) and the potential evapotranspiration (PE, mm) estimate over the catchment. Most optimization algorithms used to calibrate the model parameters require knowledge of an initial parameter set. Given the small number of model parameters, simple optimization algorithms are generally capable of identifying parameter values yielding satisfactory results. The choice of an objective function depends on the objectives of the model user. The choice and use of the GR4J model is mainly due to its simple and relatively quick to calibrate structure, ensuring high levels of performance and robustness (Ficchì *et al.*, 2019; Mostafaie *et al.*, 2018; Oudin *et al.*, 2004; Van Esse *et al.*, 2013; Pushpalatha *et a*l., 2011).

The four free parameters of the GR4J model were calibrated using the default optimisation algorithm provided in the airGR package (Coron *et al.*, 2019; Delaigue *et al.*, 2019). This simple optimization algorithm, mainly based on a local optimisation, proved to be equally efficient in locating a robust optimum compared to more complex global search algorithms (Coron *et al.*, 2019) and proved effective in terms of the number of model runs required for convergence (Mathevet *et al.*, 2006). The Michel method (Michel, 1983) is based on two steps:

    I.   A systematic inspection of the global parameter space is performed to determine the most likely zone of convergence. In our study, this is done by direct grid-screening.

    II.   A steepest descent local search procedure is carried out to find an estimate of the optimum parameter set starting from the best parameter set from step 1.

Table 11: Overview of the global reanalysis and the blended (Satellite and observation) Chirps precipitation dataset(s) used in the study.

Short Name	Full Name and details	Data sources (*)	Spatial resolution	Spatial coverage	Temporal coverage	Temporal resolution	Reference
CHIRPS V2.0	Climate Hazards group Infrared Precipitation (CHIRP) V2.0 (http://chg.ucsb.edu/data/chirps/)	G, S, R [*]	0.05º	Land, <50	Daily	1981–NRT[*]	(Funk *et al.*, 2015)
ERA5	European Centre for Medium-Range Weather Forecasts Reanalysis (https://www.ecmwf.int/en/research/climate-reanalysis/era-5)	R [*]	~ 0.31º	Global	Hourly	1979 – NRT[*]	(Hersbach, 2018)
ERA-Interim	European Centre for Medium-Range Weather Forecasts ReAnalysis Interim (https://www.ecmwf.int/en/research/climate-reanalysis/era-interim)	R [*]	~ 0.75º	Global	3-Hourly	1979-2019	(Dee *et al.*, 2011)
JRA-55	Japanese 55-year Reanalysis (JRA-55) (https://jra. kishou.go.jp/JRA-55 ; https://data.diasjp.net/dl/storages/filelist/dataset:204)	R [*]	~ 0.56º	Global	3-Hourly	1951-NRT[*]	(Kobayashi et al., 2015)
NCEP-CFSR	National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR; http://cfs.ncep.noaa.gov/cfsr/)	R [*]	~ 0.38º	Global	Hourly	1979– 2010	(Saha *et al.*, 2010)

[*] NRT= Near Real Time with a delay of several days, G = Gauge, S = Satellite, R = Reanalysis

Figure 5.2: Schematic representation of the GR4J rainfall-runoff model (Source: Perrin et al., 2003). P is rainfall depth; E is potential evapotranspiration estimate; Q is total streamflow; X_i are the model parameters; all other letters are model variables or fluxes summarized in Table 4.

The four free model parameters were calibrated by applying the Kling-Gupta Efficiency (KGE) (Gupta *et al.*, 2009) as the objective function and the daily observed river discharge data of the selected catchments as reference. We use different inputs (precipitation datasets) from CHIRPS, ERA5, ERAI, CFSR and JRA55 to calibrate the GR4J model. The KGE was used also for evaluating the performance of the GR4J model when forced with different reanalysis data. The KGE objective function represents a weighting of three components that correspond to bias, correlation, and variability, ensuring that KGE is sensitive to errors in the overall distribution of streamflow (Kling, Fuchs and Paulin, 2012; Adeyeri *et al.*, 2020). We therefore calculated the hydrological model performance statistics for the calibration and validation periods and compared them across the different reanalysis datasets to investigate the overall suitability of the different reanalyses as input data to simulate river flows. We adopted a threshold of model performance in the range $-0.41 < KGE \leq 1$ as reasonable, following Knoben (2019) being -0.41, the KGE value corresponding to a mean flow benchmark.

A split-sample validation technique (Klemeš, 1986) was used to test model performance beyond the calibration period. For this study, 36 years (1981-2016) of streamflow data for each catchment were available, so we split them into two equal 18-year Split-Sample Testing (SST) periods hereafter referred to as SST1 and SST2 (Figure 3.5).

Thirdly, we perform Sensitivity Analysis by applying the global Sobol' sensitivity method for the GR4J model parameters using the KGE as our target function and the daily observed data of the 19 catchments as reference. We adopt the Sobol' method because it estimates the relative contribution of individual model parameters and their interactions through the decomposition of model output variance (Nossent, Elsen and Bauwens, 2011). A sensitivity analysis allows a reduction of the number of parameters incorporated in the optimization by determining the most influential parameters of a model and their identifiability (Saltelli, Tarantola and Campolongo, 2000). As no prior information is available on the parameters, the input parameter values for the Sensitivity Analysis are sampled from a uniform distribution (Nossent, Elsen and Bauwens, 2011). The different parameter ranges are scaled between 0 and 1 with a linear transformation. Then, we obtain one value of the Sensitivity Indices (SI) per parameter, and we investigate the relative role of each parameter in explaining the output variance and assess possible over-parameterization issues by counting the number of sensitive parameters. Because the value of the objective function for the calibration of parameters can be used as the model performance statistics for sensitivity analysis, we adopted the KGE.

Finally, we assess the overall suitability of the rainfall–runoff model when forced with different meteorological inputs by calculating the Model Suitability Index (MSI). We compare the performance of the four reanalysis datasets across the 19 catchments and investigate which of the input datasets are suitable and which require caution, because of low model performance and possible parameter identifiability or over-parameterization problems. The well-known problem of over-parameterisation due to insensitive parameters in models with large number of parameters (van Griensven *et al.*, 2006) makes sensitivity and performance statistics important. This may result in uncertain model simulations arising from equifinality in model calibration but yielding unequifinal model simulations in validation (Beven, 2012b). This mostly arises from the application of calibrated multiple optimal parameters sets with significantly variable parameter values (Shin et al., 2015;Shin and Kim, 2017). Therefore, in most cases the problem arising from prediction uncertainty may pose problems to modellers when it comes to decision making. By applying the quantitative method of Sobol's SA, we were enabled to couple the results with the performance statistics. The MSI aggregates both sensitivity indices & performance statistics (Shin and Kim, 2017), providing a clear index to judge the relative global performance of the reanalysis products with respect to observations. The computed MSI can be used in comparison studies with any catchment data. If all the model parameters were sensitive, this would yield a MSI of 1 and a hydrograph with perfect matching between the simulations and observations.

We adapt Shin' and Kim (2017) Model Suitability Index (MSI), which is a combined measure of performance statistics and Sensitivity Analysis results. The MSI can be expressed Equation **5.1**:

Equation 5.1:- Model Suitability Index.

$$MSI = 0.5 \times \left(\frac{1}{n} \sum_{i=1}^{n} SR_i \right) + 0.5 \times \left(\frac{1}{m} \sum_{j=1}^{m} PS_j \right)$$

where the SR is the sensitivity ratio (i.e., the ratio of the number of sensitive parameters to the total number of model parameters) ranging from [0, 1] and PS is the performance statistics, *n* is the

number of years over which the sensitivity analysis is run, and *m* is the number of split sample periods in model calibration. It is necessary to set a sensitivity threshold to ascertain the sensitive parameters, hence we adopted a minimum value of 0.2 for the TSI of a sensitive parameter. This value has been suggested and used in some past studies (e.g Van Werkhoven et al 2009 & Shin et al 2013). It is worth noting that this is an arbitrary value, so we acknowledge the need to practise caution when the parameters' TSI values are nearing the threshold. PS is computed by obtaining the average value of all the periods considered (i.e., two split sample periods). To calculate the average PS, we considered the calibration and validation performance statistics (KGE). As both measures are equally important, we gave the same averaged weight to PS and SR in calculating the MSI.

5.6. Results and discussion

5.6.1. Results

5.6.1.1. Overall performance evaluation using observations

The performance of ERA5, ERAI, JRA55 and CFSR on monthly, seasonal, and annual scales is presented in this section. We used the monthly scale as a base time scale and calculated CC, RMSE, MAE and ME for all the four reanalysis products.

5.6.1.1.1. Performance on monthly scale

ERA5, ERAI, JRA55 and CFSR were first evaluated on a monthly timescale with respect to observations at the country level. All the datasets passed the significance test of the correlation coefficient at the 99% confidence level and to eliminate the influence of the seasonal cycle on the values, each Correlation Coefficient was calculated per month as shown in Figure 5.3. ERA5 shows the highest average correlation coefficient value of 0.71 on a monthly timescale compared to observations and is consistently higher across all months (Figure 5.4, and Table 12) than the other reanalysis products. ERAI and CFSR have good average correlation but show larger drops in some months (especially in the drier month of August). JRA55 obtained a poor correlation coefficient of 0.46 on average. In general, ERA5, ERAI and CFSR show higher correlations to observations in rainy months (March-April-May and October- November-December) and lower in the dry months (June-July- August), whereas JRA55 shows the worst correlations during both rainy seasons.

The average twelve months evaluation indices for each of the reanalysis product is shown in Table 12. Overall, ERA5, ERAI and CFSR show a similarly good ability to simulate the precipitation for all the indices under consideration. ERA5 has a better CC, BIAS and RMSE whereas JRA55 has the lowest CC and the largest BIAS and RMSE, suggesting that JRA55 is the worst performing reanalysis dataset over the Kenyan catchments.
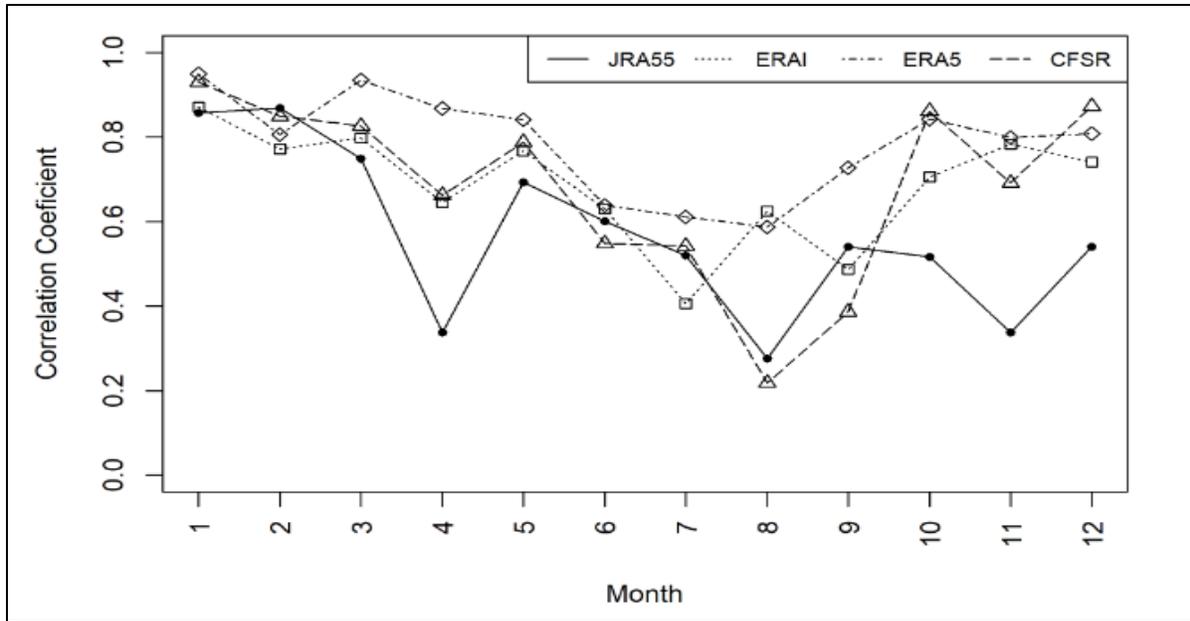
Figure 5.3:- Line graph of correlation coefficients (CC) between monthly observations and ERA5, ERAI, JRA55 and CFSR precipitation for the period 1981 – 2016 on average across the 19 study catchments.

Table 12:- Average CC, BIAS, RMSE, MAE and ME between the four reanalysis precipitation datasets and observations on monthly timescale for the period 1981 – 2016 over all the study catchments.

Index	ERA5	ERA-Interim	JRA55	CFSR
CC	0.71	0.63	0.46	0.68
BIAS (%)	49.72	-26.97	146.66	-76.68
RMSE (mm)	31.59	43.73	115.12	67.36
MAE (mm)	25.11	37.02	79.92	59.95
ME (mm)	1.64	-30.00	71.52	-59.91

5.6.1.1.2. Performance on seasonal and annual timescales

The overall performance of the four reanalyses (ERA5, ERAI, CFSR and JRA55) were evaluated on seasonal and annual timescales to explain the propagation of errors at these timescales. The results of the different performance indices are shown in Table 13.

Table 13: CC, BIAS, RMSE, MAE and ME between the reanalyses and observation precipitation data at a seasonal and annual timescale averaged over the 19 study catchments in Kenya.

Season	Dataset	CC	BIAS (%)	RMSE (mm)	MAE (mm)	ME (mm)
MAM	JRA55	0.34	155	965.10	898.12	898.12
	ERAI	0.59	-28.1	193.87	168.09	-160.36
	ERA5	0.88	47.6	322.44	282.42	275.60
	CFSR	0.78	-84.6	481.03	475.14	-475.14
JJA	JRA55	0.52	67.4	361.72	290.77	267.76
	ERAI	0.24	-39.5	182.05	157.12	-154.93
	ERA5	0.25	-1.4	85.66	70.90	-5.446
	CFSR	0.22	-84.6	332.89	327.95	-327.95
OND	JRA55	0.44	271.7	1036.4	983.25	983.25
	ERAI	0.81	-27.9	123.34	111.09	-100.56
	ERA5	0.52	96.6	397.00	349.70	349.70
	CFSR	0.84	-84.3	307.76	296.47	-296.47
ANNUAL	JRA55	0.25	171	2902.21	2760.63	2760.63
	ERAI	0.46	-26.3	462.513	421.14	-421.14
	ERA5	0.52	44.7	801.75	728.57	720.70
	CFSR	0.60	-85.2	1354.36	1349.21	-1349.21

The overall correlation coefficients on seasonal and annual timescale are shown in Figure 5.4. Higher CC across all the datasets were obtained in the wet seasons of MAM and OND, whereas lower CC were obtained in the dry season of JJA, with the performance index higher in OND than in MAM. ERA5 obtained the highest CC (0.88) in MAM, whereas CFSR the highest (0.84) in OND. JRA55 showed lower CC of 0.34 and 0.44 in the two seasons respectively and a CC of 0.52 in the dry season, depicting a tendency of a wet bias over the dry months. On average, the variability in the CC index across the four datasets was relatively lower in the OND season and higher in the MAM season. The BIAS across the four datasets was lower in the dry season (JJA) and higher in the wet seasons (MAM &OND) with JRA55 showing a higher positive BIAS across all the seasons. There are large values in the RMSE and the MAE across the four datasets in the two wet seasons and this may be linked to the high precipitation concentrations during those seasons across most of the catchments. Generally, it can be noted that JRA55 shows the worst performance in comparison to observation, especially in the wet seasons of MAM and OND, but obtained relatively better scores in the dry season of JJA. ERA5 shows better agreement with observations across the three seasons, so may be an appropriate option for simulating precipitation over the Kenyan catchments.
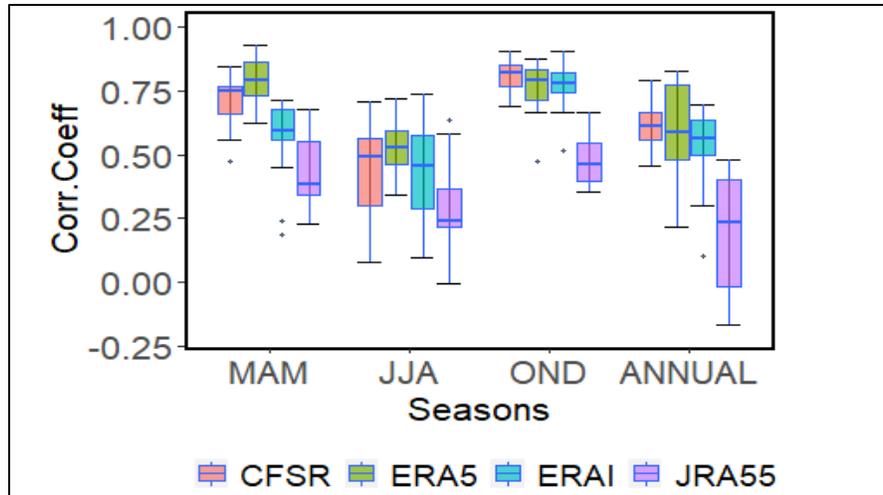
Figure 5.4: Boxplots of the seasonal (MAM, JJA and OND) and annual Correlation Coefficients (CC) for four reanalysis CFSR (pink), ERA5(green), ERAI (blue) and JRA55 (purple) across the 19 catchments. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.

On an annual timescale, the average annual precipitation of CFSR, ERA5, ERAI and JRA55 was computed and compared with the observation (CHIRPS) (Figure 5.5). ERA5, ERAI and CFSR show a similar trend compared to observations across all the years, with CFSR and ERAI underestimating the precipitation. JRA55 shows a higher tendency to overestimate the annual precipitation over the study catchments. In terms of performance indices, CFSR, ERAI and ERA5 showed better CC indices of 0.60, 0.46, 0.52 respectively whereas JRA obtained lower CC of 0.25 (Figure 5.4). The variability in the CC was higher in JRA55 (Figure 5.4). ERA5 and JRA55 show a positive bias of 45% and 171% respectively, whereas ERAI and CFSR show negative bias of -26 and -85%. ERA5 has a lower RMSE and ME whereas JRA55 has the highest. These results show that ERA5 is the best performing reanalysis dataset compared to observations on annual timescales whereas JRA55 is the worst performing.
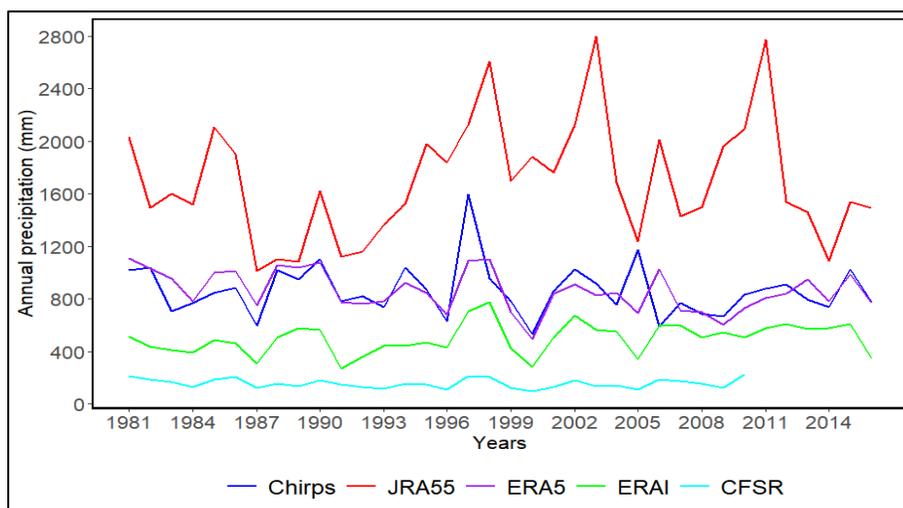
Figure 5.5: Areal average annual precipitation from the observations and the reanalysis datasets, averaged across the 19 study catchments.

123

The mean monthly and seasonal standardized precipitation anomalies in the four-reanalysis precipitation for a base climatological period 1981–2016 is shown in Figure 5.6. On a monthly timescale, the observations show a positive anomaly over the central highlands and the western parts of Kenya (Figure 5.6, pan1). The arid and semi-arid parts in the eastern and coastal lowlands show a negative anomaly (dry bias). This pattern is also captured in ERA5, ERAI and JRA55, although JRA55 has excessively high and widespread negative anomalies compared to the former two. On seasonal timescales, ERA5, ERAI and CFSR show positive anomalies in the western and central highlands in all three seasons, except for JRA55 which has a stronger negative and positive anomaly in the MAM and OND seasons respectively.
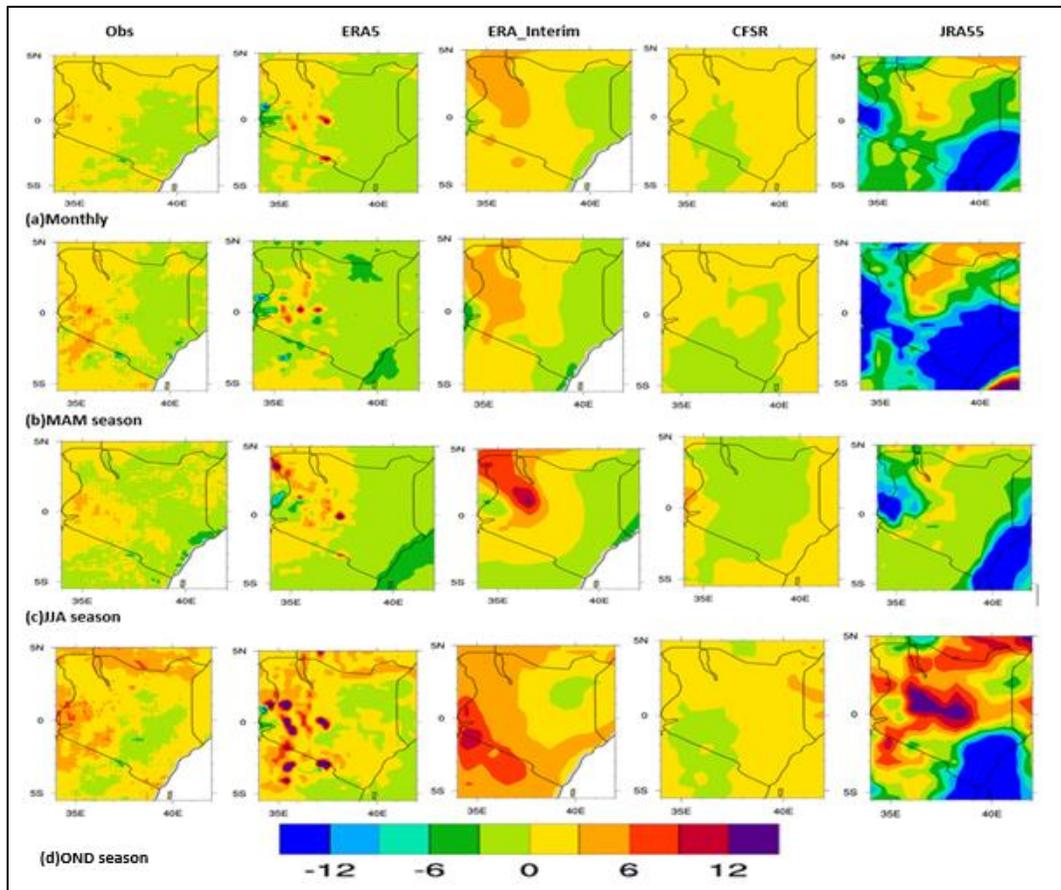
Figure 5.6: Mean monthly and seasonal standardized precipitation anomalies in the four reanalysis products for the 1981–2016 period: (a) Monthly anomalies, (b) MAM season, (c) JJA and (d) OND anomaly index in seasonal precipitation, for ERA5 (2nd column), ERA-I (3rd column), CFSR (4th column) and JRA-55 (5th column). Column 1 shows the observations (CHIRPS).

An evaluation of the extreme precipitation in the four reanalyses was also performed (**Figure 5.7**). For this case, we focused on the 95th percentile of rainy days for the MAM, JJA and OND seasons during the period 1981-2016. A rainy day represents a day for which the recorded precipitation amount is greater than or equal to 1mm (Gudoshava *et al.*, 2020). The observed extreme precipitation varied between 60mm and more than 240mm across western, central highlands and coastal catchments for the rainy seasons (MAM and OND). The observed extreme precipitation

during the dry season JJA varied between 100mm and 160mm across the western catchments only, whereas in the other regions the observed precipitation was less than 60mm. Our results show that CFSR and ERAI show a positive bias for extreme precipitation across most parts of the country in all the three seasons, like the results in Garibay *et al.* (2019). JRA55 has an enhanced negative bias for extreme precipitation in most parts of the country, except for an isolated positive bias in the central highlands region in the JJA and OND seasons. ERA5 has a positive bias in MAM and OND in most parts of the country with some patches of negative bias in the western and central highlands catchments. It has an enhanced negative bias in the JJA season with a positive bias in the western and coastal strip. We conclude that ERA5 outperforms other reanalysis products as it captures the wet extremes over the regions in which observations show enhanced precipitation in the respective seasons. The results are consistent with the findings in Gleixner *et al.* (2020) which showed both ERA5 and ERAI to have the capability to capture wet extremes in the dry seasons, with ERA5 matching observations more closely than the excessively wet ERA-interim. A promising performance by ERA5 in simulating wet extremes can be attributed to an improved bias correction method which incorporates aircraft measurements, satellite radiances, radiosonde measurements and surface pressure (Probst and Mauser, 2022). In addition, better performance in the central highlands can be attributed to the improved horizontal resolution in ERA5, which results in better estimates in orographic precipitation.
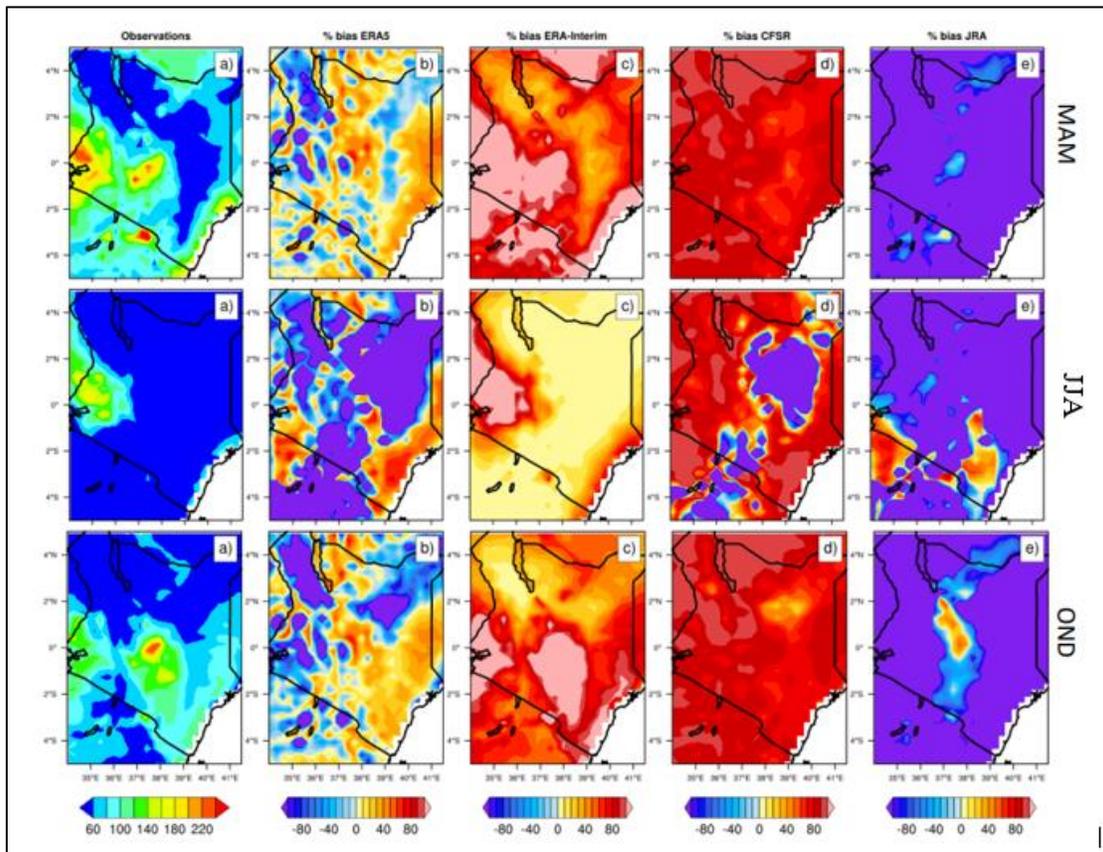
Figure 5.7: Seasonal observed precipitation (mm) and mean bias (%) of the extreme rainy days at 95th percentile in the four reanalysis products for 1981–2016 period: (a) Long-term observed (OBS) seasonal average precipitation (mm) from CHIRPS, (b– e) mean relative bias (%), in seasonal precipitation in ERA5 (b), ERA-

Interim (c), CFSR (d), JRA-55 (e), with respect to CHIRPS. MAM season (top), JJA (middle) and OND (bottom panel).

5.6.1.2. Evaluation of the reanalyses as inputs for hydrological modelling

5.6.1.2.1. Assessment of the overall model performance using different reanalysis

The performance of the four reanalysis datasets was evaluated using the GR4J model in the 19 catchments for the period spanning 1981 -2016. The KGE in calibration (top panel) and validation (bottom panel) scores obtained using different datasets for each of the catchments are represented in **Figure 5.8**. Overall, wetland catchments in the western and highlands of Kenya obtained relatively better calibration scores than those in the semi-arid regions, with Yala, Sio, Nzioa and Gucha (wetland catchments) performing best and Perkerra, Ndo, Tsavo, Thiba and Tana (semi-arid catchments) performing worst. For each of the catchments, ERA5 showed better calibrated KGE scores compared to observations, while CFSR and JRA55 obtained poorer KGE scores. However, we are cautious in the interpretation of our results in terms of performance criteria, because these catchments are strongly influenced by human activities such as irrigation schemes and dams. Consequently, the low performance in some catchments may not be solely due to uncertainty in the input data.
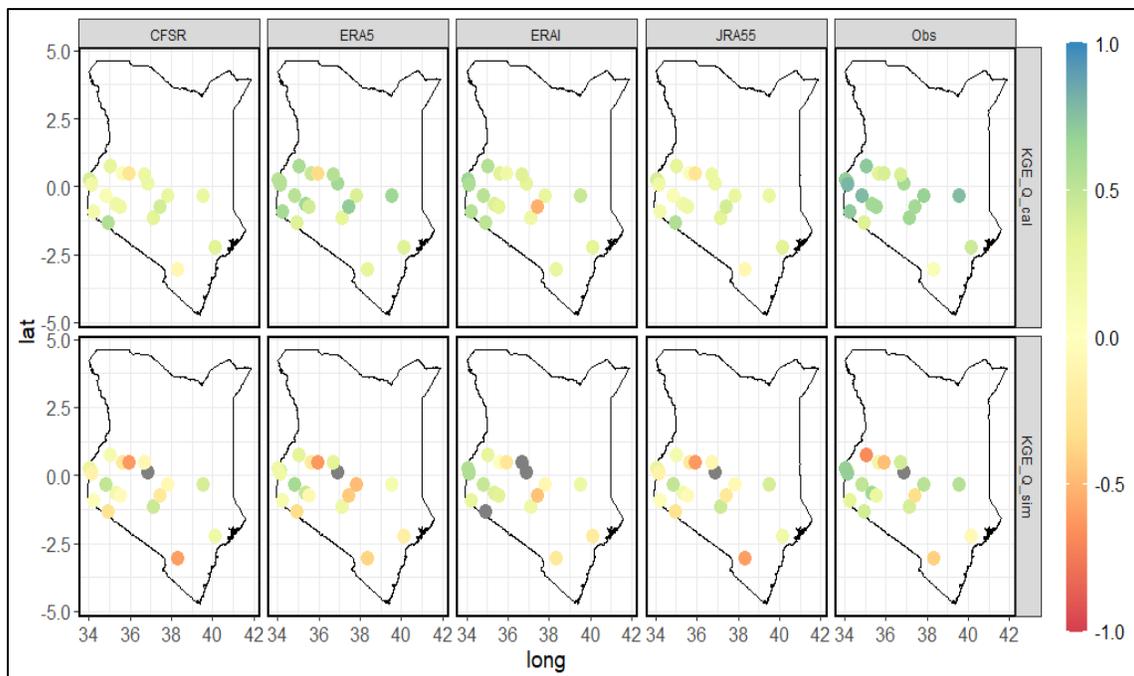
Figure 5.8: GR4J model performance (KGE) in calibration (top panel) and validation (bottom panel) across the 19 catchments for different input datasets, Pan.1 (CFSR), Pan.2 (ERA5), Pan.3 (ERAI), Pan. 4 (JRA55), Pan.6 (CHIRPS).

The overall variability in GR4J model KGE scores across the four reanalyses are shown in **Figure 5.9**. There are overall high-performance scores (KGE>0.5) in calibration mode in about half of the catchments for all datasets except CFSR, which suggests that using CFSR as hydrological model inputs in the region causes problems that cannot be solved or compensated by calibration. ERA5, ERAI and JRA55 show similar overall performance compared to observation in **Figure 5.9a**. The

range for the performance statistic is narrower in the ERA5, ERAI and JRA55, indicating a more stable model performance in the region, while it is wider in the CFSR data (Figure 5.9a). In validation mode, the performance markedly decreases, as expected, for all datasets (Figure 5.9b): ERAI, ERA5 and JRA55 have the highest median KGE value (just above or about 0) whereas CFSR has the lowest median values (KGE<-0.5). The range of KGE values is relatively larger compared to observations; thus, a relatively unstable prediction ability is expected for streamflow in reanalysis in the region. The range of performances is more variable in ERA5 and JRA55 and less variable in ERAI. Overall, the variability in KGE values is higher in validation than in calibration across all the reanalyses compared to observations, as expected. The percentage bias of the KGE component in each catchment in calibration (top panel) and in validation (bottom panel) is shown in (Figure 5.10). The bias in all the four reanalysis is higher in calibration, whereas in validation most catchments exhibit lower biases, except for Perkerra.
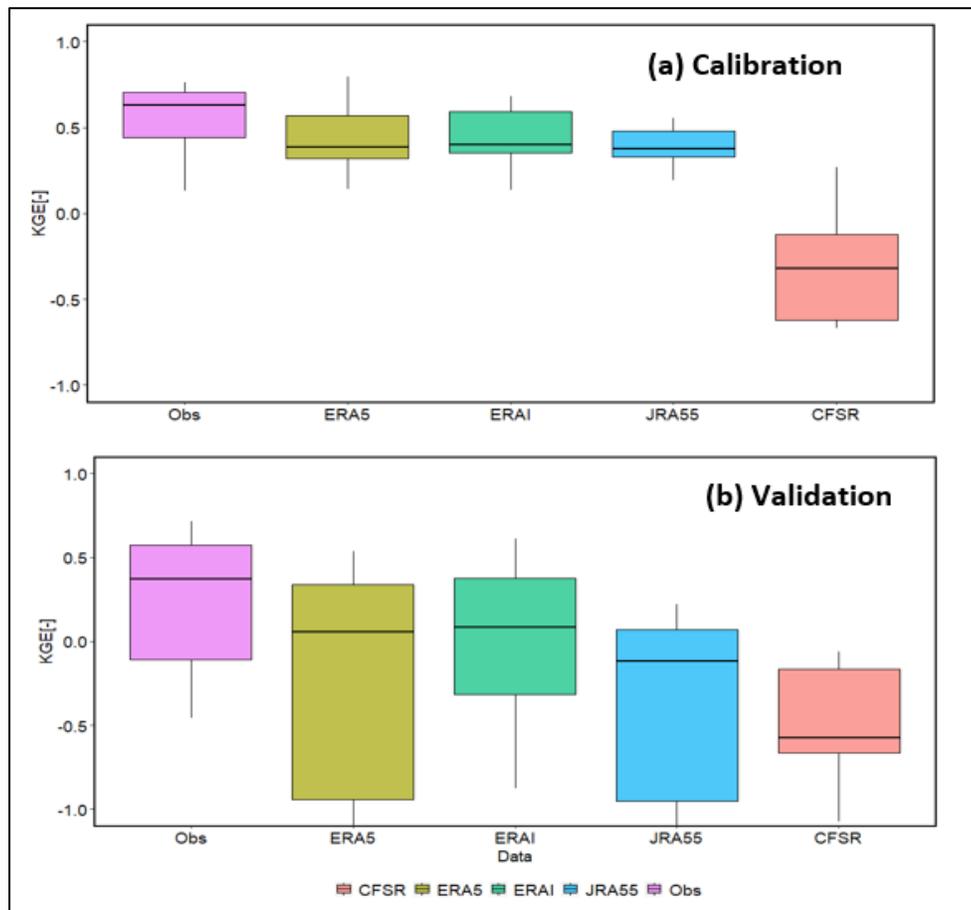
Figure 5.9: Boxplots of the overall GR4J model performance (KGE [-]) in (a) calibration and (b) validation mode over the 19 catchments.
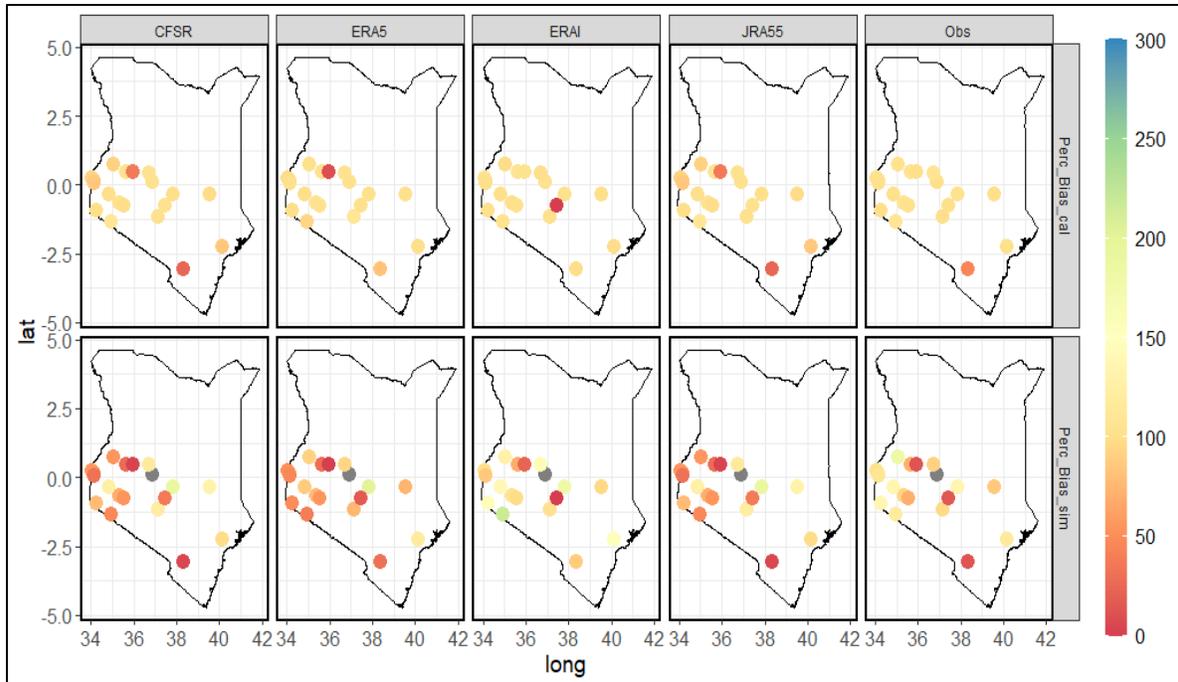
Figure 5.10: GR4J model performance (Percentage KGE- Bias) in calibration (top panel) and validation (bottom panel) across the 19 catchments.

5.6.1.3.   Sensitivity analysis results

5.6.1.3.1. Variability in sensitivity of model parameters

The GR4J's model maximum and minimum TSIs for the four reanalysis datasets is illustrated in **Figure 5.11**. The maximum and minimum TSIs represent the variability of parameter sensitivity within the catchment with respect to KGE over the sampling periods and the variation across the four-reanalysis relative to observations. If the maximum and minimum TSIs for a parameter are equal (on the one-to-one line), that parameter has the same TSI for the sampling period, implying that the parameter is more stable across time, and would be expected to vary depending on the catchment characteristics and input data as well. In all the four datasets, the routing parameter (X4) related to the unit hydrograph is evidently the least sensitive as it is far below the threshold, followed by the capacity of the routing store (X3), whereas the two parameters governing the water balance, i.e. the soil moisture accounting store (X1) and the groundwater exchange coefficient (X2), are the most sensitive across the datasets in most of the catchments, except for CFSR, where X1 is less sensitive.

128
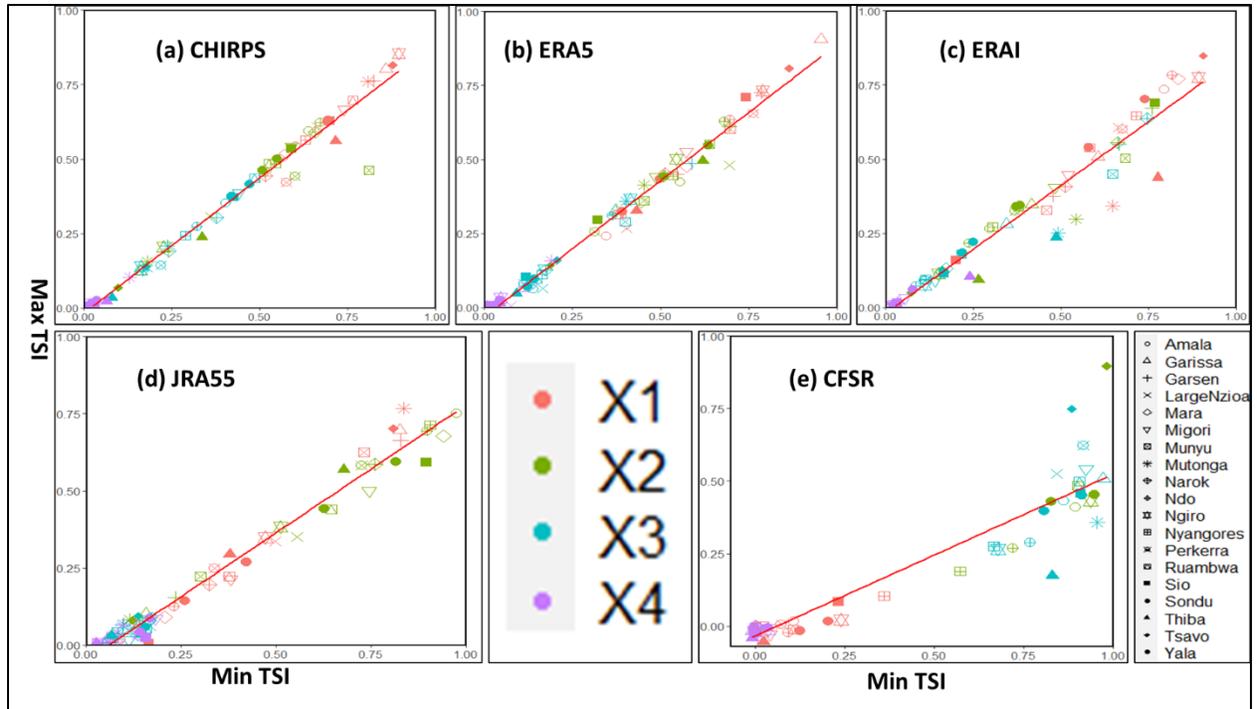
Figure 5.11: Scatter plot of Sobol' Total Sensitivity indices (TSI) for the different reanalysis datasets and the GR4J model parameters for the nineteen catchments. Minimum and maximum TSI were calculated for the whole data period. (a) Chirps, (b) ERA5, (c) ERAI, (d) JRA55 and (e) CFSR . The diagonal line is the line of best fit.

Observations show more stability for the parameters for all catchments except six (Munyu, Thiba, Ndo, Ewaso Ngiro, Perkerra and Tsavo) with respect to reanalysis datasets (Figure 5.11a). In ERA5, most of the catchments showed stability in parameters except in Thiba, Tsavo, Large Nzioa and Ewaso Ngiro catchments (Figure 5.11b). In ERAI, there is high variability in model parameter stability with less stability for some catchments such as Thiba, Munyu, Mutonga and Tsavo catchments (Figure 5.11c). In JRA55 and CFSR (Figure 5.11d and Figure 5.11e respectively), the departure in sensitivity of model parameters from the diagonal is pronounced across most of the catchments. Overall, the variability in sensitivity of model parameters is high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, so we can conclude that the reanalysis datasets are not suitable for model calibration in these catchments of Kenya, that are characterized by arid and semi- arid conditions. However, the catchment's water balance may be strongly affected by the dams constructed in the upstream areas and the massive irrigation schemes, which result in water attenuation.

5.6.1.3.2. Overall sensitivity of GR4J model parameters

The parameters related to water balance, i.e., the soil moisture accounting store (X1) and ground water exchange (X2), show higher sensitivity across all the four datasets except for CFSR, in which the production accounting store is less sensitive and falls below the threshold value of 0.2 from the model TSIs (Figure 5.12). The first parameter responsible for water routing (X3) is less sensitive for most datasets (except CFSR), whereas the unit hydrograph parameter (X4) is the least sensitive across all the catchments in all datasets. In comparison to observations, ERA5, ERAI and JRA55 show similar parameter sensitivities to model parameters, while CFSR show distinctly higher

variability and a difference in the parameters' sensitivity, which points to high uncertainty in the CFSR dataset. This result shows that the sensitivity of the model parameters can change with the input datasets, having very different hydrological characteristics.
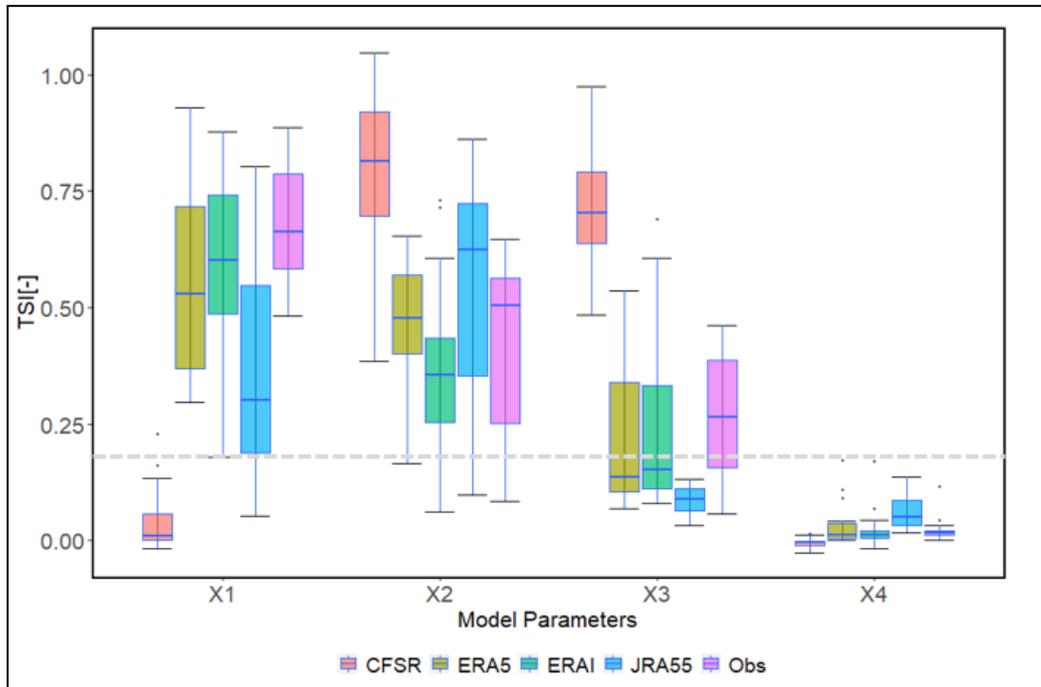
Figure 5.12: Boxplots of the Sobol' Total Sensitivity Indices (TSI) for the GR4J parameters for Obs. (pink), CFSR (orange), JRA55(blue), ERAI (green) and ERA5(forest-green) over the nineteen catchments. Dashed grey line represents the sensitivity threshold. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.

5.6.1.3.3. Comparison of reanalysis datasets using model suitability index

When the sensitivity indices and performance statistics are considered, it is difficult to determine which dataset is most appropriate**.** ERA5 and ERAI datasets, for example, had good and clear parameter sensitivities that captured their purposes, and the model performance score median values were higher than in CFSR and JRA55. However, the range of the performance statistics across the catchments was sometimes wider than in the other datasets, resulting in higher simulation uncertainty. When compared to the other two methods, the MSI, which considers both sensitivity indices and performance statistics, has the advantage of making it easier and clearer to judge the superiority and inferiority of the datasets in terms of both model performance and parameter identifiability. We determined that a value of 0.5 for the MSI is a good MSI threshold (Moriasi *et al.*, 2007). We give the same weight to model performance and sensitivity, as described in the subsection 'MSI'; thus, the threshold value for good MSI is 0.5.
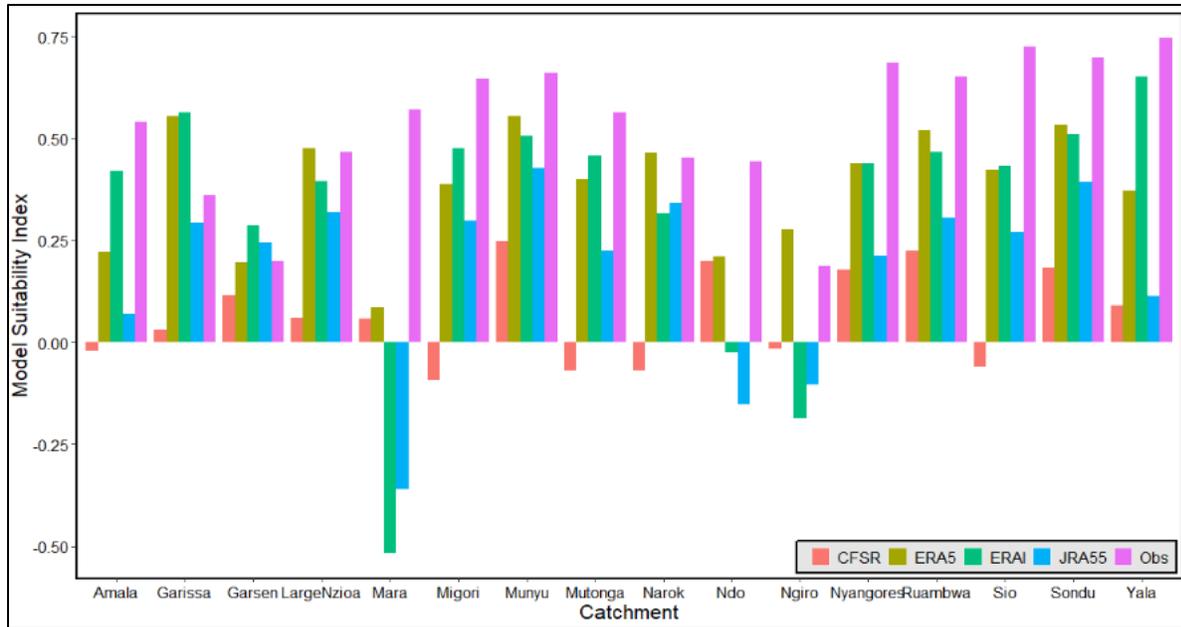
Figure 5.13: Bar chart showing a comparison of model suitability in terms of performance and parameter sensitivity across different reanalysis using the Model Suitability Index (MSI).

Combining the model performance and sensitivity indices discussed in Section 5.6.1.1 and 5.6.1.3, MSIs for all the reanalysis datasets are shown by the bar graph (Figure 5.13) . The ERA5 has the highest MSI compared to observations across the nineteen catchments, followed by the ERAI reanalysis. As a result, the ERA5 and ERAI reanalyses are appropriate, at least for the selected sample of Kenyan catchments, whereas CFSR and JRA55 are least appropriate as they show lower MSI values across most of the catchments. CFSR shows negative MSI values for Amala, Migori, Mutonga, Narok, Ewaso Ngiro and Sio catchments, meaning it is not appropriate for application in these Kenyan catchments. Overall, the four reanalysis datasets obtained relatively lower MSI values in Mara, Ndo, Ewasi Ngiro and Tana Garsen catchments. These catchments are mainly in arid and semi-arid areas of Kenya (Table 14).

5.7.   Discussion

5.7.1. Overall performance of reanalysis precipitation products

In this study, we assessed four reanalysis precipitation products relative to observations for the period 1981 to 2016 on monthly, seasonal, and annual timescales. We also assessed how best they simulate streamflow, using the GR4J model and sensitivity analysis for 19 catchments located in distinct geographical and climatic environments. Results show that the ERA5 reanalysis outperforms the other reanalysis products on monthly and seasonal scales, whereas CFSR outperforms ERA5 on annual and seasonal timescales. In general, ERA5 data were often closer to observations than other reanalysis data, which corresponds with earlier research on the datasets in different regions (e.g., Betts *et al*., 2019; Gleixner *et al*., 2020; Tarek *et al*., 2020), even though these studies considered different evaluation periods, spatiotemporal resolutions, hydrologic models, and climates. However, the performance scores for the reanalysis products over the

Kenyan catchments were lower, in contrast to some of the studies carried out in other parts of the world with varying climates (e.g., Dhanya *et al*., 2017; Harada *et al*., 2016; Mahto *et al*., 2019; Wang *et al*., 2019), which obtained higher scores for their study areas. The low performance scores may be due to variations in the initial resolution of the datasets (Chen, Brissette and Chen, 2018; Lemma, Upadhyaya and Ramsankaran, 2019) and the interpolation approach is likely to have some influence on the evaluation of various reanalysis data (Rapaić *et al.*, 2015; Zhang and Academy, 2016). It is also worth noting that while the observed precipitation data are the best estimates available, they are likely to be subject to errors too (Beck *at al*., 2017, Dinku *et al*., 2019). In addition, the seasonality of rainfall over Kenya is greatly influenced by weather phenomena such as El Niño -Southern Oscillation (ENSO) and Indian Ocean dipole (IOD) (Ayugi *et al*., 2020; Ojara *et al*., 2021; Onyutha, 2016) and play a major role in extreme rainfall events and inter-annual variability (Ongoma *et al*., 2015). For example, the warm phase of ENSO/El Niño results in unusually heavy rainfall, causing rare floods like the 1997/1998 occurrence (Takaoka, 2005). ERA5, ERAI and JRA55 picked the enhanced annual precipitation totals of the strong El Niño years such as 1997/98 and 2015. However, relative to observations, ERA5 and ERAI underestimated the rainfall, and this may be attributed to incorrect configuration in the reanalysis products. For example, ERA5 precipitation is not customized to pick up the perturbations caused by the changes in the ocean-atmosphere interactions and the mountainous regions and so may miss picking up the extremes caused by events such as ENSO, thus leading to the low performance scores.

Standardized precipitation anomalies in ERA5, ERAI and CFSR show a positive anomaly over the central highlands and the western parts of Kenya (Figure 5.6a, pan1) and a negative anomaly in arid and semi-arid regions in the eastern and coastal lowlands in the three seasons (MAM, OND, JJA) except for JRA55, which has a stronger negative and positive anomaly in the MAM and OND seasons respectively. This is consistent with a study by Ongoma *et al*. (2018), which indicates a rise in the severity of severe precipitation events shown by a positive standardized rainfall anomaly over East Africa, including the aforementioned regions in Kenya. With the changing climate, temperatures in the region are projected to rise by the end of the twenty-first century, leading to an increase in extreme rainfall occurrences (Ongoma and Chen, 2017), thus exacerbating flood risk.

Our analysis of the accuracy of precipitation reanalysis with respect to observations, over different timescales from monthly to annual, showed a positive but relatively small bias in CFSR, ERA5 and ERAI and a larger negative bias in JRA55 in the MAM and OND seasons. Moreover, the first three reanalysis datasets showed a good average correlation at the monthly and seasonal scales. Therefore, the three reanalysis products have the potential to capture the rainfall seasonality and events in the study area. Recent worldwide research show that the frequency, severity, geographical range, length, and timing of severe climatic events are changing (Wainwright *et al*., 2021). A rise in severe rainfall events such as very wet days (R95p) and extremely wet days (R99p) anticipated for the future (2021–2100) (Gudoshava *et al*., 2020), is likely to cause the loss of life and devastation to property owing to an increase in flood intensity (Finnley, 2020). Therefore,

further work should assess the capacity of the reanalysis datasets in capturing extreme rainfall event characteristics, such as timing and daily peaks.

5.7.2. Performance of reanalysis as inputs into a hydrological model

Using a hydrological model as integrator to compare simulated and observed streamflow, which can operate as an independent validation variable, is one approach to assessing the quality of observation and reanalysis precipitation data. Each set of reanalysis data on precipitation and estimated potential evapotranspiration was supplied to the GR4J model, which was subsequently calibrated for each combination (consistently using precipitation and potential evapotranspiration from the same dataset), to analyze independently the quality of input data for each dataset relative to the observed streamflow gauge data. Streamflow gauges, of course, are subject to a variety of inaccuracies (Baldassarre and Montananari, 2009), but they represent the best available estimates for this study. Results of KGE scores show that ERA5 is better than ERAI, JRA55 and CFSR but, on overage, all the reanalyses are less skillful relative to observations across the catchments in this study and this is entirely due to the precipitation data quality. However, there is a marked improvement in the KGE scores for the catchments in the central highlands and western wet catchments which agrees with some studies on the datasets in other regions (e.g., Terek *et al.*, 2020; Essou *et al.*, 2017; Lakew *et al.*, 2016), pointing to the fact that reanalysis data can be used as a replacement for observations.

There are overall high-performance scores (KGE>0.5) in calibration mode in about half of the catchments for all datasets except CFSR, which suggests problems in using CFSR to reproduce the hydrological water balance in the region that cannot be solved or compensated for by calibration (Diro *et al.*, 2009).

5.7.3. Sensitivity analysis of model parameters

Sensitivity analysis is useful in identification of the parameters that have a strong impact on the model outputs, which in turn influence the effectiveness of the model. The greater the sensitivity of the model response to a parameter, the more effectively and promptly will that parameter be optimized, so high sensitivity is good. Such an in-depth analysis of a hydrological model may (i) help to identify any potential deficiencies in model structure and formulation; (ii) provide guidance for model parameterization; and (iii) provide the information content of available input data.

Based on provision of information content of the input data, different reanalyses show different sensitivities of model parameters, and one that provides a higher sensitivity of model response has less uncertainty and it may be a lot easier to parameterize the values, but then this does not reach the real value in actuality; on the other hand, the dataset with low sensitivity has got high uncertainty and model parameterization may be very difficult (Zeng *et al.*, 2019). In comparison to observations, ERA5, ERAI and JRA55 show similar sensitivities in model parameters while CFSR shows distinctly higher variability and a difference in the sensitive parameters, which points to high uncertainty in the CFSR dataset. This result shows that the sensitivity of the model parameters can change with the input datasets, having very different hydrological characteristics. Overall, the variability in sensitivity of model parameters is high in Thiba, Munyu, Pekerra and

Ewaso Ngiro across all the datasets, so we can conclude that the reanalysis datasets are not suitable for model calibration is those catchments of Kenya that are characterized by arid and semi- arid conditions. MSI considers both model performance and uncertainty quantitatively, therefore it can be used to compare any catchment. The ERA5 has the highest MSI compared to observations across the nineteen catchments, followed by the ERAI and JRA55, whereas CFSR has least MSI values. MSI's dependability may be increased by including more sensitivity indices and performance scores as well as assigning weights to the scores.

5.8.   Summary and conclusions

This study addresses a notable gap that was found in the literature on evaluating the accuracy of multiple precipitation reanalysis datasets across data-scarce regions like Kenya, and on assessing their potential to supplement scarce rain gauge observations for hydrological modelling. Four different state-of-the-art reanalysis datasets were assessed. Precipitation data from ERA5 shows the highest average correlation coefficient value (0.71) on a monthly timescale compared to observations and is consistently higher across all months than the other reanalyses. ERAI and CFSR have good average correlation but show larger drops in some months (especially in the drier month of August). JRA55 obtained a poor correlation coefficient of 0.46 on average. ERA5, ERAI and CFSR show higher correlations with observations in rainy months (March-April-May and October- November-December) and lower in the dry months (June-July- August), whereas JRA55 shows the worst correlations during both rainy seasons. On annual timescales, CFSR, ERAI and ERA5 showed better CC indices of 0.60, 0.46, 0.52 respectively whereas JRA obtained lower CC of 0.25. ERA5 and JRA55 show a positive bias of 45% and 171% respectively, whereas ERAI and CFSR show a negative bias of -26 and -85%.

Spatial rainfall patterns directly affect temporal distribution, key in driving runoff and soil erosion processes, which is useful in management of hydrological risks and generation of sediments from rainwater (Peña-Angulo et al 2020). Monthly standardised anomaly maps in ERA5, ERAI and JRA55 showed a positive anomaly over the central highland and western parts. In the arid and semi-arid parts in the eastern and coastal lowlands parts of Kenya the three datasets showed an enhanced negative anomaly. On seasonal timescales, ERA5, ERAI and CFSR show positive anomalies in western and central highland regions in the three seasons, except for JRA55, which has a stronger negative and positive anomaly in MAM and OND seasons respectively. Extreme precipitation showed a positive bias in CFSR, ERA5 and ERAI in MAM and OND seasons, whereas JRA55 has enhanced negative bias in most parts of the country except for isolated positive bias in the central highland region in JJA and OND seasons.

The performance of the GR4J model, when forced by different reanalysis in the 19 catchments, reveals a bigger role forf localized catchment characteristics and processes in model calibration. Wetland catchments in the western and highland regions of Kenya obtained relatively better calibration scores compared to those in the semi-arid regions, with Yala, Sio, Nzioa and Gucha (wetland catchments) performing best and Perkerra, Ndo, Tsavo, Thiba and Tana (semi-arid catchments) performing worst. For each of the catchments, ERA5 showed better calibrated KGE scores compared to observations, while CFSR and JRA55 obtained poorer KGE scores. The range of KGE values was relatively larger compared to observations; thus, a relatively unstable

prediction ability is expected for streamflow in reanalysis for Kenyan catchments. The range of performances is more variable in ERA5 and JRA55 and less variable in ERAI. Overall, the variability in KGE values is higher in validation than in calibration across all the reanalyses compared to observations, as expected.

Sensitivity analysis facilitates the reduction of the number of parameters incorporated in optimization by determining the convergence of the most influential model parameters. Sensitivity analysis revealed that in all the four datasets, the routing parameter (X4) related to the unit hydrograph was evidently the least sensitive, followed by the capacity of the routing store (X3), whereas the two parameters governing the water balance, i.e. the soil moisture accounting store (X1) and the groundwater exchange coefficient (X2), were the most sensitive across the datasets in most of the catchments, except for CFSR, where X1 was less sensitive, with ERA5 showing the highest sensitivity in the model parameters. However, the variability in sensitivity of model parameters was high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, so we conclude that model calibration using reanalysis date in arid and semi- arid catchments of Kenya does not yield satisfactory results. The MSI aggregates both sensitivity indices & performance statistics, providing a clear index to judge the superiority (or inferiority) of a reanalysis with respect to observations. On average ERA5, ERAI (& JRA55) have better MSI scores across most of the Kenyan catchments: ERAI & ERA5 perform better than JRA55 & CFSR, and lead to more robust model parameters. Using a catchment model and combined sensitivity-model performance analysis allows an evaluation of the impact of the variability in the rainfall products throughout the catchment modelling process

In conclusion, in this study we have demonstrated the usefulness of reanalysis rainfall products as potential alternatives for hydrological applications in Kenya. We assessed the suitability of reanalysis precipitation datasets for hydrological modelling across Kenyan catchments, but first assessed the propagation of errors when reanalysis is compared to observations. We performed the assessment on monthly, seasonal, and annual timescales. Then, using a lumped bucket-style hydrological model, we assessed the model performance via the KGE criterion and parameter uncertainty via Sobol's Sensitivity Analysis for four different reanalyses: - ERA5, ERAI, JRA55 and CFSR across 19 catchments. The parametric and model input uncertainty is investigated using the sensitivity indices and the comprehensive model performance analysis is used to examine the model's input strength, i.e., the extent to which the model captures the dynamics of rainfall–runoff processes with respect to different forcing. We also coupled the results of the performance scores and sensitivity indices to compute MSI for the 19 catchments.

We acknowledge the value and necessity of additional work, if reliable data at higher temporal frequency becomes available and can be used, as it contains more information. However, this is a big limitation for the current study due to extensive data gaps in the daily data (river discharge data used in the current study) and the lack of higher temporal resolution hydrological data. Future work should concentrate on assessing the sub-daily performance of hydrological modelling with reanalysis, testing its quality on other additional catchments in countries in the region with quality observed gauge data, but prior investments in data collection in Kenya seem to be needed. Our approach may be extended to various conceptual rainfall -runoff models as well as physically based

distributed rainfall–runoff models. The MSI analysis is a practical method for weeding out the appropriate model input on a catchment scale basis, but a more robust analysis where weights are assigned would yield some improvements in the results. To fully ascertain the potential of alternative model forcing, catchments' individual characteristics and human influence, such as dams and reservoirs, should be modelled

Finally, it is essential to note that this work does not promote the use of products such as reanalysis to replace observed data from weather stations, nor can it be understood as giving reasons to continue the present trend of retiring additional stations. Quality controlled ground observations still act as the best data for research. The ERA5 results demonstrate that atmospheric reanalysis has probably reached the stage where it can consistently supplement data from weather stations and offer trustworthy proxies in places with less dense station networks, at least across Kenya. Overall, reanalysis can be a viable alternative to observations in ungauged catchments, but the associated uncertainties need to be carefully communicated to ensure an informed choice of hydrological modelling tools.

5.9.   Acknowledgements

Chapter 6

6. Detecting trends in flood series and shifts in flood timing across Kenya

6.1. Objective addressed and publication details

Observations show a shift in timing and variability in flood occurrences in most parts of the country. Trend analysis of peak over threshold (POT) and annual maximum (AMAX) flood series are useful in detecting and supporting the evidence of change in flow series, as well as variability in flood timing. Flood peaks are identified using a threshold technique from Kenyan daily discharge data, and notable patterns in the AMAX series are compared to those in the POT series, which is created for three distinct exceedance criteria. A comparison is made between trends in the observed and simulated AMAX. The timing and variability of the annual floods is determined from the AMAX flow. Findings show that the AMAX series detects more trends in flood magnitude than the POT series, while the POT series detects more significant trends in flood frequency than flood magnitude. Sensitivity of trends to different exceedance thresholds selection reveals differing trend patterns across the stations. Flood timing is in the peak rainfall months of April, May and November and is highly predictable in most of the coastal and western stations, and less predictable in stations whose annual floods occur in the dry months of June, July, and August. This information on flood characteristics can help to inform policy for disaster risk management, infrastructure design and agriculture, and ultimately support the improvement of livelihoods in Kenya.

At the time of submission of this thesis, this paper had been submitted to the *Hydrological Sciences Journal,* Manuscript reference HSJ-2022-0371, for publication and the exact copy of the submitted version is in Appendix **A4** of this thesis.

6.2. Authors' contributions

Maureen A. Wanzala (65%): - Conceptualization, Data curation- Lead, Methodology- Lead, Investigation, Formal analysis - Lead, Visualization, Writing – original draft, Writing - Review & Editing. Hannah L. Cloke (15%): Supervision, Writing - Review & Editing, Project administration, Elisabeth M. Stephens (10%): Supervision, Methodology, Writing - Review & Editing- Support, Project administration, Andrea Ficchì (10%): Conceptualization, Supervision, Methodology- Support, Analysis - Support, Visualization, Shaun Harrigan (5%): Methodology - Support, Writing – Review & Editing-Support.

6.3. Introduction

Flooding is among the most detrimental natural hazards worldwide (Berghuijs *et al.*, 2019), and with a changing climate there is an expected increase in flood risk globally (Arnell and Gosling, 2016; Liu *et al.*, 2018). For Kenya, floods are the most common climatic extreme and the leading hydro-meteorological disaster (Huho and Kosonei, 2014). There is a growing concern that major flooding events in many parts of Kenya in the past decade are indicative of the effects of a changing climate (Wainwright *at al* 2021; Wanzala and Ogallo, 2020, Kilavi *et al*., 2018; McLeod *et al.,* 2021).

Understanding flood characteristics such as frequency, magnitude and timing is important for informing policy for disaster risk management, infrastructure design and agriculture, amongst

other hydrological applications (Rosner, Vogel and Kirshen, 2014; Bezak, Brilly and Šraj, 2016). Such assessments require information on the probable year to year variations in flood characteristics (Parry *et al.*, 2007; Kundzewicz *et al.*, 2014). In addition, consideration of the trends in flood data series may result in more accurate flood timing, magnitude and frequency estimations (Berghuijs *et al.*, 2017, 2019; Mangini *et al.*, 2018; Sa'adi *et al.*, 2019). Trend analysis can be used to investigate whether there is any evidence of an increase in river floods in the observational river discharge data. Such analysis requires long records (i.e., more than 30 years) not only to distinguish climate variability explicitly from climate change induced trends (Svensson, Kundzewicz and Maurer, 2005; Vogel, Yaindl and Walter, 2011), but also to incorporate the impacts of human induced activities such as deforestation and water management practices (e.g., reservoirs and irrigation).

Trend analysis of river flow series have been undertaken at global and regional scales in many parts of the world (see, e.g., Svensson, Kundzewicz and Maurer, 2005; Cunderlik and Ouarda, 2009; Burn, Whitfield and Sharif, 2016; Vormoor *et al.*, 2016; Berghuijs *et al.*, 2017; Mangini *et al.*, 2018; Paprotny *et al.*, 2018; Ávila *et al.*, 2019; Ishak and Rahman, 2019; Zadeh, Burn and O'Brien, 2020). However, there is overgeneralization of trend patterns when considering a larger spatial extent, therefore a need for trend analysis at smaller scales, e.g., country scale (see Wilcox *et al.*, 2018; Giuntoli, Renard and Lang, 2019; Tramblay *et al.*, 2019). Relatively few studies have undertaken river flow trend analysis in Africa (Nka *et al.*, 2015; Diop *et al.*, 2018; Degefu *et al.*, 2019), mainly due to data quality issues which may affect trend detection (Slater and Villarini, 2017).

For Kenya, a few studies have attempted to quantify trends in streamflow (e.g., Mwangi *et al.*, 2016; Langat, Kumar and Koech, 2017; Cheruiyot, Gathuru and Koske, 2018). The studies mentioned here examined at the trends and variability in rainfall in Kenya and did not look at the changes in its frequency and how this varies across different catchments in Kenya. Despite its practical significance, little is known about the temporal characteristics of streamflow and these studies considered only trends in the annual maximum flow' and only for one catchment: little has been done to quantify trends in both annual floods and peak over threshold (frequency and magnitude). Those studies which have incorporated the frequency and magnitude of floods were focused on single river basins, incorporating only a single station, e.g., Tana River (Langat, Kumar and Koech, 2017), Malewa river (Nyokabi, Wambua and Okwany, 2021), and Naivasha (Kyambia and Mutua, 2015). These studies showed that flow series for these stations had a statistically significant upward monotonic trend and seasonal variability, indicating that the streamflow regime had changed significantly.

Flood trend analysis looks at trends in the annual maximum river discharge (AMAX), i.e., a one value per year flood series (Kundzewicz *et al*. 2004; 2005). The advantage of this strategy is that the events chosen in two consecutive years are independent. However, the AMAX technique ignores flood occurrences that are less than the annual maximum in each year but are nevertheless significant for society, particularly in terms of losses, and perhaps inappropriate for climates with two distinct rainy seasons. The Peak Over Threshold (POT) technique (Burn, Whitfield and Sharif, 2016; Mangini *et al.*, 2018) selects all floods over a specific threshold that occur throughout a flow record. This makes it possible for a trend in the frequency (counts) of floods rather than merely their magnitude to be estimated (Svensson, Kundzewicz and Maurer, 2005). No study looking into trends in POT series has focused on

Kenya, even though there has been an increase in the frequency of reported flood events across the country.

Understanding variability in flood timing and seasonality is important for water resources planning and management (Stephens *et al.*, 2015). Changes in the timing of the yearly flood have far-reaching consequences for flood-based farming systems, especially for the livelihoods of people who adjust their floodplain management and agricultural activities to the rise and fall of the flood wave (Ficchì and Stephens, 2019). Significant changes in flood timing have been found around the world, as shown in studies by Cunderlik and Ouarda (2009), Burn, Whitfield and Sharif (2016) and Ficchì and Stephens (2019). There has been an overall shift in flood timing in East Africa in recent decades (Ficchì and Stephens, 2019), but there has been no detailed study for catchments across Kenya.

Hence, there is a notable gap in the literature associated with temporal characteristics of streamflow data in terms of frequency and magnitude in both AMAX and POT as well as the country scale flood seasonality and timing in Kenya, and this gap was an important motivation for the present study. The objective of this paper is to detect the evidence of statistically significant flood trends in observations and the GR4J model simulated discharge across Kenya, where flood-based farming systems and livelihoods are extensive. High flow indices are derived from river flow discharge series, with AMAX and POT indices using different magnitude and frequency thresholds. First high flow indices were derived and subjected to a trend test. Then a sensitivity analysis of the detected trends was performed for different flood peak selection criteria. Next, the country-scale seasonality in flooding was characterised. Finally, the changes in the timing of the annual floods were established. The following three research questions are addressed:

1. What are the trends in the observed flood magnitude and frequency across Kenya for the period 1981–2016?

2. What are the trends in the GR4J model simulated discharge across Kenya for the period of 1981 – 2016. How do they compare with the trends in the observations?

3. What is the sensitivity of the detected trends to the selection criterion used to define different flood peak series?

4. What are the characteristics of country scale seasonality in flooding?

5. What are the changes in seasonality and timing of the annual floods across Kenya?

6.4. Study area

The study is undertaken across Kenyan catchments at 19 river gauging stations (Figure 6.1) with varying characteristics (Table 14). These were selected due to the frequency and magnitude of the impacts of floods, as well as the availability of river flow observations (Table 1). Kenya exhibits high variability in physiographic and hydroclimatic conditions. The highest point is at about 5,000m a.s.l. (mostly areas around the central highlands) while the lowest point is about 20m a.s.l. (mainly around coastal areas, Figure 4.4). The vegetation cover is mainly a mixture of tree cover, grass, and sparse vegetation in most parts of the country, with shrubs and bare land in the arid and semi-arid areas of northern Kenya (Figure 4.5). As a result,

Kenya experiences different climate related extremes in terms of intensity, magnitude, and timing.

Rainfall pattern follows a bimodal rainfall seasonality (Ongoma and Chen, 2017) with high spatiotemporal variability (Hession & Moore, 2011). Three seasons are experienced: the 'long rains' of March - - April - May (MAM), the non-rainy months of June - July - August (JJA), and the 'short rains' of October - November - December (OND) (Ogallo, 1988; Ongoma *et al.*, 2015) (Figure 4.3). About 42% of the total annual rainfall is observed during the MAM rainfall season (Ongoma and Chen, 2017), with the highest intensity observed near the water bodies of the Indian Ocean, Lake Victoria, and the Kenyan highlands.
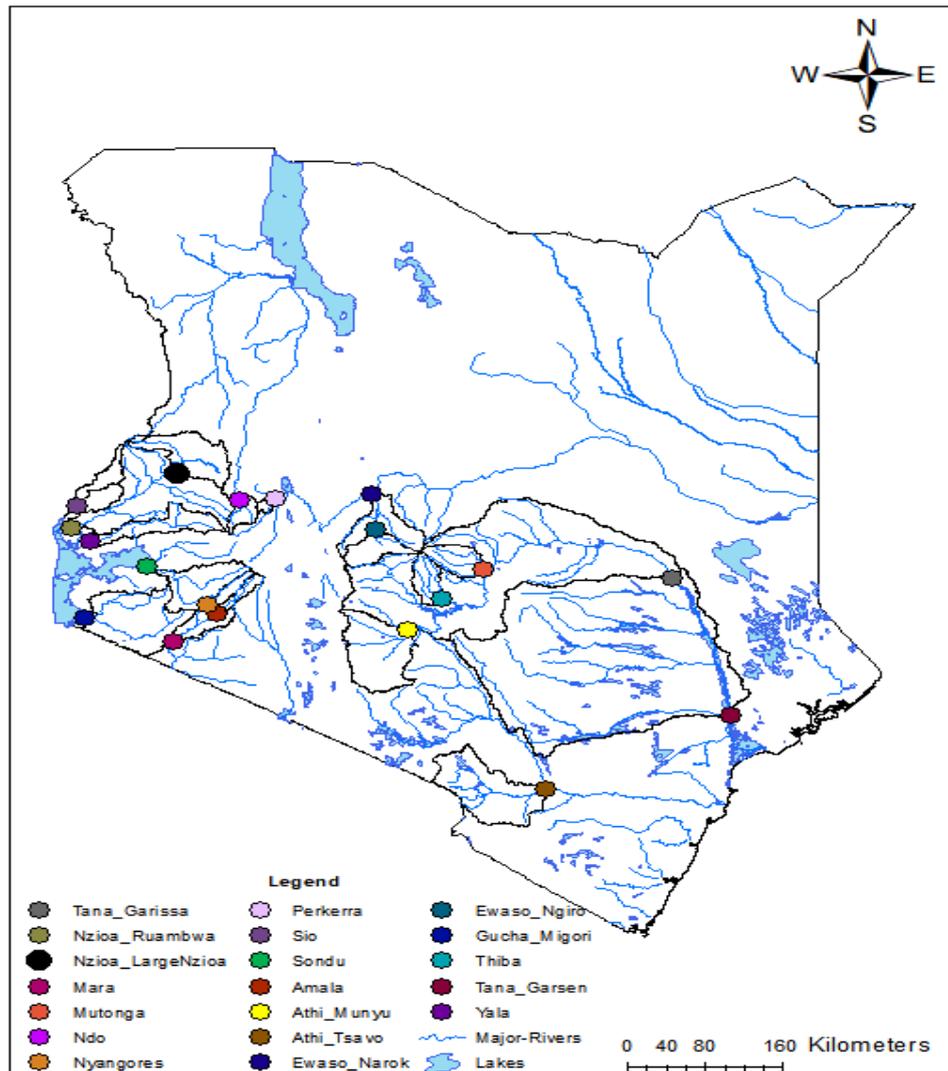
Figure 6.1: Location of the 19 river gauging stations located in the Kenyan catchments studied.

6.5. Data and methods

6.5.1. Observed and simulated discharge data

For the study, observed daily river flow data and the GRJ4 model simulated discharge obtained from the modelling experiment in Chapter 5 from 19 sites distributed across the country were employed (Table 14, Figure 6.1). A common data period of 1981 – 2016 was adopted for the analysis. Two of the stations (Ndo and Perkerra) did not have the most recent

data records and only 13 and 21 years of data were available respectively, which may possibly affect the shifts in the trend patterns in the two stations.

Six distinct high flow indices were derived from the daily mean river discharge data for the 19 stations, and then subjected to the Mann-Kendall test to check for trends. The magnitude of the trend slope was also determined for each index for each station. The observed discharge data used in this study do not incorporate in-depth metadata such as the water abstraction data including dam construction and irrigation activities, so the trends may include the effects of changes in flow attenuation (e.g., reservoirs) and land use as well as any impacts of climate change. Consequently, the GR4J model simulated discharge from Chapter 5 modelling experiment was used to extract the simulated AMAX and the trends tests carried out to compare with the trends in the observations.

6.5.2. Flow indices

The features of high flow regimes were described using six distinct indices. First, the yearly maximum daily mean river flow (AMAX) was derived. The second index is peak over threshold (POT). Three peak magnitudes were tested as thresholds with the size of the criterion established so that, on average, one, three and five POT occurrences were chosen every year (Mangini *et al*., 2019): POT1mag; POT3mag and POT5mag. If the peaks in a POT series were separated by a specific time, they were deemed independent of each other. After an inspection of the flow series, a declustering method was employed that involved the introduction of a runs parameter which used a 7day separation time interval between the peaks as most of the catchments are less than 40,000km2 (Collet, Beevers and Prudhomme, 2017) and thus had a concentration time of less than 7 days. Counting the number of POTs that occur each year can be used to quantify flood frequency (annual counts), and thus three flood frequency indices were derived corresponding to the POT magnitudes: POT1freq, POT3freq and POT5freq.The three POT1, POT3 and POT5 series represent the magnitude and frequency of the most severe, moderate, and minor floods respectively. In this work, the sensitivity of trend results to this threshold selection is evaluated and trends in the POT1 flood magnitude data are compared to those of AMAX.

A multiple index (MI) was adopted from Mangini *et al.,* (2019). Multiple Index (MI) shows the ratio of the mean discharge magnitude to the maximum/peak discharge magnitude of a flood sequence, as well as the average yearly outflow for specific stations, which can be used to show differences in hydrological flood characteristics. Thus, this is useful in describing the varied flood regimes across Kenya's different hydroclimatic regions (e.g., wet, arid and semi-arid), as well as the effects of human influence such as dam construction and irrigation activities on the catchments' water balance. The MI is the ratio of the mean discharge of a flood series to the peak annual flow at each individual station and can be expressed as in Equation **6.1**:

Equation 6.1:- Multiple Index.

$MI = Q_F/Q_A$

Q_F represents the mean discharge value of the flow series and Q_A is the peak discharge recorded for an individual station. A significant divergence from the mean flow is indicated by higher MI values; for smaller exceedance thresholds, low MI values would be obtained.

.

Table 14: Characteristics of the Kenyan catchments studied, human influence, daily mean river flow gauge, and data availability information. Missing data are expressed as a percentage of the available period.

River Name	Catchment Outlet point	Station ID	Lon	Lat	Drainage Area (km²)	Mean Elevation (m.a.s.l)	Mean Annual Rainfall (mm)	Annual Discharge (m³s⁻¹)	Catchment Characteristics	Human Influence		First & Last year of record	Record length (years)	Amount missing (%)
										Dams	Irrigation schemes			
Tana	Tana Garsen	4G02	40.11	-2.28	80 760	720	672	135.8	Semi-arid plains	9	11	1981-2016	36	58.2
Tana	Tana Garissa	4G01	39.7	-0.45	32 695	870	868	169.3	Highlands on the upstream & semi-arid plains in the lowlands	8	7	1981-2018	38	14.2
Nzioa	RuambwaBridge	1EF01	34.09	0.12	12 643	1740	1488	151.2	Dense forest cover (highlands) & low trees & bushes (lower reaches)	2	4	1981-2018	38	13.6
Galana	Galana Tsavo	3G02	38.47	-2.99	6560	930	628	3.3	Semi-arid savannah plains	3	1	1981-2015	35	59.6
Gucha	Gucha Migori	1KB05	34.21	-0.95	6 310	1650	1435	45.0	Eastern lowlands with dense vegetation cover	0	2	1981-2015	35	47.8
Athi	Athi Munyu	3DA02	37.19	-1.09	5689	1730	822	18.8	Highlands & forest cover	3	1	1981-2017	37	21.6
Nzioa	Large Nzioa	1BD02	35.06	0.76	3878	1720	1267	15.3	Dense forest cover	1	1	1981-2011	31	28.8
Sondu	Sondu Miriu	1JG04	34.80	-0.33	3444	2017	1614	53.9	Low lying plains (western) &	2	2	1981-2018	38	64.4

									highland (eastern)					
Mara	Mara	1LA04	35.04	-1.23	2977	2100	1262	11.8	Low lying shrubs, semi-arid	0	1	1981-2015	35	77.7
Yala	Yala	1FG02	34.27	0.04	2700		1696	40.8	Swampy	0	0	1981-2019	39	59.6
Ewaso	Ewaso Narok	5AC10	36.73	0.43	2597	1600	880	5.3	Low lying shrubs & mainly semi-arid	0	2	1981-2018	38	26.5
Tana	Mutonga	4EA07	37.89	-0.38	1867	1830	1427	35.5	Highlands and forest cover	0	1	1981-2016	36	44.2
Ewaso	Ewaso Ngiro	5BC04	36.91	0.09	1837	1700	972	20.6	Low lying shrubs & mainly semi-arid	0	0	1981-2019	39	35.0
Sio	Sio	2EE07A	34.14	0.39	1011	1650	1822	15.5	Low trees & bushes & swampy in lower reaches	0	1	1981-2018	38	18.1
Turkwel	Ndo	2C07	35.65	0.45	897	1133	1371	9.1	Extensive palaeo-floodplain & arid conditions	0	1	1981-1993	13	47.2
Mara	Amala	1LB02	35.44	-0.89	695	2100	1377	6.8	Low lying shrubs, semi-arid	0	0	1981-2017	37	25.6
Mara	Nyangores	1LA03	35.35	-0.79	692	2008	1262	11.8	Semi-arid savannah plains, low lying shrubs, semi-arid	0	0	1981-2017	37	15.5
Turkwel	Perkerra	2EE07A	35.97	0.46	371	1023	832	5.7	Extensive palaeo-floodplain & arid conditions	1	1	1985-2005	21	50.1

6.5.3. Trend detection in AMAX and POT series

The non-parametric Mann Kendall's (MK) test (Mann, 1945; Kendall, 1975) was used to detect trends in AMAX and POT*mag* flood series. The modified MK test was applied to test for trend: this incorporates Yue and Wang's (2004) variance correction approach. The effective sample size was calculated using serial correlation coefficients for all lags and the slope magnitude was estimated using the Thiel-Sen slope algorithm (Sen, 1968).

Trends in POT$_{freq}$ were estimated using the Chi-Squared test with parametric Poisson regression because, unlike the non-parametric Mann Kendall test, it accounts for the hierarchical count series that may contain several paired values (Vormoor *et al.*, 2016; Mangini *et al.*, 2018). A two-tailed trend test was applied at 10% significance level to test for the statistically significant trends. This was required because the direction of the trends to be tested was unknown.

6.5.4. Sensitivity analysis of trends to threshold selection

The number of peaks considered in each flood series are affected by the exceedance threshold (λ), where POT1 (λ =1), yields the highest floods recorded at each of the stations and POT*n* (λ =*n*) yields n flood events in the series. This means that selecting a higher λ would result in a lower threshold, thus yielding higher number of flood peaks in the series. Different thresholds (λ =1, 3 and 5) corresponding to POT1, POT3 POT5 flood series were derived and the sensitivity of trends to different POT series leading to the selection of different thresholds (λ) analysed.

Threshold selection was aided by the creation of different plots, such as a mean residual life plot, that can aid in the determination of a suitable threshold level (Burn, Whitfield and Sharif, 2016). The mean residual life (MRL) plot is a plot of the mean flood excess over a given threshold versus a range of threshold values.

6.5.5. Flood seasonality

Seasonality measures (Parajka *et al.*, 2010) are used to characterize the timing and variability of extreme flood events. These are defined by directional statistics (Mardia, 1975). The date of occurrence of a flood event is defined as a directional statistic through conversion of the Julian date of the occurrence of an event to an angular value (Berghuijs *et al.*, 2019; Ficchì and Stephens, 2019), where January 1 is Day 1 and December 31 is Day 365 of the flood occurrence for the event *i* following Equation **3**.**54**. The subsequent equations of seasonality measures are outlined in Section 3.3.3.3 of Chapter 3 of this thesis.

6.6. Results

6.6.1. Frequency of peak events in the flood series

The mean number of discharge peaks per year varies spatially across Kenyan catchments from about 4 to 13 events (Figure 6.2). The highest numbers of flood events per year are recorded at Perkerra, Mara and Garsen stations. The notably high number of events at Perkerra may be due to the shorter data series (a large percentage of data is missing, as outlined in Table 14). However, the Garsen and Mara stations have relatively good data records and are mainly arid and semi-arid, therefore we can be more confident that the high numbers of events are likely

to be due to the sporadic torrential rains that occur during the rainy months. The mean number of flood events per year for the other stations ranges between 6 and 8.
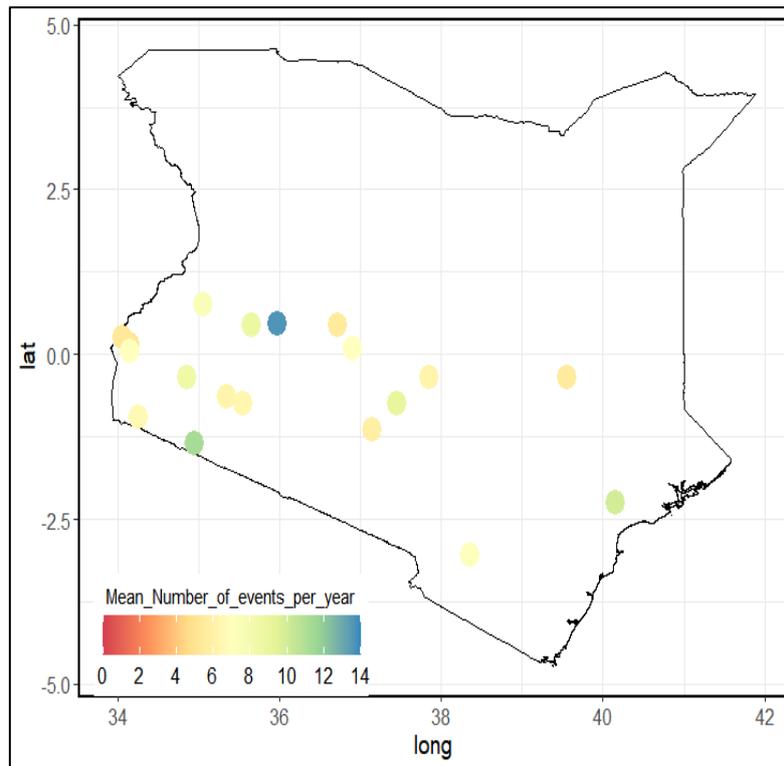
Figure 6.2: Mean number of independent flood peaks per year at the 19 gauging stations in Kenya.

6.6.2. Trends in the observed AMAX and POT1 flood series

There is considerable heterogeneity in the trend results calculated for the 36-year research period (Error! Reference source not found.**Error! Reference source not found.**). However, a positive trend across twelve stations dominates Kenyan stations in the AMAX flood series, only three stations (Nyangores, Ewaso-Ngiro and Gucha-Migori) show negative trends and four (Mara, Miriu, Thiba and Mutonga) do not show distinct trend direction.

Statistically significant positive and negative trends at 10% confidence interval in the AMAX series are detected in 12 out of 19 stations (Figure **6.3**). The significant positive trends are dominant at Garissa, Garsen, Mutonga, Athi-Munyu, Athi-Tsavo, Sondu, Perkerra, Sio and Ruambwa stations. The significant negative trends are dominant at Gucha-Migori, Nyangores and Ewaso-Ngiro stations. Only three stations (Perkerra, Ewaso-Narok and Ndo) do not show trends in the AMAX series. Interestingly, Pekerra has the highest mean number of independent discharge peaks but no visible trends in the peaks. However, in general trends in the same direction can be seen for stations in proximity to each other, which is evidence of a spatially coherent pattern.
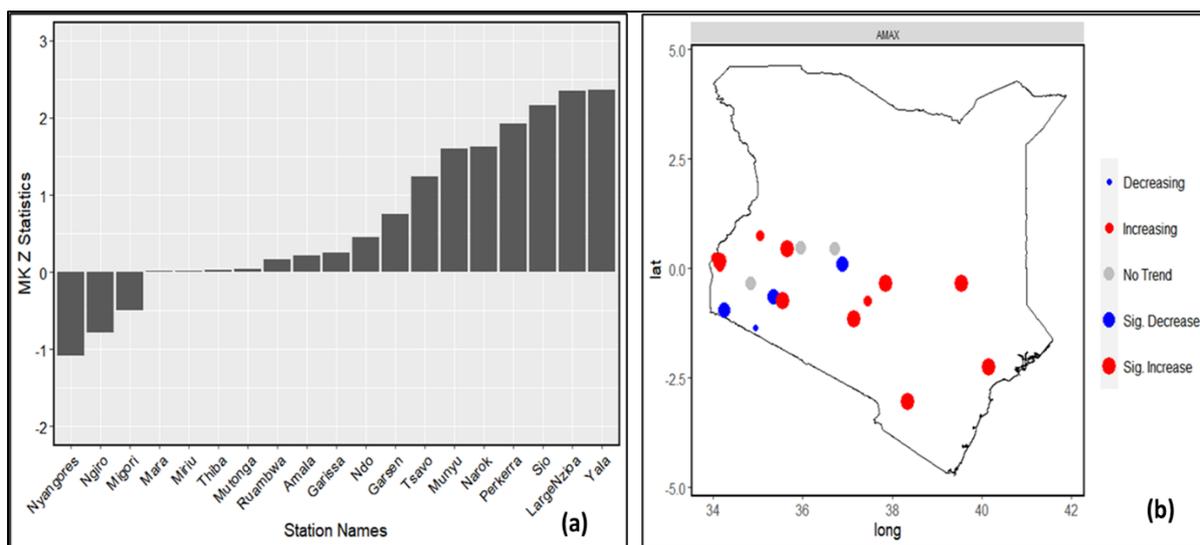
Figure 6.3:- Mann–Kendall (MK) test Zs statistics (a) and trends (b) in the observed annual maximum flood series (AMAX) for the period 1981 -2016. In (b), filled circular symbols indicate the direction of the trend slope with positive (red), negative (blue) and grey (no trend) trends at 10% significance level. The size of the circles indicates the statistical significance of the trends.

Positive trends in POT1 flood magnitude are found for 11 stations and negative trends for 8 stations respectively (Figure 6.4 right). For flood frequency only three stations (LargeNzioa, Munyu and Migori) show negative trends for POT1 (Figure 6.4 left), whereas the rest of the stations show positive trends. Garsen, Yala, Narok, Perkerra, Miriu, Sio, Nyangores and Garissa stations show increasing positive trends, and Thiba, Munyu, and LargeNzioa show decreasing trends both in flood frequency and magnitude.

The total number of stations exhibiting statistically significant trends is higher in the POT1mag than in AMAX (Figure 6.5). There is a consistent pattern in the flood change in some catchments depicting significantly increasing trends in both the MAX and POT1mag series: Garissa, Athi- Munyu, Athi- Tsavo, Ruambwa, Sio, Mutonga etc. Only one station (Ewaso – Ngiro) has a significantly decreasing trend both in the AMAX and POT1mag flood series, whereas Gucha-Migori has a significantly increasing and decreasing trend in the AMAX and POT1mag respectively.

POT1freq flood series show statistically significant trends in 8 out of the 19 stations, predominantly in stations in the western part of the country Figure 6.5 left. This suggests a clear pattern in that the trends in flood frequency are more significant in the western parts of Kenya than in the rest of the country, except for the Garissa and Garsen stations.

146
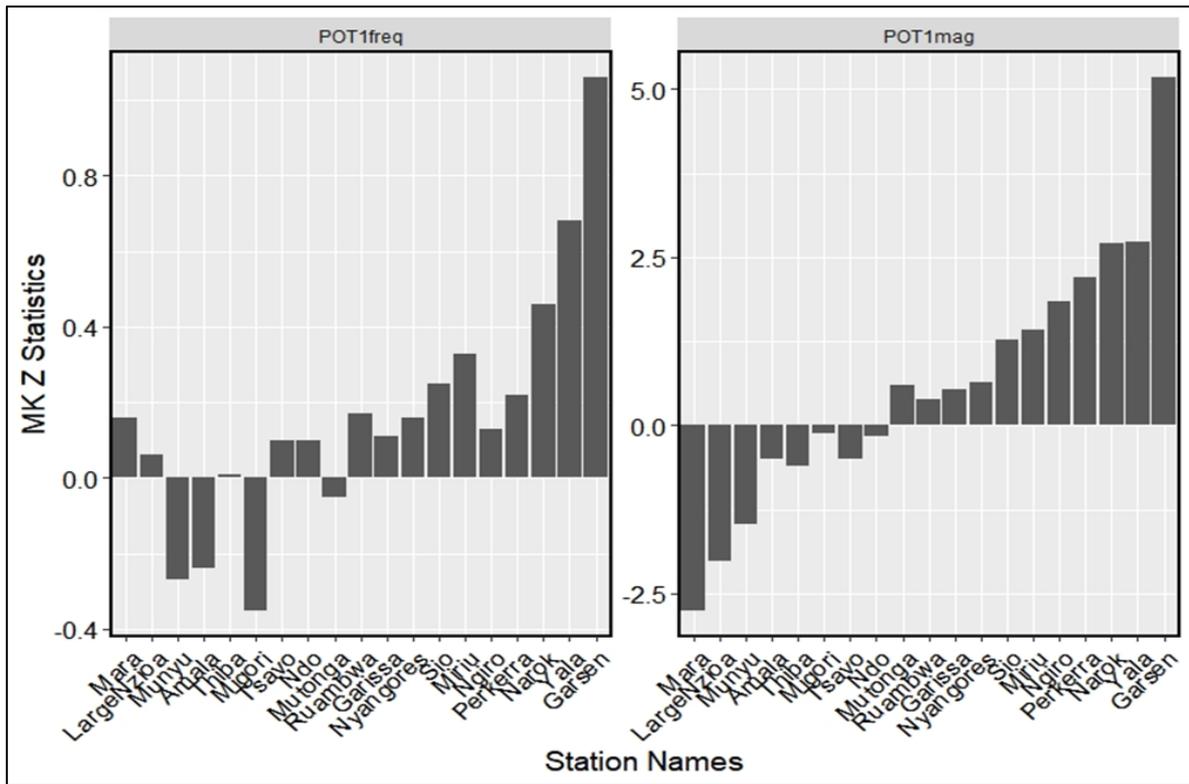
Figure 6.4: Mann–Kendall (MK) test Zs statistics in the POT1freq (left) and POT1mag (right) flood series across the 19 stations studied for the for the period 1981- 2016.
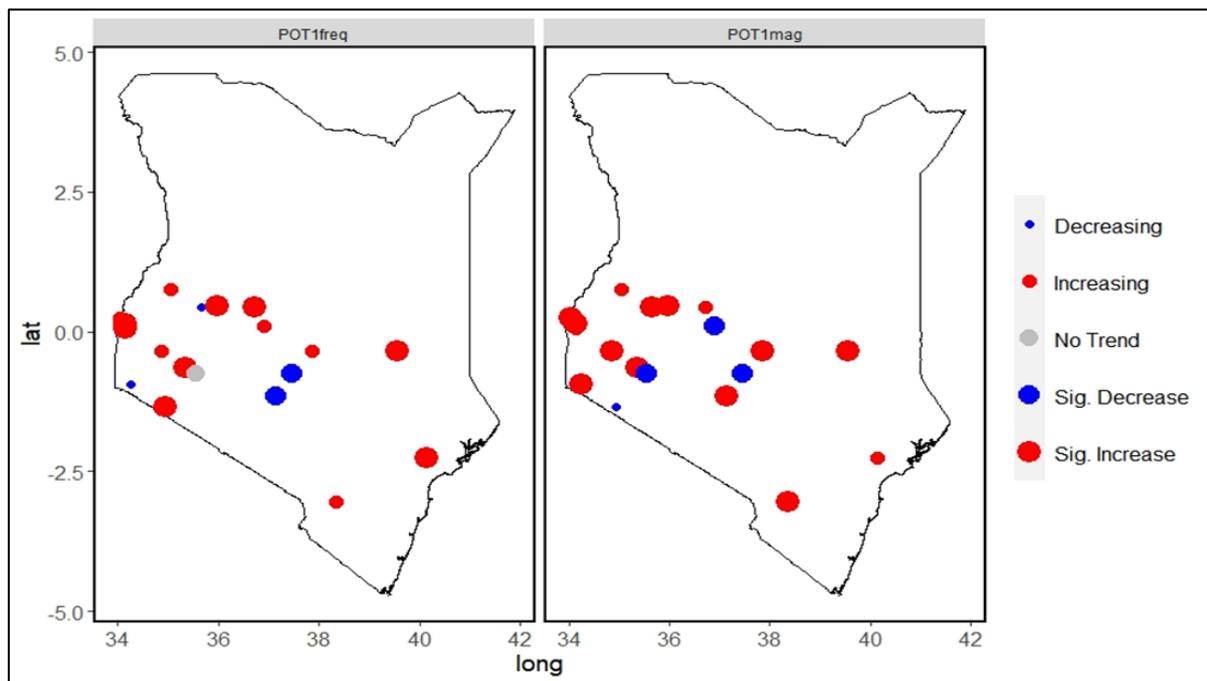
Figure 6.5: Trends in the POT1freq (left) and POT1mag (right) for the period 1981 -2016. Filled circular symbols indicate direction of the trend slope, significant positive (red), negative (blue) and grey (no trend) trends at 10% significance level. The size of the symbol indicates statistically significant trends.

6.6.3. Sensitivity of flood trends to the selection of different flood exceedances thresholds

There are notable differences in trends in some catchments for the AMAX and POT1mag flood series, which points to the effects of selecting exceedance thresholds in deriving POT series. The higher exceedance threshold ($\lambda = 1$) gives the most extreme flood events, with an average of one event per year, whereas the lower exceedance threshold ($\lambda = 5$) gives a lower threshold, thus an increase in the number of (smaller) peak events (Figure 6.7).

MIs from the AMAX are the most significant, followed by the MI of the POT1mag series as well as the remaining POT series, whereas smallest MIs are derived for the POT6mag flood series (Figure 6.6a).MI values are relatively higher in the small sized arid and the semi-arid catchments such as Ewaso-Ngiro, Ewaso-Narok, Perkerra and Mara (Figure 6.6b). The arid and semi-arid climate of these catchments causes generally low mean discharge across the year which is offset by intense precipitation events leading to high discharge values. Additionally, small catchments are likely to show an intrinsic high variance in the daily hydrograph, thus producing the high MI values. Mutonga and Thiba show the highest MI values. This could be attributed to the large dam releases into the rivers resulting in higher mean discharge, especially during rainy months. In general, catchments in the western part of the country have medium MI values due to rain falling all year round. This is because rainfall falling outside the pattern of typical seasonal rainfall contributes substantially to the mean annual rainfall.

There is a high sensitivity in the results of the trend analysis to different exceedance thresholds in terms of both magnitude and frequency (Figure 6.8). Half of the stations show significant positive and negative trends for high λ. More than half of the stations show significant positive trends for $\lambda = 3$ and $\lambda = 4$. However, there is no defined pattern in sensitivities (increasing with increase in λ or decreasing with a decrease in λ) across most of the stations for flood magnitude (Figure 6.8a). Trend results show high sensitivity to thresholds for flood frequency series as seen in Figure 6.8b. There are notably significant trends detected in the POTfreq series across most of the stations for increasing values of λ. Half of the stations depict significant increasing positive and negative trends with clear negative trends in Ewaso-Narok, Perkerra and Gucha-Migori. for threshold values of $\lambda = 3$ and $\lambda = 5$. There is a clear pattern of significantly increasing trends in flood frequency across all the stations.

6.6.4. Flood timing and variability

The occurrence of the annual flood in most of the stations is around the months of March, April, May, November and December (Figure 6.9 panel a). These coincide with the occurrence of the 'long rains' (March – May) and the 'short rains' (October – December). Some stations, such as Ndo, Perkerra, Mara, Ewaso- Ngiro and Ewaso-Narok, have a mean date of occurrence around the months of June, July and August.

Predictability is high ($r > 0.4$) in most of the stations whose annual floods occur during the short and long rains, whereas values of $r < 0.4$ can be seen in the stations in which the flood timing is around the dry months of June, July and August (Figure 6.9 panel b). There is a consistent spatial pattern of predictability: stations in the western (Nzio, Sio, Yala, LargeNzioa), central (Thiba, Mutonga, Munyu), and coastal regions (Tsavo, Garsen, Garissa) show a high tendency to predictability, whereas stations in the Rift valley region (Mara, Amala, Nyongores, Ndo and Perkerra), have a low tendency to predictability.
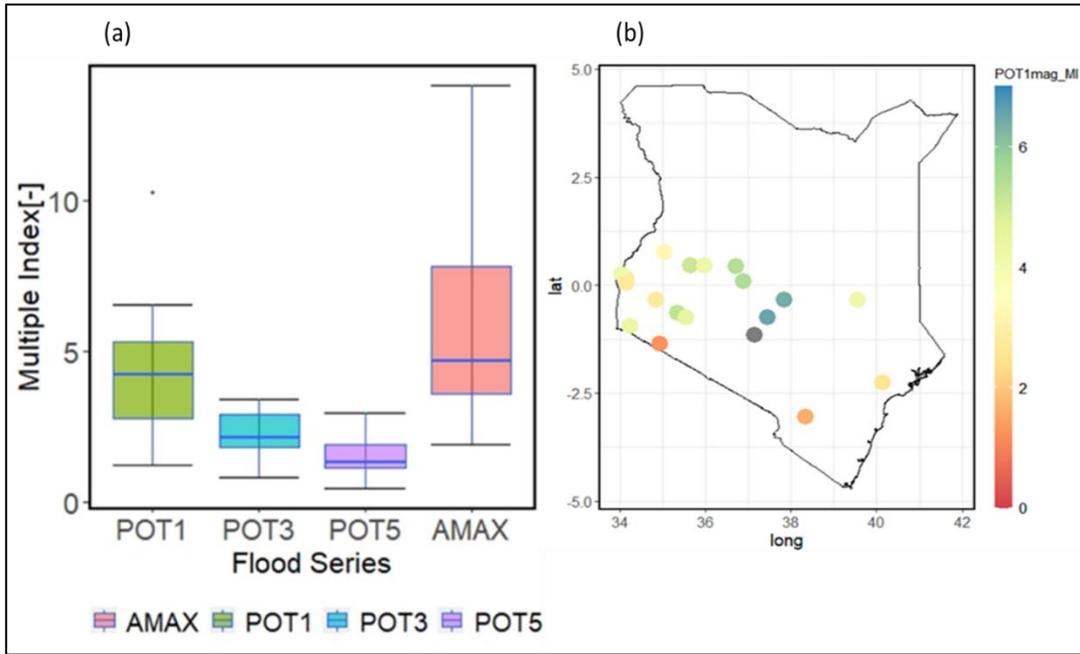
Figure 6.6:- (a) Box plots for the Multiple Index (MI) for the different peak over thresholds (POT) flood series and the annual maximum (AMAX) for all the 19 stations. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.The mean number of events per year for POT thresholds (λ) considered are one (POT1), three (POT3) and five (POT5). (b) MI for POT1mag flood magnitude at the each of the 19 gauging stations across Kenya.
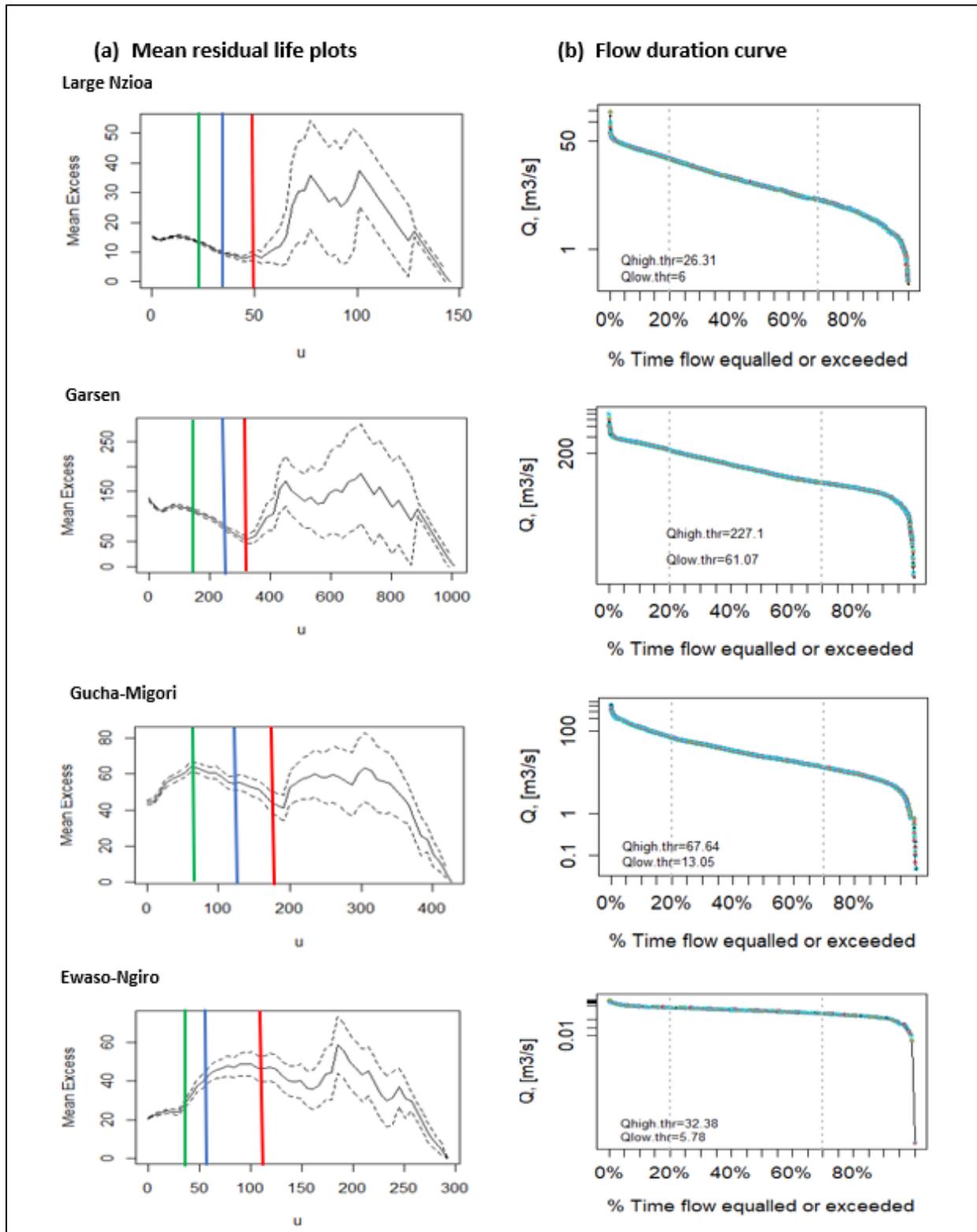
Figure 6.7: Visual plots for exceedance threshold selection using (a) mean residual life plots (left panel) and (b) flow duration curves for selected study stations (right panel). The vertical lines show different thresholds (λ =1 for red, λ =3 for blue and λ =5 for green and u represents a range of discharge threshold values).
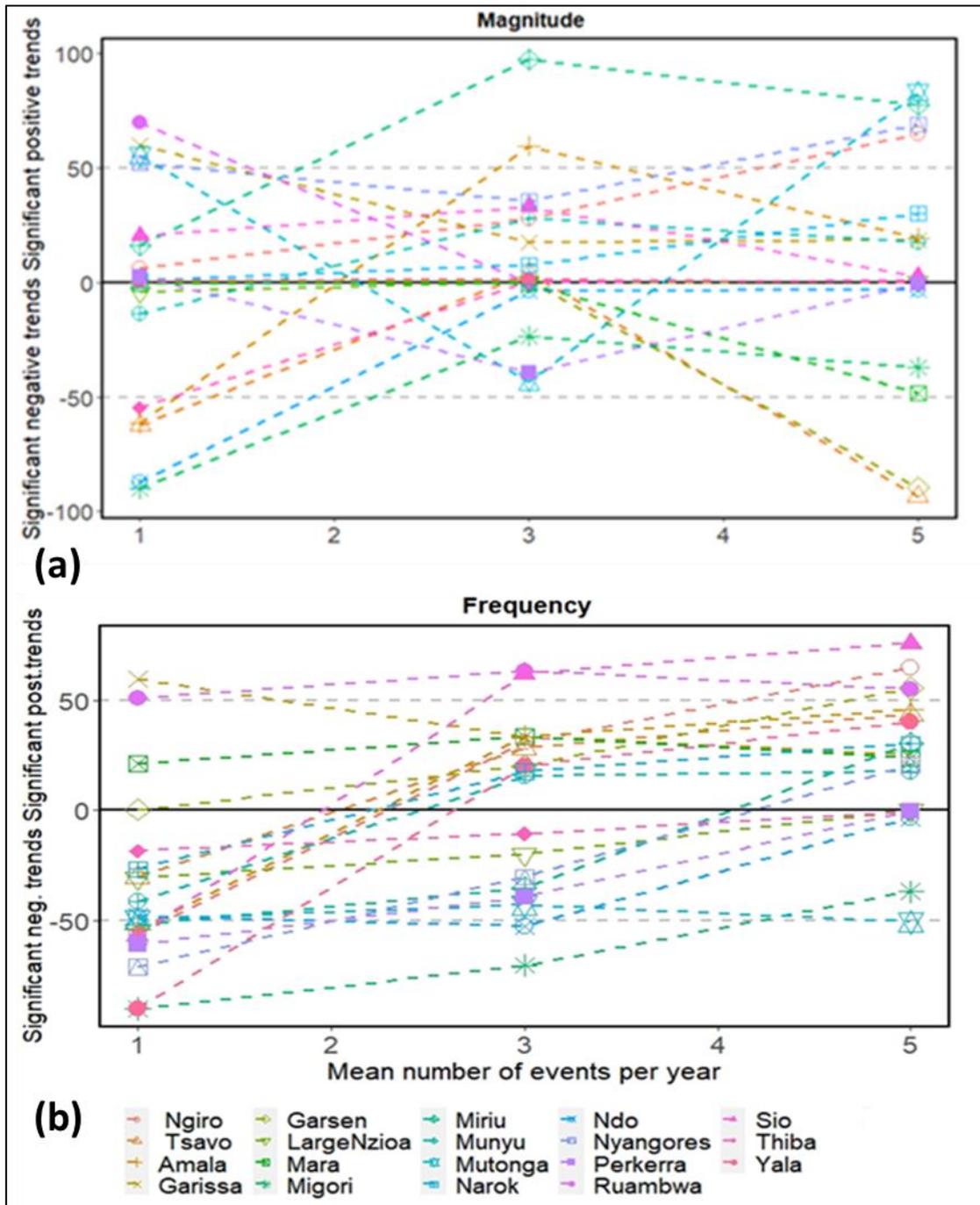
Figure 6.8: The sensitivity of trends in the POT series to the selection of different exceedances thresholds (λ) for different mean number of floods per year for POT1, POT3 and POT5 (a) magnitude and (b) frequency. Significant trends (expressed as percentages) in flood magnitude at 10% significance level with a threshold level λ range of 1, 3 and 5 mean events per year Different colours and symbols represent the different 19 stations considered in the study.
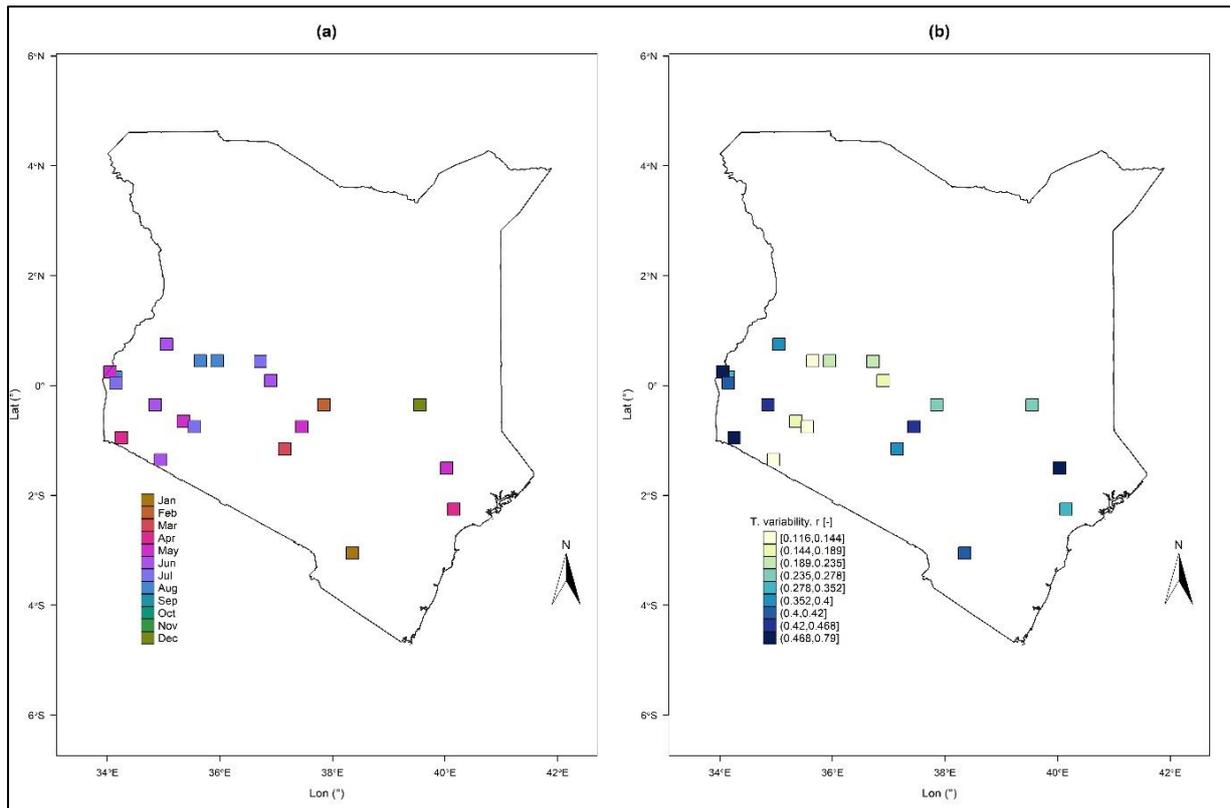
Figure 6.9:- Annual flood seasonality measure for Kenya stream-gauge data. (a) Average flood timing (θ) and (b) Interannual variability of flood timing (r [−]) shown by different colour saturation (the higher the variability is, the lighter the saturation is). The filled rectangular dots are reported at each of the 19 river gauging stations.

6.6.5. Trends in the GR4J model simulated AMAX series

There is considerably distinct trend directions in the simulated AMAX trend across the 19 catchments (Figure 6.10 a). Like in the observed AMAX, simulated MK-Z statistics show dominance in the positive trends compared to the negative trends across the stations for period 1981-2016. A positive trend is however observed across thirteen stations in GR4J simulated flood series and only 5 stations (Perkerra, Thiba, Garsen Tsavi and Amala) show negative trends. The stations depicting downward trends (negative) in the simulated flood series are different from those in the observed flood series (see Figure

Statistically significant positive and negative trends at 10% confidence interval in the simulated AMAX series are detected in 12 out of 19 stations (Figure 6.10 b). Significant positive trends dominate stations in the western catchments, particularly Nzioa Ruambwa, Large Nzioa, Miriu, Migori, Sio, Amala, Ndo and one in central and northern parts (Thiba and Ngiro) respectively. The significant negative trends are dominant are depicted in only three stations i.e., Thiba, Amala and Sio. Unlike in observations, Garissa and Garsen stations show non-significant decreasing and increasing trends respectively. Nyangores, Mutonga, Munyu and Narok do not show visible trends in the peaks, as seen in the observed AMAX series. However, a similar spatial pattern of trends in the same direction can be seen for stations in proximity to each other, which is evidence of a spatially coherent pattern and the linearity in the model simulations for catchments within the same physiographic settings.
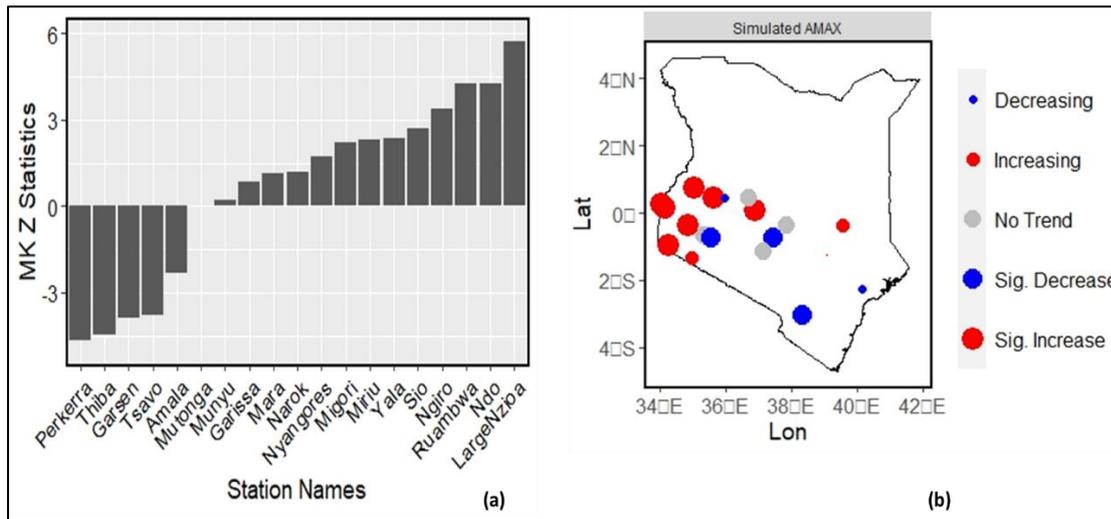
Figure 6.10: Mann–Kendall (MK) test Zs statistics (a) and trends (b) in the simulated annual maximum flood series (AMAX) for the period 1981 -2016. In (b), filled circular symbols indicate the direction of the trend slope with positive (red), negative (blue) and grey (no trend) trends at 10% significance level. The size of the circles indicates the statistical significance of the trends.

6.7.   Discussion

6.7.1. Trends in observed AMAX and POT1 series

In this study, statistically significant trends are detected in the AMAX and POT flow series in most of the 19 stations across the country. with only three showing significantly decreasing trends in the AMAX. The decreasing trends at Nyangores, Ewaso Ngiro and Migori stations are mainly due to the high sensitivity of trend detection in the flow series used in this study. The available data do not include the most recent years as compared to other stations and this may fail to capture the most recent floods in those stations. The trends are consistent spatially for stations on Tana River (Thiba, Garissa and Garsen), Athi (Tsavo, Munyu) and Nzioa (Ruambwa). This may be attributed to the increasing frequency of rainfall events in these parts of the country (Huho and Kosonei, 2014; Wainwright *et al.*, 2021).

The significant trend in both AMAX and POT1*mag* varies from one station to another. For instance, there is a general tendency of increasing trend in flood frequency and magnitude at Garissa, Athi- Munyu, Athi- Tsavo, Ruambwa, Sio and Mutonga, while Gucha-Migori, Nyangores and Ewaso-Ngiro show decreasing trends. Trend patterns at Garissa, Ruambwa and Sio agree with the findings of Nasambu *et al.* (2018) and Langat *et al*. (2020), who found significantly increasing trends due to increased frequency of rainfall events at these stations. However, the above results also point to a mixed pattern in trends in stations that may be from the same region. These trend patterns are important in indicating the existence of different flood drivers in those regions. For example, Ruambwa and Gucha are both located in the western part of Kenya, which receives rainfall all year round, but they have contrasting trend patterns. Thus, there is a lot of variability within the trends and their significance, and there is often a difference between magnitude and frequency. This finding is like those of studies from around the world, underlining the complexity of making regional generalisations about trends (see, e.g., Mangini *et al*., 2019; Arheimer and Lindström, 2015). Garsen station at the downstream point on Tana River has a significantly increasing trend in flood frequency but not in magnitude, whereas Mara and Amala have decreasing trends in two indices. This may be

153

mainly because rivers in semi-arid environments, such as the Tana and Mara, are particularly vulnerable to fluctuations in water availability caused by decreases in rainfall, increases in water withdrawals and changes in seasonal flows (Langat *et al.,* 2020). Tana, for example, is heavily affected by human activities such as irrigation schemes, domestic consumption, and hydropower generation (Langat, Kumar and Koech, 2017; Langat *et al.*, 2019).

In addition, both the Tana and Mara are classified as water limited catchments with a dryness index of 1.1 annually. This means that there is an imbalance in water demand and any increase in the rainfall is accompanied by higher temperatures resulting in high evapotranspiration, thus increasing the dryness index further. Therefore, most of the rainfall in these catchments is used up in evapotranspiration (Mwangi *et al.*, 2016).

6.7.2. Comparison of trends in observed and simulated AMAX series

There is heterogeneity in the number of stations showing increasing and decreasing trends across the Kenyan catchments. Simulated AMAX shows 13, whereas observed flood series shows 12, with increasing trends and 5 and 3 with decreasing trends respectively. A distinct slope direction in the MK-Z statistics is seen in the simulated as compared to those in observations, which shows less distinct slope directions in 5 stations (Mara, Miriu, Thiba and Munyu). However, there are distinct reverse trends (in terms of direction) in some of the catchments, although in catchments located in different hydoclimatic and physiographic settings. For example, in observations, Nyangores, Ngiro and Migori show decreasing trends whereas in simulated, they show increasing trends. In simulated discharge, Garsen, Tsavo and Perkerra show decreasing, whereas in observations they show the increasing trends. The later catchments are in the ASALs of Kenya and Perkerra is relatively small in terms of catchment area, likely to be affected by small and sporadic rains. These results indicate that small differences in amounts or shifts in time or space of modelled rainfall, in comparison with observed precipitation, can strongly modify the hydrologic response of small watersheds to extreme events (Camera *et al.,* 2020).

In terms of the significance of the trends, there is notable similarity in the trend pattern for most catchments in the western parts of Kenya, with, Nzioa Ruambwa, Large Nzioa, Migori, Sio Miriu, depicting significantly increasing trends whereas some stations in the central highlands show no trends and mixed trends in the southern and coastal catchments in the simulated AMAX flood series. However reverse trends are observed for Garsen Tsavo and Garissa station, in which observations shows significant increasing trends and the model simulations show decreasing trends. This may be due to difficulties in the GR4J model simulating the water balance component as a result of the human influence and water abstraction activities such as irrigation and dam construction (Langat *et al.*, 2019) in the upstream areas (see Figure **4.4**), data which was not incorporated in them modelling experiment due to data scarcity challenged outlined in Chapter 1 and 5.

6.7.3. Sensitivity of trends to different POT threshold selection

Flood frequency trends can be seen across all thresholds whereas for POT-M series, with more than one event per year on average, the percentage of stations exhibiting significant trends in flood magnitude is high, and only a few significant positive trends are detected for $\lambda = 1$. Sensitivity of trends to different POT series when different thresholds are selected shows different trend patterns and varies from one station to another. Generally, there is no clear

pattern in trends in POT*mag* series in response to different thresholds. This is because the flood series is highly sensitive to the threshold selection and varies considerably from one station to another. This may also be alluded to non-homogeneity in flood characteristics at different stations across Kenya.

When we considered smaller floods (POT5), we observed a general increasing trend in flood magnitude for stations in western parts, with a clear pattern of decreasing flood magnitude in the southern and coastal stations. This is because western stations are characterised by less intense rainfall all year round, unlike the coastal and southern regions, which receive rain mainly in the rainy months (March, April, May, October, November) and are mostly dry for the other months. These decreasing trends are pronounced at Tsavo and Mara station. The overall spatial pattern of decreasing flood trends in the POT3 series is similar to studies in other regions such as to Mediero *et al.,* (2015).

6.7.4. Flood timing and variability

The timing of the annual flood is assessed using the AMAX flow for the 19 stations across the country. Results show flooding occurs in the peak rainfall months of April and May for the 'long rains' (March – April – May) and in October and November for the 'short rains', mostly in western, central highlands and coastal stations. Our findings are like those in Stephens *et al.* (2015), which showed some extent of correlation of precipitation and floodiness in East Africa by comparing the anomalies in precipitation and floodiness on a monthly basis. However, some stations lying on the same river show varied timings of floods in different months, but the reasons for this are not clear. Tana River and Athi river in the coastal area and Nzioa river in the western areas are examples. The upstream gauging stations on Tana River (Mutonga and Garissa) show the occurrence of annual floods in November and December. The downstream gauging stations (Garsen and Hola) show this occurrence in April and May. The reverse is observed for stations located along the Athi river. One possible reason for this observation may be the high predictability of the short rains (October-November-December) (Kilavi *et al.*, 2018), which is captured in the daily hydrograph of these downstream stations. However, more research is needed to determine the extent of this nonlinearity at different temporal or spatial scales, considering the role of different precipitation periods in flood generation in different regions (Froidevaux *et al.*, 2015). Also, there are stations (e.g., Mara, Amala, Ewaso-Ngiro, Perkerra) showing the occurrence of floods in the dry months of July and August. These stations are in arid and semi-arid areas of the country and are mainly characterised by low flows and any sporadic rainfall falling in those offseason months may lead to a rise in flow, as depicted in the AMAX index. This implies the increased likelihood of flooding events both in rainy and non-rainy seasons, which may result in upward trends in flood frequency in these stations.

There is a high tendency of flood predictability in rivers that flood during peak rainfall months in Kenya, and the reverse is true. For example, flood predictability is high for the stations (Ruambwa, Garissa, Garsen, Munyu) whose flood timing occurs in peak rainfall months and low for those whose floods occur in dry months (Perkerra, Ndo, Ewaso-Ngiro, Mara, Amala). The results also show a defined regional spatial pattern in flood predictability across the country. For example, there is a correspondence in predictability for stations located in the western and costal stations, which are characterised by the annual floods occurring in the peak

rainfall months. The stations located in the Rift Valley region, however, display a clustered pattern with annual floods occurring in dry months.

6.8. Conclusion

Our efforts to identify regional and basin scale trend signals yielded no spatial flood trend pattern, especially for the POT*mag* flood series. In every catchment in Kenya, mixed trends emerged for most of the flood indices used. Mixed and inverse trend signals were also observed between adjacent gauging stations in the same and neighbouring river catchments. The observed complexity of trend signals between adjacent streamflow stations may be attributed to the presence of very complex climate, topography, land cover and land use systems in different parts of the country, which showed great variation at short horizontal distances.

With a few exceptions, such as the trends in flood frequency, this study shows that the number of catchments exhibiting significant trends differs across the country and is not consistent across all flood series. Also, the trend significance in both AMAX and POT1*mag* varies from one station to another. For this reason, it would be more informative to consider trend analysis within larger scale hydro-climatic regions, because trend signals within a region can be considered less sensitive. Additionally, it is therefore necessary to understand such trends, but this requires careful identification of triggers and hydrological processes (Slater and Wilby, 2017; Berghuijs *et al.*, 2019). This is useful in unveiling the degree of this nonlinearity over different temporal and spatial scales, whilst considering the influence of the role of different precipitation periods for flood generation in different regions across Kenya. However, such studies require more reliable data, which are currently lacking in Kenya. For example, the causes of historical flood trends in Kenya are still unclear, due to a limited understanding of regional variations in flood-generating mechanisms, e.g., in Tramblay *et al.* (2019), Tramblay *et al.*, 2019), land-use changes, reservoir construction and other local effects (Svensson, Kundzewicz and Maurer, 2005) and the uncertainty of projections of future flooding under climate change (Kundzewicz *et al.*, 2014; Berghuijs *et al.*, 2019).

In conclusion, the presence of statistically significant trends in observed and simulated flood series across Kenyan catchments is investigated in this study. In comparison to previous trend studies, three novel aspects are explored. First, significant trends are detected across the country in two flood series: annual maximum (AMAX) and the peak over threshold (POT) are compared. A comparison is made between the observed and simulated AMAX series. Then a sensitivity analysis of trends in floods to the selection of different exceedance thresholds in the POT flood series is performed. Finally, the timing and variability of the annual floods across the stations using the annual maximum flow is explored. This research acknowledges massive data gaps (see Table 14) in some of the station discharge data used in this study do not incorporate in-depth metadata such as the water abstraction data including dam construction and irrigation activities, so the trends may include the effects of changes in flow attenuation (e.g., reservoirs) and land use as well as any impacts of climate change. This study supports the importance of analysis of trends at country level as it highlights key characteristics that may not be captured in regional or global analysis. The findings are crucial in providing information on flood characteristics that can help to inform policy for disaster risk management, infrastructure design and agriculture, ultimately supporting the improvement of livelihoods in Kenya

6.9. Acknowledgements

Chapter 7

7. Discussion and concluding remarks

7.1. Discussion overview

This discussion and conclusion chapter summarises key research findings, scientific advances, challenges, and limitations of my research. The chapter also provides the main conclusion and future recommendations.

The next subsections look at the key messages from each of the chapters, challenges faced during each of the analyses and the limitations associated with the research findings.

7.2. Key messages from the research papers

This section summarizes the key findings from each of the three research papers presented in this thesis. Each paper addresses in turn the three objectives set in section 1.2.2 of the thesis.

7.2.1. Objective 1: - Design and propose an objective model pre-selection criterion with a filter sequence for a Kenyan national flood forecasting centre. Wanzala et al., 2022b. (Chapter 4, Appendix 1: - Hydrological model pre-selection with a filter sequence for the national 2 flood forecasting system in Kenya)

Choosing a model for operational flood forecasting is not simple because of differences in process representations, data scarcity issues and propagation of errors and uncertainty down the modelling chain. It is necessary to consider objective selection of the modelling tools at national scale.

- In this chapter, research findings show that not all models are good at capturing and/or representing the important processes relevant to flood generation (e.g., transmission losses along the river channel, re-infiltration, and subsequent evaporation of surface water) both in wetlands and ASALs of Kenya). Therefore, a single model would not be applicable to the entire country, because of stark differences in the hydroclimatic characteristics of catchments (such as wetlands and drylands). As such characteristics are dynamic (changing in space and time) due to natural and human influence, model developments and upgrades should allow the incorporation of such differing characteristics.

- A significant buy-in is required to develop operational forecasting capacity in a specific model, and so in recognition of changes in the importance and impact of many processes as a result of land use change, water management etc., it may mean that it is more efficient to choose a modelling approach that can represent a larger range of processes.

- The objective choice of modelling tools involves, for example, ensuring that the models are numerically stable, have reliable error and inconsistency checks, are able to flag missing data errors (e.g., when input sources fail), fit into an operational environment and are, preferably, user friendly.

This piece of work demonstrates the importance of understanding the qualitative and quantitative knowledge base of different catchments and how this influences the choice of modelling tools at catchment scale, whilst acknowledging the practical challenges (e.g., computing power and data scarcity challenges) that may prevent the use of complicated or resource hungry tools.

7.2.2. Objective 2: - Evaluation of different reanalysis precipitation datasets for hydrological modelling through performance statistics and parameter identifiability using sensitivity analysis to establish their influence on catchment streamflow simulations. Wanzala et al., 2022a (Chapter 5, Appendix 2: - Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: a case study of Kenya)

Hydrological modelling is essential to produce forecasts but is a challenging task, especially in poorly gauged catchments, because of the inadequate temporal and spatial coverage of hydro-meteorological observations. Open access global meteorological reanalysis datasets can fill in this gap, but they have significant errors, and careful evaluation is important to inform both the users and the developers of the datasets. However, more generalised evaluations may fail to capture the evolution and magnitude of events (e.g., floods). Research findings reveal that: -

- Assessment of these reanalysis precipitation products at catchment scale is important due to stark differences between river catchments. There is thus a need for combined performance statistics and uncertainty quantifications.

- Aggregating both sensitivity indices and performance statistics via the Model Suitability Index score provides clear evidence of the superiority (or inferiority) of a reanalysis with respect to observations.

- While ERA5 is the best performing dataset overall, performance varies by season and catchment, and therefore there are marked differences in the suitability of reanalysis products for forcing hydrological models.

- Overall, wetland catchments in the western regions and highlands of Kenya obtained better scores than those in the semi-arid regions.

These findings are important because they can inform future applications of reanalysis products to setting up hydrological models that can be used for flood forecasting, early warning, and early action in data-scarce regions, such as Kenya, whilst carefully communicating the associated errors and uncertainties.

7.2.3. Objective 3: - Assessing the historical trends in flood series and possible shifts in flood timing and seasonality across Kenyan catchments. (Chapter

6, Appendix 3: - Detecting trends in flood series and shifts in flood timing across Kenya)

The frequency and magnitude of flood events in Kenya have increased over the past decade. Trend analysis is useful in detecting and supporting the evidence of change in flow series, as well as variability in flood timing.

- In this chapter, research findings show that the mean number of discharge peaks per year varies spatially across Kenyan catchments from about 4 to 13 events (Figure 6.2).

- Observations show a shift in timing and variability in flood occurrences in most parts of the country. For example, floods occur in the peak rainfall months of April, May and November and are highly predictable in most of the coastal and western stations, and less predictable in stations whose annual floods occur in the dry months of June, July, and August (Figure 6.9).

- In what ways has the frequency and magnitude of flood events in Kenya changed over the past decade? AMAX series detects more trends in floods than the POT series, while the POT series detects more significant trends in flood frequency than flood magnitude when comparing the frequency and magnitude of floods observed in POT and AMAX flood series from 1981 to 2016 in 19 Kenyan catchments.

- In general, trends in AMAX series are in the same direction and can be seen for stations in proximity to each other, which is evidence of a spatially coherent pattern.

- POTfreq flood series show statistically significant trends, predominantly in stations in the western part of the country, suggesting the possible existence of a pattern in flood frequency over western parts of Kenya, but the pattern is unclear.

- Multiple Index values are relatively higher in the small arid and the semi-arid catchments. The arid and semi-arid climate of these catchments causes generally low mean discharge across the year which is offset by intense precipitation events leading to high discharge values which are compounded by an intrinsic high variance in the daily hydrograph, thus yielding the high MI values.

This information on flood characteristics can help to inform policy makers about the most effective ways to plan and put in place flood defences and emergency response funds for disaster risk management, infrastructure design and agriculture, and ultimately to support the improvement of livelihoods in Kenya.

7.3. Scientific advances, challenges and limitations of the research paper results

This section summarizes the scientific advances and challenges encountered in the implementation stages of the research as well as the modelling experiments across Kenyan catchments. It also explores some of the key issues raised by the anonymous reviewers during the review process of the papers at the time of submission for publication, as presented in Chapters 4 and 5 of this thesis, and the responses provided during the review process; finally, it the limitations of the research.

7.3.1. Proposed model selection framework: Wanzala et al., 2022b (Chapter 4, Appendix 1: - Hydrological model pre-selection with a filter sequence for the national 2 flood forecasting system in Kenya)

In Chapter 4, hydrological model selection has emerged as an ever-present challenge that can determine the direction of a flood forecasting system for decades. Model selection based on some criteria (outlined in Section 4.5) for suitability for a certain application (or the lack thereof in many cases: see Addor and Melsen, 2019) is also a wider problem. This research paper has determined what important aspects (hydroclimatic conditions, computing power and human factors) should be considered in the selection of a hydrological model from the many available choices for a national hydrological forecasting centre. The filter sequence presented in Section 4.6.1 offers a good starting point for a step-by-step selection, in cases where there is no customised model, which may reduce the bias in the choice and application of modelling tools.

However, there were challenges associated with the proposed selection criteria because some may be very important yet hard to implement. For example, there are several challenges in trying to model physical and hydroclimatic inhomogeneities, both within individual catchments and when there are such distinct zones of hydro-climatology within a country, as is the case for Kenya. This becomes even more challenging when the choice of modelling approach is constrained by low data availability and modest human resources and technical capacity. It may be ideal to use several different modelling approaches, particularly where there are such distinct hydroclimatic regions, or use a multi model approach to deal better with the modelling uncertainties, but pragmatically this is not an option for institutions who are able to take only one approach.

Additionally, in Chapter 4, I used MCA, as outlined in Sherlock and Duffy (2019), to assess multiple alternatives based on a mix of quantitative and mostly qualitative information from multiple sources to arrive at a model selection subset. However, the proposed MCA relies heavily on evaluation data and is very time consuming for the number of models available. Hence, for data-scarce regions, and/or agencies with limited resources, or in general, an additional decision tree is helpful to trim down the number of options. It is necessary to further evaluate the limited selection with, for example, an MCA and the FFC experiment, which may be difficult to implement, especially in Kenya.

Whilst this paper informed the choice of model for hydrological applications in Kenya, there are some limitations for the proposed aspects as well as the presented filter sequence which informed the evaluation of models to suggest the possible candidates. Some of the specific limitations are as follows:

- There is no documented research outlining the pros and cons of each of models in a single platform which a potential model user may easily use to identify which model is a suitable match for the presented criteria.

- The filter steps are not operationalized to the level where the process can be said to be objective. For example, a model may be excluded on 'many parameters' and the pre-selection criteria presented here follows a flow chart which may be subjective.

- Many of the criteria for selection have been based on expert judgement and linked to models that have been applied to diverse environments deemed suitable candidates for transfer to a Kenyan context.

- The suggestions of certain models depend on the computational capabilities (skills) of individual users as well as the NMHS in general. As a result, the model selection process is necessarily biased towards ease of application, assuming a relatively basic modelling capacity.

- This work does not look at direct analysis of each of the proposed models to evaluate its performance and accuracy with relation toa some past flood events, because it is massive job requiring high-level computational facilities and a long series of data.

7.3.2. Assessment of reanalysis precipitation products: Wanzala et al., 2022a (Chapter 5, Appendix 2: - Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: a case study of Kenya)

In Chapter 5, assessment of reanalysis precipitation datasets for hydrological modelling over Kenya provides a basis for the use of freely available alternatives in sparsely gauged and ungauged catchments. It also provides the developers of global reanalysis datasets with useful information for future developments of the products. Also, this paper contributes valuable information to the rapidly growing research area which has a very strong interest in hydrological modelling with earth system approaches supported by global forecasting centres such as ECMWF (see Harrigan, Cloke and Pappenberger, 2020).

There were a few challenges encountered in the analysis of this paper, including the following:

- **Unavailability of observational data** – Like many places in the world, Kenya suffers from several severe problems with observational data. The available precipitation gauging stations are sparse and/or not available at all. Most of the catchments in the study have either one or two stations which leads to very sparse coverage. Not only that, but the data series is short and includes massive gaps. For example, the Tana River catchment, one of our study catchments, is a 96,000 km^2 catchment but has only four precipitation gauging stations. These stations are unreliable and have large data gaps. Therefore, we had to use the satellite observation data (CHIRPS). This point was raised by the anonymous reviewers of the manuscript upon submission to the *Journal of Hydrology – Regional Studies*; they required further supporting information as to why this was the best option for Kenya. Below are the comments from an anonymous reviewer and the response provided during the review process.

"Why is the CHRPS considered as the benchmark? There are other observational precipitation datasets available such as IMERG, MSWEP and GSMaP. The authors should provide more information to claim that CHIRPS is more accurate than reanalysis products."

"We acknowledge the availability of other precipitation datasets with a high spatial resolution (≤10 km) such as the Multi-Source Weighted-Ensemble Precipitation (MSWEP; (Beck *et al.*, 2019), CPC morphing technique (CMORPH; 7 km Joyce *et al.*, 2004), Global Satellite Mapping of Precipitation (GSMaP; 10 km; Mega *et al.*, 2014), Integrated Multisatellite Retrievals for Global Precipitation Measurement (IMERG; 10 km ; Huffman *et al.*, 2001, 2015,

and Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS; 4 km; Hong *et al.*, 2004). However, all these datasets except for MSWEP have a data record of ≤20 years and do not take advantage of river discharge observations for bias correction, and do not incorporate reanalysis-based precipitation estimates. Additionally, precipitation evaluation studies particular to Eastern Africa (including Kenya) have less explored the potential of MSWEP as an alternative to observations and shown better performance for CHIRPS relative to other observations (see Dinku *et al.*, 2018) even though the daily temporal resolution of CHIRPS renders it less suitable in highly dynamic precipitation analysis and CHIRPS include spurious drizzle and underestimation of peak magnitudes of the most extreme rainfall ((Beck *et al.*, 2017), but our focus was not necessarily extreme rainfall should be insensitive to these biases.”

- **GR4J model calibration problems across Kenyan catchments:** GR4J model parameters were very unstable across most of the study catchments and as a result there were some cases where there was a huge difference between the observed and simulated discharge (Appendix **A4**). The instability of the model parameters may have been caused by the instability in the input data. Precipitation products often use all available information at a given time, which may not be stable in the long term. Strong changes may yield unstable model parameters or behaviour. Secondly, the model may have failed to manage to simulate a good water balance at the annual or monthly time steps or instability may be due to a redundancy in model function due to the specific characteristics of the catchment. For example, there are many reservoirs and dams as well as large irrigation schemes in across Kenya (see Figure 4.4). Most of the dams are regulated by the WRA and the abstraction activities are not monitored in the upstream areas. Consequently, dam releases and water abstraction result in fluctuations of the amount of water in the catchments and this may fail to be captured in the model.

Whilst this paper informed the choice of reanalysis precipitation datasets for hydrological modelling and applications in Kenya, there are some limitations of the study. The specific limitations include the following:

- The study used proxy satellite observation datasets instead of *in situ* observations and this may not provide ideal conditions for accurate performance of the reanalysis precipitation datasets. This is due to inherent uncertainties associated with the blended observations and satellite data used.

- Model calibration and modelling experiments did not take into the account the influence of human activities such as reservoir and dam constructions, irrigation, deforestation, etc., which modulate the hydrological cycle and form a key part of the hydrometeorological modelling chain.

- Modelling results are based on the performance of a single hydrological model which may have had problems in simulating a good water balance in the study catchments. Model simulation results in some catchments, such as Tana and Athi, having spurious peaks, i.e., peaks appearing earlier than they should, or underestimating the high flow peaks. An alternative was to trial several models and compare the results, but due to the time constraints of this research, calibrations of several models could not be achieved.

7.3.3. Trends in river flow series and shifts in flood timings: Wanzala et al. (Chapter 6, Appendix 3: - Detecting trends in flood series and shifts in flood timing across Kenya)

In Chapter 6, trend detection in flow series provides an understanding of flood characteristics (frequency and magnitude) at country scale. This is important for informing policy for disaster risk management, infrastructure design and agriculture, amongst other hydrological applications. Additionally, insights into the variability of flood timing and seasonality is important for water resources planning and management because changes in the timing of the yearly flood have far-reaching consequences for flood-based farming systems, especially for the livelihoods of people who adjust their floodplain management and agricultural activities to the rise and fall of the flood wave.

Some of the challenges encountered in this study included:

- Difficulty in detecting statistically significant trends in river flow time-series due to the high degree of natural variability (low signal to noise).

- Shorter River flow time series with massive gaps as well as some stations lacking the most recent data records, which probably affects the trend patterns and thus makes it very difficult to establish clear trends.

- Difficulty in exceedance threshold selection in the POT series. The number of peaks considered in each flood series are sensitive to the exceedance threshold. This means that selecting a higher exceedance threshold would result in a lower threshold, thus yielding a higher number of flood peaks in the series, and the reverse is true. This may not be a true reflection of the expected flood peaks. The reflection of the resultant Mean Residual Life plots was used to pick the desirable threshold.

Whilst this paper provided new insights into the changing trends in river flow series, using AMAX and POT high flow indices and the seasonality and timing of floods across Kenya, the study has some specific limitations. They include the following:

- Trend analysis studies recommend long and continuous data series, which was a big limitation for this research. Most of the stations did not have long series and recent data records to correctly identify shifts in trends: this led to the decision to settle for a common analysis period (1981-2016)

- Trend analysis does not incorporate the impacts of climate change on the flow series over the study period, which is assumed to be constant in time, as well as the impacts of human activities such as reservoir/dam construction, deforestation, and irrigation schemes. These heavily influence the rise and fall of river levels and make the flow more dynamic over time.

7.4. Concluding remarks

Overall, the work in this thesis has explored and provided an enhanced understanding of the avenues to improving flood modelling and forecasting in Kenya in terms of models, data, and historical trends of floods as well as seasonality and shifts in flood timing. The results have provided objective criteria for model selection following a filter sequence for a national flood

forecasting centre, the application of alternative sources of data to counter observed data scarcity challenges and the understanding of historical flood trends and timing.

Understanding variability in flood timing and seasonality is important for water resources planning and management. Changes in the timing of the yearly flood have far-reaching consequences for flood-based farming systems, especially for the livelihoods of people who adjust their floodplain management and agricultural activities to the rise and fall of the flood wave. Therefore, the next steps should look into establishing the attribution of floodiness across Kenyan catchments and the changes instigated by climate change impacts. To conclude, research presented in this thesis was conducted in the context of and was cognizant of hydrological modelling and operational flood risk management and preparedness action conducted at that time in Nzoia River basin: therefore, the research findings can have beneficial effects to upscaling research in other flood prone areas and improving the flood forecasting useful to emergency response units and policy makers responsible for proper planning and decision making.

Some of the specific future work informed by the research findings in this thesis is summarised in the following sub-section 8.1.

7.5. Recommendations for further study

The work in this thesis has raised several questions and topics for future research, some of which have been touched upon in each chapter. This section provides specific examples of ways in which this research could be extended.

Future work that builds on the findings presented in this thesis should look at establishing the accuracy of the suggested models in simulating past flood events, in order to discover which is most efficient, according to quantitative analysis. One of the biggest obstacles is the unavailability of long and consistent datasets on streamflow and precipitation that accurately reflect flood peaks in the observed data. This is due to the limited number of stations with long records of reliable data (due to massive data gaps) across most Kenyan catchments. Thus, the next steps rely heavily on the improvement of satellite and reanalysis data to extend the period of analysis and provide stronger conclusions. With the precipitation input found to be the most dominant component of the hydrometeorological modelling chain, it is logical to focus efforts on improving the precipitation products not only for Kenya, but in the applications of the larger hydrological modelling research community.

- **A comparison study of hydrological models' analysis**. One of the recommendations raised in Chapter 4 was to carry out a further analysis of the proposed hydrological model candidates for Kenyan catchments to assess their performance skills in simulating past events. For example, due to time constraints and unavailability of observational data, the research did not investigate the individual model skill assessment. However, ideally, a further analysis would be undertaken whereby the number of suggested models would be evaluated with a combination meteorological input, hydrological and river routing component, to allow a better evaluation, leading to the best model selection for the criteria in general and Kenya's National Hydrological Forecasting Centre in particular.

- **Separating the operationalised forecasting system using a modular approach.** The gold standard for operational systems is to separate different parts of the modelling chain in a modular approach, e.g., Delft-FEWS (Werner, Reggiani and Weerts, 2014; Werner *et al.,* 2013). For example, the model selection process considered in Chapter 4 was based on the selection criteria and linked to Kenyan hydroclimatic characteristics. Ideally, an approach considering a modular system component with independent criteria would be beneficial as it provides a better assessment of the aspects to consider when making a model selection for operational flood forecasting.

- **A comparison study of the satellite and reanalysis datasets analysis.** One of the concerns raised in Chapter 5 was the use of CHIRPS as a benchmark dataset amid a variety of multi-satellite reanalysis products. For example, Chapter 5 exclusively evaluates the performance of reanalysis datasets and, due to time constraints, this thesis did not compare the performance statistics with other available observation datasets. However, ideally, a further analysis would be undertaken whereby the freely available satellite observations would be evaluated with a combined model calibration and validation experiment to allow a better evaluation of the best dataset selection for hydrological modelling in data-scarce Kenyan catchments.

- **Incorporation of Kenyan dams in the GR4J.** Overall, dam construction in the upstream areas is increasing across Kenyan catchments (Mwangi *et al.,* 2016). When a comparison was made between simulated and observed discharge, it was noticed that the flood peaks tended to occur too early along some rivers, such as the Tana River and the Athi River in the coastal region (see Appendix **A4**). The current study did not represent most of the Kenyan dams in the GR4J model. However, reservoirs were found to affect model performance substantially in some catchments when they were incorporated into the model. This gives hope that the inclusion of existing dams in the Kenya GR4J model could help to increase the accuracy of flood forecasts in the basin.

- **Testing simplified versions and the internal variables of the GR4J model:** The instability of GR4J model parameters may also come from a redundancy in model function due to the specific characteristics of the study catchments. Model simplifications may be a way to get more robust simulations: for example, running the model without groundwater exchanges (X2 fixed at zero) or with a fixed capacity for the production store. Additionally, the model's internal variables could be checked: analyses of water contents in model stores and water fluxes simulated by the model (AE, exchanges, etc.) are often very interesting, providing better understand of possible compensations between model components (e.g., Ficchi *et al,* 2016). This can be done at the long term annual or monthly scales, to establish whether a large routing store and water exchange does something very similar to the production store and AE losses. This would be useful in establishing the behaviour of the PE and discovering whether it is too close to constant over the year because of the seasonality of the production store, which would therefore be a likely cause of model underperformance for Kenyan catchments.

- **Testing another hydrological model or a multi-model ensemble:** it may be a good option to consider a better evaluation of the possible role of structural and input uncertainty.

- **Flood generating mechanisms across Kenyan catchments**: Chapter 6 investigates the trends and shifts in flood timing. There is a marked difference in the trend direction, frequency, and magnitude of flooding in catchments within close proximities, pointing to variations in flood generating mechanisms/drivers in Kenya. Causes of flooding have been studied in Kenya to establish the link between flooding and meteorological drivers (e.g., Macleod *et al.,* 2021, Kilavi *et al.,* 2018). However, no documented research has investigated the mesoscale and local drivers of flooding in Kenya, even though similar work has been done in different parts of the world (e.g., Berghuijs *et al.,* 2016; 2017; 2019; Blöschl *et al.,* 2013). Acknowledging that there has been a shift in the frequency and magnitude of floods in recent decades across Kenya and assessing different flood generating drivers would be a useful step towards understanding their contribution to floodiness, with particular interest in quantitatively establishing their relative importance and change over time in different catchments. For example, see Appendix **A6**, which was the next plan, but due to time constraints on the thesis, the research was not actualised.

References

Abbott, M.B., Babovic, V.M. and Cunge, J.A. (2001). Towards the hydraulics of the hydroinformatics era. *Journal of Hydraulic Research*, **39** (4): 339-349.

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J. (1986). An introduction to the European Hydrological System — Système Hydrologique Européen, "SHE", 2: structure of a physically-based, distributed modelling system. *Journal of Hydrology*, **87** (1–2): 61–77.

Abbott, M.B., Babovic, V.M. and Cunge, J.A. (2001). Towards the hydraulics of the hydroinformatics era. *Journal of Hydraulic Research*, **39** (4): 339-349.

Abraham, T., Woldemicheala, A., Mulumeha, A. and Abateb, B. (2018). Hydrological Responses of Climate Change on Lake Ziway Catchment, Central Rift Valley of Ethiopia.*Journal of Earth Science & Climatic Change,* **9** (6): doi: 10.4172/2157-7617.1000474.

Acharya, S.C., Nathan, R., Wang, Q.J., Su, C.H. and Eizenberg, N. (2019). An evaluation of daily precipitation from a regional atmospheric reanalysis over Australia. *Hydrology and Earth System Sciences*, **23** (8): 3387–3403.

Archer, G.E.B., Saltelli, A. and Sobol', I.M. (1997). Sensitivity measures, ANOVA-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation*, **58** (2): 99–120.

Arnold, J.G., Srinivasan, R., Muttiah, R.S. and Williams, J.R., (1998). Large area hydrologic modeling and assessment part I: model development 1. *JAWRA Journal of the American Water Resources Association*, *34*(1), pp.73-89.

Adeyeri, O.E., Laux, P., Arnault, J., Lawin, A.E. and Kunstmann, H. (2020). Conceptual hydrological model calibration using multi-objective optimization techniques over the transboundary Komadugu-Yobe basin, Lake Chad Area, West Africa. *Journal of Hydrology: Regional Studies,* **27:** 100655. doi: 10.1016/j.ejrh.2019.100655.

Addor, N., and Melsen, L. A. (2019). Legacy, rather than adequacy, drives the selection of hydrological models. *Water Resources Research*, **55** (1): 378-390.

Agrawal, N. and Desmukh, T.S. (2016). Rainfall using MIKE 11 nam– review. *International Journal of Innovative Science, Engineering & Technology*, **3** (6): 659–667.

Alemayehu, T., Kilozo, F., van Griensven, A. and Bauwens, W (2018). Evaluation and application of alternative rainfall data sources for forcing hydrologic models in the Mara Basin. *Hydrology Research*, **49** (4): 1271–1282.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. (1998). *Crop Evapotranspiration: Guidelines for computing crop water requirements*. FAO Irrigation and drainage paper 56. 1st ed. Rome: Food and Agriculture Organization of the United Nations.

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J. and Pappenberger, F. (2013). GloFAS-global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, **17** (3): 1161–1175. doi: 10.5194/hess-17-1161-2013.

Ali, A., Amani, A., Diedhiou, A. and Lebel, T. (2005). Rainfall estimation in the Sahel. Part II: evaluation of rain gauge networks in the CILSS countries and objective intercomparison of rainfall products. *Journal of Applied Meteorology and Climatology*, **44** (11): 1707-1722.

Anagnostopoulos, G.G., Fatichi, S. and Burlando, P. (2015). An advanced process-based distributed model for the investigation of rainfall-induced landslides: effect of process representation and boundary conditions. *Water Resources Research,* **51** (9): 7501–7523.

Anderson, M.P., Woessner, W.W. and Hunt, R.J. (2015). *Applied groundwater modeling: simulation of flow and advective transport*. 2nd ed. London: Elsevier Academic press.

Andersson, J., Pechlivanidis, I., Gustafsson, D., Donnelly, F. and Arheimer, B. (2015). Key factors for improving large-scale hydrological model performance. *European Water*. **4:** 77–88.

Arheimer, B. and Lindström, G. (2015). Climate impact on floods: changes in high flows in Sweden in the past and the future (1911–2100). *Hydrology and Earth System Sciences*. **19** (2): 771–784.

Arlimasita, A. and Lasminto, U. (2020). Sensitivity Analysis of Rainfall-Runoff Model in Malino Sub-Watershed. *International Journal on Advanced Science, Engineering and Information Technology*, 10 (4): 1578-83.

Arnal, L. (2014). An intercomparison of flood forecasting models for the Meuse River basin. MSc Thesis. Deltares, IRSTEA and University of Amsterdam.

Arnell, N.W. and Gosling, S.N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, **134** (3): 387–401.

Arnold, J.G., Srinivasan, R., Muttiah, R.S. and Williams, J.R. (1998). Large area hydrologic modeling and assessment part I: model development', *Journal of the American Water Resources Association*, **34** (1): 73–89. doi: 10.1016/S0899-9007(00)00483-4.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, D., van Griensven, A., van Liew, M., Kannan, K. and Jha, M.K. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE,* **55** (4): 1491-1508.

Arshad, M., Ma, X, Yin, J., Ullah, W., Liu, M. and Ullah, I. (2021). Performance evaluation of ERA-5, JRA-55, MERRA-2, and CFS-2 reanalysis datasets, over diverse climate regions of Pakistan. *Weather and Climate Extremes*, **33:** 100373.

Artan, G., Gadain, H., Smith, J.L., Asante, K., Bandaragoda, C.J. and Verdin, J.P. (2007). Adequacy of satellite derived rainfall data for stream flow modeling. *Natural Hazards*, **43** (2): 167–185. doi: 10.1007/s11069-007-9121-6.

Artan, G.A., Asante, K.O. and Verdin, J.P. (2004). Continental Scale Flood Hazard Monitoring System. *American Geophysical Union Spring Meeting Abstracts*. H4 1D-14.

Artan, G., Verdin, J. and Asante, K. (2001). A wide-area flood risk monitoring model. *Fifth international workshop on application of remote sensing in hydrology, Montpellier, France*.

Asante, K.O., , G.A., Pervez, S. and Rowland, G. A. (2008). A linear geospatial streamflow modeling system for data sparse environments. *International Journal of River Basin Management*, **6** (3): 233–241.

Asante, K.O., Artan, G.A., Pervez, S., Bandaragoda, C. and Verdin, J.P. (2008). *Technical Manual for the Geospatial Stream Flow Model (GeoSFM): US Geological Survey Open-File Report 2007-1441*, Reston: US Geological Survey.

Asnani, G. C. and Kinuthia, J. H. (1979). *Diurnal variation of precipitation in East Africa*. Nairobi: Republic of Kenya, Meteorological Department, East African Institute for Meteorological Training and Research.

Atkinson, S.E., Woods, R.A. and Sivapalan, M. (2002). Climate and landscape controls on water balance model complexity over changing timescales. *Water Resources Research*, **38** (12): 50–51.

Aura, C.M., Nyamweya, C.S., Odoli, C.O., Owiti, H., Njiru, J.M., Otuo, P.W., Waithaka, E. and Malala, J. (2020). Consequences of calamities and their management: The case of COVID-19 pandemic and flooding on inland capture fisheries in Kenya. *Journal of Great Lakes Research*, **46** (6): 1767-1775.

Ávila, Á., Guerrero, F.C., Escobar, Y.C. and Justino, F. (2019). Recent precipitation trends and floods in the Colombian Andes. *Water*, **11** (2): 379.

Awan, U.K., Liaqat, U.W., Choi, M. and Ismaeel, A. (2016). A SWAT modeling approach to assess the impact of climate change on consumptive water use in Lower Chenab Canal area of Indus basin. *Hydrology Research*, **47** (5): 1025–1037.

Ayugi, B.O., Tan, G., Ongoma, V. and Mafuru, K. B. (2018). Circulations associated with variations in boreal spring rainfall over Kenya. *Earth Systems and Environment*, **2** (2): 421–434.

Ayugi, B.O, Tan, G., Niu, R., Dong, Z.Y., Ojara, M., Mumo, L., Babaousmail, H., Ongoma, V. (2020). Evaluation of meteorological drought and flood scenarios over Kenya, East Africa. *Atmosphere,* **11** (3). doi: 10.3390/atmos11030307.

Ayugi, B.O., Wen, W. and Chepkemoi, D. (2016). Analysis of spatial and temporal patterns of rainfall variations over *Journal of Environment and. Earth Science*, **6** (11): 69-83.
Bai, P., Liu, X., Kang, L. and Liu, C. (2015). Comparison of performance of twelve-monthly water balance models in different climatic catchments of China. *Journal of Hydrology* **529** (3): 1030–1040.

Baker, T.J. and Miller, S.N. (2013). Using the Soil and Water Assessment Tool (SWAT) to assess land use impact on water resources in an East African watershed'. *Journal of Hydrology*, **486:** 100–111.

Baldassarre, G.D. and Montanari, A. (2009). Uncertainty in river discharge observations: a quantitative analysis' *Hydrology and Earth System Sciences* **13** (6): 913–921.

Barasa, B.N. and Perera, E.D.P. (2018). Analysis of land use change impacts on flash flood occurrences in the Sosiani River basin Kenya. *International Journal of River Basin Management*, **16** (2): 179–188.

Barasa, E., Kazungu, J., Orangi, S., Kabia, E., Ogero, M. and Kasera, K. (2021). Indirect health effects of the COVID-19 pandemic in Kenya: a mixed methods assessment. *BMC Health Services Research*, **21** (1):1-16.

Bárdossy, A. and Singh, S.K. (2008). Robust estimation of hydrological model parameters. *Hydrology and Earth System Sciences*, **12** (6): 1273–1283.

Barnston, A.G., Li. S., Mason, S.J., DeWitt, D.G., Goddard, L. and Gong, X (2010). Verification of the first 11 years of IRI's seasonal climate forecasts. *Journal of Applied Meteorology and Climatology*, **49** (3): 493–520.

Bartholmes, J.C., Thielen, J., Ramos, M.H and Gentilini, S. (2009). The european flood alert system EFAS - Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrology and Earth System Sciences*, **13** (2):141–153. doi: 10.5194/hess-13-141-2009.

Bashar, K.E. (2012). Comparative performance of soil moisture accounting approach in continuous hydrologic simulation of the Blue Nile. *Water Science and Engineering*, **5**: 1–10.

Bauer, H.S., Schwitalla, T., Wulfmeyer, V., Bakhshaii, A., Ehret, U., Neuper, M. and Caumont, O. (2015). Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF–a performance test. *Tellus A: Dynamic Meteorology and Oceanography*, **67** (1): 25047.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A.I.J.M., Weedon, G.P., Brocca, L., Pappenberger, F., Huffman, G.J. and Wood, E.F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, **21** (12): 6201–6217.
Beck, H.E., van Dijk, A.I.J.M., De Roo, A., Dutra, E., Fink, G, Orth, R. and Schellekens, J. (2017). Global evaluation of runoff from ten state-of-the-art hydrological models. *Hydrology and Earth System Sciences*, **21** (6): 2881–2903.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., Van Dijk, A.I.J.M., McVicar, T.R. and Adler, R.F. (2019). 'MSWEP V2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment', *Bulletin of the American Meteorological Society*, **100** (3): 473–500.

Beck, H.E., Pan, M., Miralles, D.G., Reichle, R.H., Dorigo, W.A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R.M., van Dijk, A.I.J.M., Du J., Kimball, J., Vergopolan, N. and Wood, E.F. (2021). Evaluation of 18 satellite-and model-based soil moisture products using in situ measurements from 826 sensors. *Hydrology and Earth System Sciences*, **25** (1): 17–40.

Bennett, J.C., Robertson, D.E., Ward, P.G.D., Hapuarachchi and Wang, Q.J. (2016). Calibrating hourly rainfall-runoff models with daily forcings for streamflow forecasting applications in meso-scale.*Environmental Modelling & Software*, **76:** 20–36.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D. and Andréassian, V. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, **40:** 1–20.

Berghuijs, W.R., Aalbers, E., Larsen, J., Trancoso, R. and Woods, R.A. (2017). Recent changes in extreme floods across multiple continents. *Environmental Research Letters*, **12** (11): 114035.

Berghuijs, W.R., Slater, L.J., Harrigan, S., Molnar, P. and Kirchner, J.W. (2019). The relative importance of different flood-generating mechanisms across Europe. *Water Resources Research*, **55** (6): 4582–4593.

Berglöv, G., German, J., Gustavsson, H., Harbman, U. and Johansson, B. (2009). *Improvement HBV Model Rhine in FEWS: Final Report*. *Hydrology*, No. 112. Norrköping: Swedish Meteorological and Hydrological Institute.

Bergstrom. S. (1995). The HBV model. In: V.P. Singh (ed.) *Computer models of watershed hydrology* Highlands Ranch, Colorado: Water Resources Publications, pp. 473-476.

Berthet, L., Andréassian, V., Perrin, C. and Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences*, **13** (6): 819-831.

Bethel, B.J. and Dusabe, K. (2021). The Uncertain Influence of the African Great Lakes and the Indian Ocean Dipole on Local-scale East Africa Short Rains. *SOLA.*, **1**7: 158-163. doi: 10.2151/sola.2021-028.

Betts, A.K., Chan, D.Z. and Desjardins, R.L. (2019). Near-surface biases in ERA5 over the Canadian prairies. *Frontiers in Environmental Science*, **7,** https://doi.org/10.3389/fenvs.2019.00129.

Beven K.J. (1989). Changing ideas in hydrology: the case of physically-based models. *Journal of Hydrology,* **105**:157–172. doi: 10.1016/0022-1694(89)90101-7.

Beven, K.J. (1995). Linking parameters across scales: subgrid parameterizations and scale dependent hydrological models. *Hydrological Processes*, **9** (5-6): 507–525.

Beven, K.J. (2001). How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences,* **5** (1):1-12.

Beven, K.J. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, **320** (1–2): 18–36.

Beven, K.J. (2010). *Environmental modelling: uncertain future?* 1st ed. Boca Raton, Florida: CRC press.

Beven, K.J. (2011). *Rainfall-runoff modelling: the primer*. 1st ed. Oxford: John Wiley & Sons.

Beven, K.J. (2012). *Rainfall-runoff modelling: the primer*. 2nd ed. Oxford: John Wiley & Sons. doi: 10.1002/9781119951001.

Beven, K.J. (2016). Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, **61** (9): 1652–1665.

Beven, K.J. and Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, **6** (3): 279–298.

Beven, K.J. and Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, **249** (1–4): 11–29.

Beven, K.J., Kirkby, M.J., Schofield, N. and Tagg, A.F. (1984). Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. *Journal of Hydrology*, **69** (1–4): 119–143.

Beven, K.J. and Kirkby, M.J. (1979). A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, **24** (1): 43–69.

Bezak, N., Brilly, M. and Šraj, M. (2016). Flood frequency analyses, statistical trends and seasonality analyses of discharge data: a case study of the Litija station on the Sava River. *Journal of Flood Risk Management*, **9** (2): 154–168.

Birkel, C., Tetzlaff, D., Dunn, S.M. and Soulsby, C2010). Towards a simple dynamic process conceptualization in rainfall–runoff models using multi-criteria calibration and tracers in temperate, upland catchments. *Hydrological Processes: An International Journal*, **24** (3): 260–275.

Birundu, A.M., Mutua, B. (2017). Analyzing the Mara River Basin behaviour through rainfall-runoff modeling. *International Journal of Geosciences*, 8: 1118-1132.

Bisselink, B., Zambrano-Bigiarini, M., Burek, P. and de Roo, A. (2016). Assessing the role of uncertain precipitation estimates on the robustness of hydrological model parameters under highly variable climate conditions. *Journal of Hydrology: Regional Studies* **8**: 112–129. doi: 10.1016/j.ejrh.2016.09.003.

Bitew, M.M., Gebremichael, M., Gebremichael, L.T and Bayissa, Y.A, (2012). Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia. *Journal of Hydrometeorology*, **13** (1): 338–350.

Black, E., Slingo, J. and Sperber, K.R. (2003). An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. *Monthly Weather Review*, **131** (1): 74–94.

Blanc, E. and Strobl, E. (2013). The impact of climate change on cropland productivity: evidence from satellite-based products at the river basin scale in Africa. *Climatic change*. **117** (4): 873–890.

Blau, M.T. and Ha, K. (2020). The Indian Ocean dipole and its impact on East African short rains in two CMIP5 historical scenarios with and without anthropogenic influence. *Journal of Geophysical Research: Atmospheres*, **125** (16): e2020JD033121.

Blöschl, G. and Sivapalan, M. (1995). Scale issues in hydrological modelling: a review. *Hydrological Processes*, **9** (3-4): 251–290.

Blöschl, G., Hall, J., Parajka, J., Perdigao, R.A.P., Merz, B., Arheimer, B. and Kjelsden, T. (2017). Changing climate shifts timing of European floods. *Science*, **357** (6351): 588–590.

Boelee, L., Lumbroso, D.M., Samuels, P.G, and Cloke, H.L. (2019). Estimation of uncertainty in flood forecasts—A comparison of methods. *Journal of Flood Risk Management*, **12** (S1): e12516.

Bongartz, K. (2003). Applying different spatial distribution and concepts in three nested mesoscale catchments of Germany. *Physics and Chemistry of the Earth, Parts A/B/C*, **28** (33-36): 1343-1349.

Booij, M.J., Tollenaar, D., van Beek, E. and Kwadijk, J. (2011). Simulating impacts of climate change on river discharges in the Nile basin. *Physics and Chemistry of the Earth, Parts A/B/C*, **36** (13): 696–709.

Booker, D.J. and Woods, R.A. (2014). Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. *Journal of Hydrology*, **508:** 227–239.

Brocca, L., Melone, F., Moramarco, T. and Singh, V.P. (2009). Assimilation of observed soil moisture data in storm rainfall-runoff modeling. *Journal of Hydrologic Engineering*, 14 (2): 153-165.

Burn, D. H. (1997). Catchment similarity for regional flood frequency analysis using seasonality measures. *Journal of hydrology*, **202** (1–4): 212–230.

Burn, D. H., Sharif, M. and Zhang, K. (2010). Detection of trends in hydrological extremes for Canadian watersheds. *Hydrological Processes*, **24** (13): 1781–1790. doi: 10.1002/hyp.7625.

Burn, D. H., Whitfield, P. H. and Sharif, M. (2016). Identification of changes in floods and flood regimes in Canada using a peaks over threshold approach. *Hydrological Processes*, **30** (18): 3303–3314.

Butts, M.B., Payne, J.T., Kristensen, M. and Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology*, **298**: 242-266.

Camberlin, P. and Okoola, R.E., (2003). The onset and cessation of the "long rains" in eastern Africa and their interannual variability. *Theoretical and Applied Climatology*, **75** (1): 43-54.

Camberlin, P., Moron, V., Okoola, R., Philippon, N. and Gitau, W. (2009). Components of rainy seasons' variability in Equatorial East Africa: onset, cessation, rainfall frequency and intensity. *Theoretical and Applied Climatology*, 98 (3): 237-249.

Cameron, D., Beven, K. and Naden, P. (2000). Flood frequency estimation by continuous simulation under climate change (with uncertainty *Hydrology and Earth System Sciences*, **4** (3): 393-405.

Camera, C., Bruggeman, A., Zittis, G., Sofokleous, I. and Arnault, J., 2020. Simulation of extreme rainfall and streamflow events in small Mediterranean watersheds with a one-way-coupled atmospheric–hydrologic modelling system. *Natural Hazards and Earth System Sciences*, **20** (10): 2791-2810.

Chen, J., Brissette, F.P. and Chen, H. (2018). Using reanalysis-driven regional climate model outputs for hydrology modelling. *Hydrological Processes*, **32** (19): 3019–3031.

Chen, J. and Wu, Y. (2012). Advancing representation of hydrologic processes in the Soil and Water Assessment Tool (SWAT) through integration of the TOPographic MODEL (TOPMODEL) features. *Journal of Hydrology*, **420-421:** 319–328.

Chen, Y., Niu, J., Kang, S and Zhang, X. (2018). Effects of irrigation on water and energy balances in the Heihe River basin using VIC model under different irrigation scenarios. *Science of the Total Environment*, **645:** 1183–1193.

Cherkauer, K.A., Bowling, L.C. and Lettenmaier, D.P. (2003). Variable infiltration capacity cold land process model updates. *Global and Planetary Change*, **38** (1–2): 151–159.

Cheruiyot, M., Gathuru, G. and Koske, J. (2018). Quantity and trends in streamflows of the Malewa River Basin, Kenya. *Journal of Environmental Science and Engineering B*, 7(1): 1–11.

Clark, M.P., Slater, A.G., Rupp, D.E., Woods, R.A., Vrugt, J.A., Gupta, H.V., Wagener, T. and Hay, L.E. (2008). Framework for Understanding Structural Errors (FUSE): a modular framework to diagnose differences between hydrological models. *Water Resources Research,* **44** (12). doi: 10.1029/2007WR006735.

Cloke, H.L. and Pappenberger, F. (2009). Ensemble flood forecasting: a review. *Journal of Hydrology*, **375** (3-4): 613-626.

Cloke, H.L., Pappenberger, F. and Renaud, J. (2008). Multi-method global sensitivity analysis (MMGSA) for modelling floodplain hydrological processes. *Hydrological Processes*, 22 (11): 1660–1674.

Cloke, H.L. and Hannah, D.M. (2011). Large-scale hydrology: advances in understanding processes, dynamics and models from beyond river basin to global scale. *Hydrological Processes*, **25** (7): 991–995.

Cloke, H. (2021). Europe's catastrophic flooding was forecast well in advance – what went so wrong? https://theconversation.com/europes-catastrophic-flooding-was-forecast-well-in-advance-what-went-so-wrong-164818: Accessed 06/07/2022.

Collet, L., Beevers, L. and Prudhomme, C. (2017). Assessing the impact of climate change and extreme value uncertainty to extreme flows across Great Britain. *Water*, **9** (2): 103.

Collier, E., Mölg, T. and Sauter, T. (2016). Analysis and simulation of recent climate variability in the high-mountain regions of East Africa: *EGU General Assembly Conference Abstracts:* EPSC2016-6136.

Connor, R. (2015) *The United Nations world water development report 2015: Water for a sustainable world*. Paris: UNESCO publishing.

Coron, L. Delaigue, O., Thirel, G., Perrin, C. and Michel, C (2019). airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling (v. 1.2. 13.16).

Cunderlik, J. M. and Ouarda, T. B. M. J. (2009). Trends in the timing and magnitude of floods in Canada. *Journal of Hydrology*, **375** (3–4): 471–480.

Daniel, E.B., Camp, J.V., LeBoeuf, E.J., Penrod, J.R., Dobbins, J.P. and Abkowitz, M.D. (2011). Watershed modeling and its applications: state-of-the-art review. *The Open Hydrology Journal*, **5:** 26-50.

Da Costa J., (2021). Report from Europe's flood zone: researcher calls out early warning system gridlock amid shocking loss of life: https://theconversation.com/report-from-europes-flood-zone-researcher-calls-out-early-warning-system-gridlock-amid-shocking-loss-of-life-164648. Accessed 06/07/2022.

Decker, M., Brunke, M.A., Wang, Z., Sakaguchi, K., Zeng, X. and Bosilovich, M.G. (2012). Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *Journal of Climate*, 25 (6): 1916–1944.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, d. P., Bechtold, P., van de Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J-J., Park, B-K., de Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J-N. and Vitart, F. (2011). The ERA-Interim reanalysis: and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137** (656): 553–597. doi: 10.1002/qj.828.

Degefu, M.A., Alamirew, T., Zeleke, G. and Bewket, W. (2019). Detection of trends in hydrological extremes for Ethiopian watersheds, 1975–2010. *Regional Environmental Change*, **19** (7): 1923–1933. Available at: https://doi.org/10.1007/s10113-019-01510-x.

Delaigue, O., Thirel, G., Coron, L. and Brigode, P. (2019). airGR and airGRteaching: two packages for rainfall-runoff modeling and teaching hydrology.In: *15th edition of the International R User Conference,* Jul 2019, Toulouse, France: 1. hal-02609956.

Demetriou, C. and Punthakey, J.F. (1998). Evaluating sustainable groundwater management options using the MIKE SHE integrated hydrogeological modelling package. *Environmental Modelling & Software*, **14** (2–3): 129–140.

Dessalegn, T.A., Moges, M.A., Dagnew, D.C. and Gashaw, A. (2017). Applicability of Galway Forecasting and GFFMS) for Lake Tana Basin, Ethiopia. *Journal of Water Resource and Protection,* **9** (12): 1319–1334. doi: 10.4236/jwarp.2017.912084.

Dessu, S.B., Seid, A., Abiy, A.Z. and Melesse, A. (2016). Flood forecasting and stream flow simulation of the upper Awash River basin, Ethiopia using geospatial stream flow model (GeoSFM). In.: A. Melesse and W. Abtew (eds) *Landscape Dynamics, Soils and Hydrological Processes in Varied Climates*, Cham: Springer, pp. 367–384.

Devak, M. and Dhanya, C.T. (2017). 'Sensitivity analysis of hydrological models: review and way forward', *Journal of Water and Climate Change*, 8(4): 557–575.

Devia, G.K., Ganasri, B.P. and Dwarakish, G.S. (2015). A Review on Hydrological Models. *Aquatic Procedia*, 4: 1001–1007. doi: 10.1016/j.aqpro.2015.02.126.

Dhanya, C.T. and Villarini, G. (2017). An investigation of predictability dynamics of temperature and precipitation in reanalysis datasets over the continental United States. *Atmospheric Research*, **183:** 341–350.

Dietterick, B.C., Lynch, J.A. and Corbett, E.S. (1999). A calibration procedure using TOPMODEL to determine suitability for evaluating potential climate change effect on water yield. *Journal of the American Water Resources Association*, **35** (2): 457–468.

Dile, Y.T. and Srinivasan, R. (2014). Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. *Journal of the American Water Resources Association*, **50** (5): 1226–1241.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H. and Ceccato, P. (2018). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society*, **144**: 292–312.

Diop, L., Yaseen, Z.M., Bodian, A., Djaman, K. and Brown, L. (2018) Trend analysis of streamflow with different time scales: a case study of the upper Senegal River. *ISH Journal of Hydraulic Engineering*, **24** (1): 105–114.

Diro, G.T., Grimes, D.I.F., Black, E., O'Neill, A. and Pardo-Iguzquiza, E. (2009). Evaluation of reanalysis rainfall estimates over Ethiopia. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, **29** (1): 67–78.

Dobler, C. and Pappenberger, F. (2013). Global sensitivity analyses for a complex hydrological model applied in an Alpine watershed. *Hydrological Processes*, **27** (26): 3922–3940.

Döll, P., Fiedler, K. and Zhang, J. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrology and Earth System Sciences*, **13** (12): 2413–2432.

Donnelly, C., Andersson, J.C.M. and Arheimer, B. (2016). Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe. *Hydrological Sciences Journal*, **61** (2): 255–273.

Doulgeris, C., Georgiou, P., Papadimos, D. and Papamichail, D. (2011). Evaluating three different model setups in the MIKE 11 NAM model. In: N. Lambrakis, G. Stournaras and K. Katsanou (eds) *Advances in the Research of Aquatic Environment: Vol. 1,* Berlin: Springer, pp. 241–249.

Doulgeris, C., Georgiou, P., Papadimos, D. and Papamichail, D. (2012). Ecosystem approach to water resources management using the MIKE 11 modeling system in the Strymonas River and Lake Kerkini. *Journal of Environmental Management,* **94** (1): 132-143.

Duan, Q., Sorooshian, S. and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water,* **28** (4): 1015–1031.

Dutra, E., Magnusson, L., Wetterhall, F., Cloke, H.L., Balsamo, G., Boussetta, S. and Pappenberger, F. (2013). The 2010–2011 drought in the Horn of Africa in ECMWF reanalysis and seasonal forecast products. *International Journal of Climatology*, **33** (7): 1720–1729.

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K. and Takahashi, K. (2011). The Japanese 55-year reanalysis "JRA-55": an interim report. *Sola*, **7**: 149–152.

Eckstein, D., Künzel, V., Schäfer, L. and Winges, M. (2019). *Global Climate Risk Index 2020.* Bonn: Germanwatch.

Edijatno, and C. Michel (1989), Un modèle pluie-débit journalier à trois paramètres, La Houille Blanche, 113-121.

Edijatno, N. O. Nascimento, X. Yang, Z. Makhlouf, and C. Michel (1999), GR3J: a daily watershed model with three free parameters, *Hydrological Sciences Journal,* **44** (3), 263-277.

Efstratiadis, A., Nalbantis, I., Koukouvinos, A., Rozos, E. and Koutsoyiannis, D. (2008). HYDROGEIOS: a semi-distributed GIS-based hydrological model for modified river basins. *Hydrology & Earth System Sciences*, **12** (4): 989-1006.

Efron, B. and Tibshirani, R.J. (1994). *An Introduction to the Bootstrap*. 1st ed. Monographs on Statistics and Applied Probability, 57. Boca Raton, Florida: Chapman and Hall/CRC press.

Emerton, R., Cloke, H.L., Stephens, E.M., Zsoter, E., Woolnough, S.J. and Pappenberger, F. (2017). Complex picture for likelihood of ENSO-driven flood hazard. *Nature Communications,* **8** (1): 1-9.

Emerton, R., Zsoter, E., Arnal, L., Cloke, H.L., Muraro, D., Prudhomme, C., Stephens, E.M., Salamon, P. and Pappenberger, F. (2018). Developing a global operational seasonal hydro-

meteorological forecasting system: GloFAS-Seasonal v1. *Geoscientific Model Development*, **11** (8): 3327–3346.

Emerton, R. E., Stephens, E.M., Pappenberger, F., Pagano, T.C., Weerts, A. H., Wood, A.W., Salamon, P., Brown, J.D., Hjerdt, N., Donnelly, C., Baugh, C.A. and Cloke, H.L. (2016). Continental and global scale flood forecasting systems. *Water*, **3** (3): 391–418. doi: 10.1002/wat2.1137.

English, S., McNally, T., Bormann, N., Salonen, K., Matricardi, M., Moranyi, A., Rennie, M., Janisková, M., Di Michele, S., Geer, A. and Di Tomaso, E. (2013). *Impact of satellite data*. Reading: European Centre for Medium-Range Weather Forecasts Technical Memorandum, No. 71.

Van Esse, W. R., Perrin, C., Booij, M.J., Augustijn, D.C.M., Fenicia, F., Kavetski, D. and Lobligeois, F. (2013). The influence of conceptual model structure on model performance: a comparative study for 237 French catchments. *Hydrology and Earth System Sciences*, **17** (10): 4227–4239.

Essou, G. R. C., Brissette, F. and Lucas-Picher, P. (2017). The use of reanalyses and gridded observations as weather input data for a hydrological model: of performances of simulated river flows based on the density of weather stations. *Journal of Hydrometeorology*, **18** (2): 497–513.

Essou, G.R.C., Sabarly, F., Lucas-Picher, P., Brissette, F. and Poulin, A. (2016). Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling? *Journal of Hydrometeorology*, **17** (7): 1929–1950.

Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.D., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M. and Tarboton, D. (2016). An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology* **53: 7**45–60. doi: 10.1016/j.jhydrol.2016.03.026.

Fazal, M.A., Imaizumi, M., Ishida, S., Kawachi, T. and Tsuchihara, T. (2005). Estimating groundwater recharge using the SMAR conceptual model calibrated by genetic algorithm. *Journal of Hydrology*, **303** (1–4): 56–78.
Fetter, C.W. (2018). *Applied hydrogeology*. 4th ed. Long Grove, Illinois: Waveland Press.

Ficchì, A., Perrin, C. and Andréassian, V. (2016). Impact of temporal resolution of inputs on hydrological model performance: analysis based on 2400 flood events. *Journal of Hydrology*, **538:** 454–470.

Ficchì, A., Perrin, C. and Andréassian, V. (2019). Hydrological modelling at multiple sub-daily time steps: model improvement via flux-matching. *Journal of Hydrology,* **575**: 1308-1327. doi: https://doi.org/10.1016/j.jhydrol.2019.05.084.

Ficchì, A. and Stephens, L. (2019). Climate variability alters flood timing across Africa. *Geophysical Research Letters*, **46** (15): 8809–8819.

Finney, D.L., Marsham, J.H., Walker, D.P., Birch, C.E., Woodhams, B.J., Jackson, L.S. and Hardy, S. (2020). The effect of westerlies on East African rainfall and the associated role of tropical cyclones and the Madden–Julian Oscillation. *Quarterly Journal of the Royal Meteorological Society,* **146**: 647-664.

Flügel, W. (1995). Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany. *Hydrological Processes*, **9** (3-4): 423–436.

Fortin, V., Roy, G., Donaldson, N. and Mahidjiba, A. (2015). Assimilation of radar quantitative precipitation estimations in the Canadian Precipitation Analysis (CaPA). *Journal of Hydrology*, **531**: 296–307.

Foster, S.B. and Allen, D.M. (2015). Groundwater—surface water interactions in a mountain-to-coast watershed: effects of climate change and human stressors. *Advances in Meteorology, Special Issue:* Article ID 861805.

Franks, S.W., Gineste, P., Beven, K.J. and Merot, P. (1998). On constraining the predictions of a distributed model: the incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resources Research*, **34** (4): 787–797.

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C. and Martius, O. (2015). Flood triggering in Switzerland: the role of daily to monthly preceding precipitation. *Hydrology and Earth System Sciences*, **19** (9): 3903–3924.

Fuka, D.R., Walter, M.T., MacAlister, C., Degaetano, A.T., Steenhuis, T.S. and Easton, Z.M. (2014). Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrological Processes*, **28** (22): 5613–5623.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - new environmental record for monitoring extremes. *Scientific Data*, **2:** 1–21. doi: 10.1038/sdata.2015.66.

Gabellani, S., Boni, G., Ferraris, L., Von Hardenberg, J. and Provenzale, J. (2007). Propagation of uncertainty from rainfall to runoff: a case study with a stochastic rainfall generator. *Advances in Water Resources,* **30** (10): 2061–2071.

Gao, J., Holden, J. and Kirkby, M. (2015). A distributed TOPMODEL for modelling impacts of land-cover change on river flow in upland peatland catchments. *Hydrological Processes*, **29** (13): 2867–2879.

Gao, J., Holden, J. and Kirkby, M. (2017). Modelling impacts of agricultural practice on flood peaks in upland catchments: application of the distributed TOPMODEL. *Hydrological processes*, **31** (23): 4206–4216.

Gardelin, M. and Lindström, G. (1997). Priestley-Taylor Evapotranspiration in HBV-Simulations: Paper presented at the Nordic Hydrological Conference (Akureyri, Iceland-August 1996). *Hydrology Research*, 28 (4–5): 233–246.

Gassman, P.W., Reyes, M.R., Green, C. H. and Arnold, J.G. (2007). The soil and water assessment tool: historical development, applications, and future research directions. *Transactions of the ASABE (American Society of Agricultural and Biological Engineers)*, **50** (4): 1211-1250.

Gathee, N.G. and Odera, P.A. (2015). Comparison of ANN-NARX, SMAR-LTF and LPM-LTF flood forecasting models in the Nzoia basin, Kenya. *Journal of Sustainable Research in Engineering*, **2** (2): 36–44.

Giuntoli, I., Renard, B. and Lang, H. (2019). Floods in France. In: Z. W. Kundzewicz (ed.) *Changes in Flood Risk in Europe* [Preprint]. 1st e-book ed. https://doi.org/10.1201/b12348.

Githui, F., Mutua, F. and Bauwens, W. (2009). Estimating the impacts of land-cover change on runoff using the soil and water assessment tool (SWAT): case study of Nzoia catchment, Kenya/Estimation des impacts du changement d'occupation du sol sur l'écoulement à l'aide de SWAT: étude du cas du basin. *Hydrological Sciences Journal*, **54** (5): 899–908.

Gleixner, S., Demissie, T. and Diro, G.T. (2020). Did ERA5 improve temperature and precipitation reanalysis over East Africa? *Atmosphere,* **11** (9): 996.

Golian, S. and Murphy, C. (2021). Evaluation of sub-selection methods for assessing climate change impacts on low-flow and hydrological drought conditions. *Water Resources Management*, **35** (1): 113–133.

Gosset, M.*,* Viarre, J., Quantin, G. and Alcoba, M.(2013). Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Quarterly Journal of the Royal Meteorological Society*, **139** (673): 923–940. doi: 10.1002/qj.2130.

Goswami, M. and O'Connor, K.M. (2010). A 'monster' that made the SMAR conceptual model 'right for the wrong reasons'. *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, **55** (6): 913–927.

Graham, D.N. and Butts, M.B. (2005). Flexible, integrated watershed modelling with MIKE SHE. In: V.P. Singh and D.K. Frevert (eds) *Watershed models*, Boca Raton, Florida: CRC Press, pp. 245–272.

Graham, R.J.*,* Yun, W.-T.,Kim, J., Kumar, A., Jones, D., Bettio, L., Gagnon, N., Kolli, R.K. and Smith, D (2011).   Long-range forecasting and the Global Framework for Climate Services. *Climate Research*, **47** (1–2): 47–55.

Grayson, R.B., Moore, I.D. and McMahon, T.A. (1992). Physically based hydrologic modeling: 2. Is the concept realistic? *Water*, **28** (10): 2659–2666.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M. and Srinivasan, R. (2006). A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*, **324** (1–4): 10–23.

Gribbon, K.T. and Bailey, D.G. (2004). A novel approach to real-time bilinear interpolation. *Proceedings. DELTA 2004. Second IEEE International Workshop on Electronic Design, Test and Applications*:126–131.

Grijsen, J.G., Brown, C., Tarhule, A., Ghile, Y.B. and Taner, (2012). Climate risk assessment for water resources development in the Niger River Basin. *Water Partnership Program (WPP)/TWIWA, HydroPredict*, 12.

Gudoshava, M., Misiani, H.O., Segele, Z.T., Jain, S., Ouma, J.O., Otieno, G., Anyah, R., Indasi, V.S., Endris, H.S., Osima, S., Lennard, C., Zaroug, M., Mwangi, E., Nimusiima, A., Kondowe, A., Ogwang, B., Artan, G. and Atheru, Z. (2020). Projected effects of 1.5 C and 2 C global warming levels on the intra-seasonal rainfall characteristics over the Greater Horn of Africa. *Environmental Research Letters*, **15** (3): 034037.

Gumindoga, W., Rwasoka, D.T. and Murwira, A. (2011). Simulation of streamflow using TOPMODEL in the Upper Save River catchment of Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C* **36** (14–15): 806–813.

Guo, B. (2018). Applicability assessment and uncertainty analysis of multi-precipitation datasets for the simulation of hydrologic models. *Water*, **10** (11): 1611.

Gupta, H.V., Sorooshian, S. and Yapo, P.O. (1998). Toward improved calibration of hydrologic models: and noncommensurable measures of information. *Water Resources Research*, **34** (4):751–763.

Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F. (2009). Decomposition of the mean squared error and NSE performance criteria: for improving hydrological modelling. *Journal of Hydrology*, **377** (1): 80–91. doi: https://doi.org/10.1016/j.jhydrol.2009.08.003.

Gupta, H.V, Beven, K.J. and Wagener, T. (2006). Model calibration and uncertainty estimation. In: M.G. Anderson and J. McDonnell (eds) *Encyclopedia of Hydrologic Sciences,* Chichester: Wiley. Wiley Online Library.

Gupta, H.V, Wagener, T. and Liu, Y. (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, **22** (18): 3802–3813.

Haddeland, I., Clark, D.B., Franssen, W., Ludwig, F., Voss, F., Arnell, N.W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S.N., Hagemann, S., Hanasaki, N., Harding. R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacks, T., Viterbo, P., Weedon, G.P. and Yeh, P. (2011). Multimodel estimate of the global terrestrial water balance: setup and first results. *Journal of Hydrometeorology*, **12** (5): 869–884.

Hamed, K.H. (2008). Trend detection in hydrologic data: the Mann–Kendall trend test under the scaling hypothesis. *Journal of Hydrology*, **349** (3–4): 350–363.

Hamon, W.R. (1960). Estimating potential. B.S. Thesis, Cambridge, Massachusetts: Massachusetts Institute of Technology.

Harada, Y., Kamahori, H., Kobayashi, C., Endo, H., Kobayashi, S., Ota, Y., Onoda, H., Onogi, K., Miyaoka, K. and Takahshi, K. (2016). The JRA-55 Reanalysis: representation of

atmospheric circulation and climate variability. *Journal of the Meteorological Society of Japan. Ser.II,* 94 (3): 269-302.

Harbaugh, A.W. (2005). *MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process.* Reston, VA: US Department of the Interior, US Geological Survey.

Hargreaves, G.H. and Samani, Z.A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture,***1:** 96-99.

Harrigan, S., Cloke, H. and Pappenberger, F. (2020). Innovating global hydrological prediction through an Earth system approach. *WMO Bulletin*, **69**:1.

Hartman, A.T. (2018). An analysis of the effects of temperatures and circulations on the strength of the low-level jet in the Turkana Channel in East Africa. *Theoretical and Applied Climatology*, **132** (3): 1003–1017.

Hattermann, F.F., Krysanova, V., Gosling, S., Dankers, R., Daggupati, P., Donnelly, C., Flörke, M., Huang, S., Motovilov, Y., Buda, S., Yang, T., Müller, C., Leng, G., Tang, Q., Portmann, F.T., Hagemann, S., Gerten, D., Wada, Y., Masaki, Y., Alemayehu,T., Satoh, Y., Samaniego, L. (2017). Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. *Climatic Change*, **141** (3): 561–576.

Hengade, N. and Eldho, T.I. (2016). Assessment of LULC and climate change on the hydrology of Ashti Catchment, India using VIC model. *Journal of Earth System Science*, **125** (8): 1623–1634.

HM Government (2016). National Flood Resilience Review (September). https://www.gov.uk/government/publications/national-flood-resilience-review

Herrera, P.A., Marazuela, M.A. and Hofmann, T. (2022). Parameter estimation and uncertainty analysis in hydrological modelling. *Wiley Interdisciplinary Reviews, Water*, **9** (1): e1569.

Hersbach, H., de Rosnay, P., Bell, B., Alonso-Balmaseda. M., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J.-R., de Boisséson, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P/. Dee. D., Dragani, R., Diamantakis, M., Flemming. J., Forbes, R., Geer, A.J., Haiden, T., Hólm, E., Haimberger, L., Hogan, R., Horányi, A., Janiskova, M., Laloyaux, P., Lopez, P., Munoz-Sabater, J., Peubey, C., Radu, R., Richardson, D., Thépaut. J-N., Vitart, F., Yang, X., Zsótér, E. and Zuo, H. (2018). *Operational global reanalysis: progress, future directions and synergies with NWP.* Reading: European Centre for Medium Range Weather Forecasts. doi: 10.21957/tkic6g3wm.

Hipel, K.W. and McLeod, A.I. (1978). Preservation of the rescaled adjusted range: 3. Fractional Gaussian noise algorithms. *Water Resources Research*, **14** (3) 517–518.

Hession, S.L. and Moore, N. (2011). A spatial regression analysis of the influence of topography on monthly rainfall in East Africa. *International Journal of Climatology*, **31** (10): 1440–1456.

Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griessbach, S., Heng, Y., Konopka, P., Müller, R. and Vogel, B. (2019). From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations. *Atmospheric Chemistry and Physics*, **19** (5): 3097–3124.

Holden, J., Kirkby, M.J., Lane, S.N., Milledge, D.G., Brookes, C.J., Holden, V. and McDonald, A.T. (2008). Overland flow velocity and roughness properties in peatlands. *Water Resources Research*, **44** (6): eWO6415. doi: 10.1029/2007WR006052.

Hong, Y., Hsu, K.-L., Sorooshian, S. and Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology*, **43** (12): 1834–1853.

Hrachowitz, M. and Weiler, M. (2011). Uncertainty of precipitation estimates caused by sparse gauging networks in a small, mountainous watershed. *Journal of Hydrologic Engineering,* **16**: 460-471.

Hu, Z., Hu, Q., Zhang, C., Chen, X. and Li, Q. (2016). Evaluation of reanalysis spatially interpolated and satellite remotely sensed precipitation data sets in central Asia. *Journal of Geophysical Research: Atmospheres*, **121** (10): 5648-5663. doi: 10.1002/2016JD024781.

Huang, S., Huang, Q., Leng, G. and Liu, S. (2016.) A nonparametric multivariate standardized drought index for characterizing socioeconomic drought: A case study in the Heihe River Basin. *Journal of Hydrology*, **542**: 875–883. Available at: https://doi.org/10.1016/j.jhydrol.2016.09.059.

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E., Sorooshian, S., Wang, J. and Xie, P. (2015). First results from the integrated multi-satellite retrievals for GPM (IMERG). *EGU General Assembly Conference Abstracts*: 7034.

Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B., Susskind, J., Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B. and Susskind, J. (2001). Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *Journal of Hydrometeorology*, **2** (1): 36–50. Available at: https://doi.org/10.1175/1525-7541(2001)002<0036: GPAODD>2.0.CO;2.

Huho, J.M. and Kosonei, R.C. (2014). Understanding extreme climate events for economic development in Kenya. *IOSR Journal of Environmental Science*, *Toxicology and Food Technology,* **8** (2): 14–25. doi: 10.9790/2402-08211424.

Van Huijgevoort, M.H.J., van Lanen, H.A.J., Teuling, A.J. and Uijlenhoet, R. (2014). Identification of changes in hydrological drought characteristics from a multi-GCM driven ensemble constrained by observed discharge. *Journal of Hydrology*, **512:** 421–434.

Humphrey, G.B., Gibbs, M.S., Dandy, G.C. and Maier, H.R. (2016). A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology*, **540:** 623–640.

Hunink, J.E., Niadas, I.A., Antonoropoulos, P., Droogers, P. and De Vente, J. (2013). Targeting of intervention areas to reduce reservoir sedimentation in the Tana catchment (Kenya) using SWAT. *Hydrological Sciences Journal*, **58** (3): 600–614.

Im, S., Kim, H., Kim, C. and Jang, C. (2009). Assessing the impacts of land use changes on watershed hydrology using MIKE SHE. *Environmental Geology*, **57** (1): 231-239.

ISDR (2004). *Living with risk: review of disaster reduction initiatives*. New York and Geneva: United Nations Publications.

Ishak, E. and Rahman, A. (2019). Examination of changes in flood data in Australia. *Water*, **11**(8): 1734.

Jain, S.K. and Singh, V.P. (2019). Hydrological Cycles, Models, and Applications to Forecasting BT. In: Q. Duan, F. Pappenberger, A. Wood, H.L. Cloke, J.C. Schaake (eds) *Handbook of Hydrometeorological Ensemble Foreasting,* Berlin, Heidelberg: Springer, pp. 311–339. doi: 10.1007/978-3-642-39925-1_20.

Jajarmizadeh, M., Harun, S. and Salarpour, M. (2012). A review on theoretical consideration and types of models in hydrology. *Journal of Environmental Science and Technology*, 5 (5): 249-261.

Jakeman, A.J., Letcher, R.A. and Norton, J.P. (2006). Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software*, **21** (5): 602–614.

Jayathilake, D.I. (2019). *Exploring the Sensitivity of Hydrologic Models to Potential Evapotranspiration Inputs.* Doctoral Dissertation. New York: Clarkson University.

Jayatilaka, C.J., Storm, B. and Mudgway, L.B. (1998). Simulation of water flow on irrigation bay scale with MIKE- *Journal of Hydrology*, **208:** 108-130.

Johnson, B.E., Zhang, Z. and Downer, C.W. (2013). Watershed scale physically based water flow, sediment and nutrient dynamic modeling system. In: B. Fu, B. Jones (eds) *Landscape Ecology for Sustainable Environment and Culture,* Dordrecht: Springer, pp. 145–171.

Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J. and Wang, J. (2021). Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland. *Journal of Hydrology*, **595**: 125660.

Joyce, R.J., Janowiak, J.E., Arkin, P.A. and Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5 (3): 487–503.

Kachroo, R.K. (1992). River flow forecasting. Part 5. Applications of a conceptual model. *Journal of Hydrology*, **133** (1–2): 141–178.

Karlsson, I.B., Sonnenberg, T.O., Refsgaard, J.C., Trolle, D., Børgesen, C.D., Olesen, J.E., Jeppesen, E. and Jensen, K.H (2016). Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change. *Journal of Hydrology*, **535:** 301–317.

Katuva, J.M., Omuto, C.T. and Obiero, J.P.O. (2018). Water allocation, assessment and hydrological simulation on Mukurumudzi River Basin in Kenya. In: C. M. Ondieki and J. U. Kitheka (eds) *Hydrology and Best Practices for Managing Water Resources in Arid and Semi-Arid Lands*, Hershey, Pennsylvania: IGI Global, pp. 51–69.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P. and Thielen, J. (2016). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software*, **75:** 68–76.

Kavetski, D. (2019). Parameter estimation and predictive uncertainty quantification in hydrological modelling BT. In: Q. Duan, F. Pappenberger, A. Wood, H.L. Cloke and J. Schaake (eds) *Handbook of hydrometeorological ensemble forecasting*, Berlin, Heidelberg: Springer, pp. 481–522. doi: 10.1007/978-3-642-39925-1_25.

Keilholz, P., Disse, M. and Halik, Ü., (2015). Effects of land use and climate change on groundwater and ecosystems at the middle reaches of the Tarim River using the MIKE SHE integrated hydrological model. *Water*, *7*(6), 3040-3056.

Kendall, M. (1975). *Rank correlation methods*. 4th ed. London: Charles Griffin.

Kenya: Floods - Apr 2012 | ReliefWeb (no date). Available at: https://reliefweb.int/disaster/ff-2012-000062-ken (Accessed: 3 March 2020).

Kenya: Floods - Aug 2010 | ReliefWeb (no date). Available at: https://reliefweb.int/disaster/fl-2010-000174-ken (Accessed: 3 March 2020).

Kenya: Floods and Landslides - Oct 2019 | ReliefWeb (no date). Available at: https://reliefweb.int/disaster/fl-2019-000138-ken (Accessed: 3 March 2020).

Kenya Food Security Outlook Update, December 2019 – May 2020 - Kenya | ReliefWeb (no date). Available at: https://reliefweb.int/report/kenya/kenya-food-security-outlook-update-december-2019-may-2020 (Accessed: 3 March 2020).

Kenya – Tragedy After Bus Swept Away by Floods in Kitui County, December 2021: https://floodlist.com/africa/kenya-floods-december-2021. Accessed 30/06/2022.

Khan, M.H. (1995). A study of two conceptual models for Rainfall-Runoff system. *Journal of irrigation engineering and rural planning,* **28**: 4–18.

Khan, S.I., Adhikari, P., Hong, Y., Vergara, H., Adler, R.F, Policelli, F., Irwin, D., Korme, T. and Okello, L. (2011). Hydroclimatology of Lake Victoria region using hydrologic model and satellite remote sensing data. *Hydrology and Earth System Sciences*, **15** (1): 107–117.

Khu, S.T. and Madsen, H. (2005). Multiobjective calibration with Pareto preference ordering: an application to rainfall-runoff model calibration. *Water Resources Research*, **41**: W03004, doi: 10.1029/2004WR003041.

Kilavi, M., MacLeod, D., Ambani, M., Robbins, J., Dankers, R., Graham, R., Titley, H., Salih, A.A.M. and Todd, M.C. (2018). Extreme rainfall and flooding over central Kenya including

Nairobi city during the long-rains season 2018: causes, predictability, and potential for early warning and actions. *Atmosphere*, **9** (12): 472. doi:10.3390/atmos9120472.

Kiluva, V.M., Mutua, F., Makhanu, S.K. and Ong'or, B.T.I. (2011). Application of the geological stream flow and Muskingum Model in the Yala River Basin, Kenya. *Journal of Agricultural Science and Technology,* **13** (2): 92-108.

Kim, K.-Y., Kim, J., Boo, K.-O., Shim, S. and Kim, Y. (2019). 'Intercomparison of precipitation datasets for summer precipitation characteristics over East Asia', *Climate Dynamics*, **52** (5): 3005–3022.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M. and Kousky, V. (2001). The NCEP–NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bulletin of the American Meteorological society*, 82(2): 247–268.

Kinuthia, J.H. (1992). Horizontal and vertical structure of the Lake Turkana jet. *Journal of Applied Meteorology and Climatology*, **31** (11): 1248-1274.

Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal,* **31** (1): 13-24.

Kling, H., Fuchs, M. and Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, **424**: 264–277.

van der Knijff, J.M., Younis, J. and de Roo, A.P.J. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, **24** (2): 189–212. doi: 10.1080/13658810802549154.

Knoben, W.J.M., Freer, J.E. and Woods, R.A. (2019). Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, **23** (10): 4323–4331.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, H., Kamahori, H., Kobayashi, C., Endo, H., Kengo, M. and Takahashi, K. (2015). The JRA-55 reanalysis: general specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, **93** (1): 5–48. doi: 10.2151/jmsj.2015-001.

Kodja, D. J., Mahé, G., Amoussou, E., Boko, M. and Paturel, J. (2018). Assessment of the performance of Rainfall-Runoff Model GR4J to simulate streamflow in Ouémé Watershed at Bonou's outlet (West Africa). Preprints. org. doi: 10.20944/preprints201803. 0090.v1.

Koirala, S., Yeh, P.J.F., Hirabayashi, Y., Kanae, S. and Oki, T., 2014. Global-scale land surface hydrologic modeling with the representation of water table dynamics. *Journal of Geophysical Research: Atmospheres*, **119**(1), pp.75-89.

Koukoula, M., Nikolopoulos, E.I., Dokou, Z. and Anagnostou, E.N. (2020). Evaluation of global water resources reanalysis products in the Upper Blue Nile River Basin. *Journal of Hydrometeorology*, **21** (5): 935–952.

Koutsoyiannis, D. and Montanari, A. (2015). Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal*, **60** (7–8): 1174–1183.

Krysanova, V., Bronstert, A. and Müller-Wohlfeil, D.-I. (1999). Modelling river discharge for large drainage basins: from lumped to distributed approach. *Hydrological Sciences Journal*, **44** (2): 313–331.

Kuczera, G. and Franks, S.W. (2002). Testing hydrologic models: fortification or falsification? In: V.P. Singh and D.K. Frevert (eds) *Mathematical models of large watershed hydrology*, Highlands Ranch, Colorado: Water Resources Publications, pp. 141-186.

Kundzewicz, Z.W., Kanae, S., Seneviratne, S.I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L.M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G.R., Kron, W., Benito, G., Honda, Y., Takahashi, K. and Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal*, **59** (1): 1–28.

Kyambia, M.M. and Mutua, B.M. (2015). Detection of trends in extreme streamflow due to climate variability in the Lake Naivasha basin, Kenya. *International Journal of River Basin Management*, **13** (1): 97–103.

Lakew, H. B., Moges, S.A. and Asfaw, D.H. (2020). Hydrological performance evaluation of multiple satellite precipitation products in the upper Blue Nile basin, Ethiopia. *Journal of Hydrology: Regional Studies*, (27): 100664.

Van Lanen, H.A.J., Wanders, N., Tallaksen, L.M. and Van Loon, A.F. (2013). Hydrological drought across the world: impact of climate and physical catchment structure. *Hydrology and Earth System Sciences*, **17**: 1715–1732.

Langat, P.K., Kumar, L. and Koech, R. (2017). Temporal variability and trends of rainfall and streamflow in Tana River Basin, Kenya. *Sustainability*, **9** (11): 1963.

Langat, P.K., Kumar, L., Koech, R. and Ghosh, M.K. (2019). Hydro-morphological characteristics using flow duration curve, historical data and remote sensing: effects of land use and climate. *Water*, **11** (2): 309.

Langat, P.K., Kumar, L., Koech, R. and Ghosh, M.K. (2020). Characterisation of channel morphological pattern changes and flood corridor dynamics of the tropical Tana River fluvial systems, Kenya. *Journal of African Earth Sciences*, **163**: 103748.

Lavers, D.A., Villarini, G., Allan, R.P., Wood, E.F. and Wade, A.J. (2012). The detection of atmospheric rivers in atmospheric reanalyses and their links to British winter floods and the large-scale climatic circulation. *Journal of Geophysical Research: Atmospheres*, **117** (D20106). doi; 10.1029/2012JD018027.

Lavergne, T., Kern, S., Aaboe, S., Derby, L., Dybkjaer, G., Garric, G., Heil, P., Hendricks, S., Holfort, J., Howell, S., Key, J.R., Lisser, J.L., Maksym, T., Maslowski, W., Meier, W., Sabater, J.M., Nicolas, J., Özsoy, B., Rabe, B., Rack, W., Raphael, M., de Rosnay, P., Smolyanitsky, V., Tietsche, S., Ukita, J., Vichi, M., Wagner, P., Willmes, S. and Zhao, X. (2022). A New Structure for the Sea Ice Essential Climate Variables of the Global Climate Observing System.

Bulletin of the American Meteorological Society [Preprint]. Doi: 10.1175/BAMS-D-21-0227.1.

Le, A.M. and Pricope, N.G. (2017). Increasing the accuracy of runoff and streamflow simulation in the Nzoia Basin, Western Kenya, through the incorporation of satellite-derived CHIRPS data. *Water*, **9** (2). doi: 10.3390/w9020114.

Leavesley, G.H., Lichty, RW., Troutman, B.M. and Saindon, L.G. (1983). *Precipitation-Runoff Modeling System: User's Manual: US Geological Survey Water-Resources Investigations Report 83-4238*. Denver, Colorado: U.S. Geological Survey.

Legesse, D., Vallet-Coulomb, C. and Gasse, F. (2003). Hydrological response of a catchment to climate and land use changes in Tropical Africa: case study South Central Ethiopia. *Journal of Hydrology*, **275** (1–2): 67–85.

Lehning, M., Völksch, I., Gustafsson, D., Nguyen, T.A., Stähli, M. and Zappa, M. (2006). ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. *Hydrological Processes*, **20** (10): 2111–2128.

Lemma, E., Upadhyaya, S. and Ramsankaran, R. (2019). Investigating the performance of satellite and reanalysis rainfall products at monthly timescales across different rainfall regimes of Ethiopia. *International Journal of Remote Sensing*, **40** (10): 4019–4042.

Le Moine, N., 2008. *The surface watershed seen from underground: a way to improve the performance and realism of rainfall-runoff models?* Dissertation for Doctorate in Geosciences and Natural Resources, Pierre and Marie Curie Paris VI University.

Lerat, J., Thyer, M., McInerney, D., Kavetski, D., Woldemeskel, F., Pickett-Heaps, C., Shin, D. and Feikema, P. (2020). A robust approach for calibrating a daily rainfall-runoff model to monthly streamflow data. *Journal of Hydrology*, **591**: 125129.

Liang, G. C. (1992). A note on the revised SMAR model, in *Memorandum to the River Flow Forecasting Workshop Group*, Department of Engineering Hydrology, National University College of Ireland, Galway. Unpublished.

Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, **99** (D7): 14415–14428.

Lin, B., Liu, Q.J., Shang, H., Wang, Y.W. and Sui, X. (2014). Application of coupled MIKE 11/NAM model in Naoli River Basin, northeastern China. *Journal of Beijing Forestry University*, **36** (5): 99–108.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, **201**(1–4): 272–288.

Liu, H., Wang, Y., Zhang, C., Chen, A.S. and Fu, G. (2018). Assessing real options in urban surface water flood risk management under climate change. *Natural Hazards*, **94** (1): 1–18.

Lone, S.A. and Ahmad, A. (2020). COVID-19 pandemic–an African perspective. *Emerging microbes & infections*, **9** (1): 1300-1308.

Lu, W., Wang, W., Shao, Q., Yu, Z., Hao, Z., Xing, W., Yong, B. and Li, J. (2018). Hydrological projections of future climate change over the source region of Yellow River and Yangtze River in the Tibetan Plateau: a comprehensive assessment by coupling RegCM4 and VIC model. *Hydrological processes*, 32 (13): 2096–2117.

Lu, J., Sun, G., McNulty, S.G. and Amatya, D.M. (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern United States 1. *Journal of the American Water Resources Association*, **41** (3): 621–633.

Di Luzio, M., Srinivasan, R., Arnold, J.G. and Neitsch, S.L. (2002). *ArcView interface for SWAT2000: User's Guide*. Temple, Texas: Texas Water Resources Institute.

Ma, L., He, C., Bian, H. and Sheng, L. (2016). MIKE SHE modeling of ecohydrological processes: Merits, applications, and challenges. *Ecological Engineering,* **96** (5): 137-149. doi: 10.1016/j.ecoleng.2016.01.008.

MacLeod, D.A., Dankers, R., Graham, R., Guigma, K., Jenkins, L., Todd, M.C., Kiptum, A., Kilavi, M., Njogu, A. and Mwangi, E. (2021). Drivers and subseasonal predictability of heavy rainfall in equatorial East Africa and relationship with flood risk. *Journal of Hydrometeorology*, **22** (4): 887-903.

McMillan, H., Krueger, T. and Freer, J. (2012). Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, **26** (26): 4078–4111.

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

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

Mahto, S.S. and Mishra, V. (2019). Does ERA-5 outperform other reanalysis products for hydrologic applications in India? *Journal of Geophysical Research: Atmospheres*, **124** (16): 9423–9441.

Malag, A., Bouraoui, F. and Roo, A.De (2018). Diagnosis and Treatment of the SWAT Hydrological Response Using the Budyko Framework. doi: 10.3390/su10051373.

Mangini, W., Viglione, A., Hall, J., Hundecha, Y., Ceola, S., Montanari, A., Rogger, M., Salinas, J.L., Borzì, I. and Parajka, J. (2018). Detection of trends in magnitude and frequency of flood peaks across Europe. *Hydrological Sciences Journal*, **63** (4): 493–512. Available at: https://doi.org/10.1080/02626667.2018.1444766.

Mann, H.B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, **13** (3): 245–259.

Mardia, K.V. (1975). Statistics of directional data. *Journal of the Royal Statistical Society: Series B (Methodological)*, **37** (3): 349–371.

Marwick, T.R., Tamooh, F., Ogwoka, B., Teodoru, C., Borges, A.V., Darchambeau, F. and Bouillon, S. (2014). Dynamic seasonal nitrogen cycling in response to anthropogenic N loading in a tropical catchment, Athi–Galana–Sabaki River, Kenya. *Biogeosciences*, **11** (2): 443–460.

Mathevet, T., Michel, C., Andréassian, V. and Perrin, C. (2006). A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. In: V.A. Andréassian, A. Hall, N. Chaninian and J. Schaake (eds) *Large sample basin experiments for hydrological parameterization: results of the Model Parameter Experiment - MOPEX*. 1st ed. Paris: International Association of Hydrological Sciences Publications, **307**: 211-219.

Mediero, L., Kjeldsen, T.R., Macdonald, N., Kohnova, S., Merz, B., Vorogushyn, S., Wilson, D., Alburquerque, T., Blöschl, G. and Bogdanowicz, E. (2015). Identification of coherent flood regions across Europe by using the longest streamflow records. *Journal of Hydrology*, **528**: 341–360.

Mega, T., Ushio, T., Kubota, T., Kachi, M., Aonashi, K. and Shige, S. (2014). Gauge adjusted global satellite mapping of precipitation (GSMaP_Gauge) in *2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)*. IEEE: 1–4.

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Mendoza, P., Clark, M. and Uijlenhoet, R., (2017). Subjective modelling decisions significantly impact the simulation of hydrological extremes. In *EGU General Assembly Conference Abstracts* **24** (24).

Menne, M.J, Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R.S., Gleason, B.E and Houston, T.G. (2018). The global historical climatology network monthly temperature dataset, version 4. *Journal of Climate*, **31** (24): 9835–9854.

Meresa, H., Murphy, C., Fealy, R. and Golian, S. (2021). Uncertainties and their interaction in flood hazard assessment with climate change. *Hydrology and Earth System Sciences*, **25** (9): 5237–5257.

Merz, B. and Thieken, A.H. (2005). Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, **309** (1): 114–132. Available at: https://doi.org/https://doi.org/10.1016/j.jhydrol.2004.11.015.

Merz, R., Parajka, J. and Blöschl, G. (2011). Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resources Research*, **47** (2). https://doi.org/10.1029/2010WR009505.

Michel, C. (1983). Que peut-on faire en hydrologie avec modèle conceptuel à un seul paramètre? *La Houille Blanche*, (1): 39–44. https://doi.org/10.1051/lhb/1983004, 1983.

Miller, D.A. and White, R.A. (1998). A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions*, **2** (2): 1–26.

Moges, E., Demissie, Y., Larsen, L. and Yassin, F., (2020). Review: Sources of Hydrological Model Uncertainties and Advances in Their Analysis. *Water*, (**13**) 28.

Moore, R.J. (2007). The PDM rainfall-runoff model. *Hydrology and Earth System Sciences Discussions*, **11** (1): 483–499.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R,L., Harmel, R.D. and Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, **50** (3): 885–900.

Moriasi, D.N., Gitau, M.W., Pai, N. and Daggupati, P. (2012). Hydrologic and water quality models: use, calibration, and validation. *Transactions of the ASABE*, **55** (4): 1241–1247.

Mostafaie, A., Forootan, E.A., Safari, A. and Schumacher, M. (2018). Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data. *Computational Geosciences*, **22** (3):789–814. doi: 10.1007/s10596-018-9726-8.

Moussa, R. and Chahinian, N. (2009). Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events. *Hydrology and Earth System Sciences*, **13** (4): 519-535.

Msaddek, M., Kimbowa, G. and Garouani, A.E. (2020). Hydrological modeling of Upper OumErRabia Basin (Morocco), comparative study of the event-based and continuous-process HEC-HMS model methods. *Computational Water, Energy, and Environmental Engineering*, **9** (04):159.

Mubareka, S., Maes, J., Lavalle, C. and de Roo, A. (2013). Estimation of water requirements by livestock in Europe. *Ecosystem Services*, **4**: 139–145.

Muleta, M.K. and Nicklow, J.W. (2005). Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of Hydrology*, **306** (1–4): 127–145.

Muli, C. (2007). Modelling the Effects of Deforestation on Stream Flows in Arror River Basin-Kenya. Tema vatten i natur och samhälle, MA thesis, Linköping University.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S. and Hersbach, H. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, **13** (9): 4349–4383.

Mutie, S.M., Home, P., Gadain, H., Mati, B. and Gathenya, J. (2006). Evaluating land use change effects on river flow using USGS geospatial stream flow model in Mara River Basin, Kenya. In: *Proceedings of the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover, September 28-30, 2006*, Bonn: Center for Remote Sensing of Land Surfaces, pp. 141-148.

Mutie, S.M. (2019). Modelling the influence of land use change on the hydrology of Mara River using USGS Geospatial Stream Flow Model. MSc Thesis, JKUAT-COETEC.

Mutimba, S., Mayieko, S., Olum, P. and Wanyatma, K. (2010). Climate change vulnerability and adaptation preparedness in Kenya. Nairobi: Heinrich Böll Stiftung Foundation.

Mwangi, H.M., Julich, S., Patil, S.D., McDonald, M.A. and Feger, K.-H. (2016). Relative contribution of land use change and climate variability on discharge of upper Mara River, Kenya. *Journal of Hydrology: Regional Studies*, **5**: 244–260.

Najafi, M.R., Moradkhani, H. and Piechota, T.C. (2012). Ensemble streamflow prediction: climate signal weighting methods vs. climate forecast system reanalysis. *Journal of Hydrology*, **442**: 105–116.

Nalbantis, I, Efstratiadis, A., Rozos, E., Kopsiafti, M. and Koutsoyiannis, D. (2011). Holistic versus monomeric strategies for hydrological modelling of human-modified hydrosystems. *Hydrology and Earth System Sciences*, **15** (3): 743.

Nasambu, O.B., Masinde, N.E. and Mwikali, K.V. (2018). Examining the Geographical Coverage of Floods Using Satellite Images and Discharge in LNRB. *International Journal of Scientific and Research Publications*, **8** (10). Doi: 10.29322/(IJSRP).8.10.2018. p8225.

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

NBC News, July 2021. Almost 200 dead, many still missing after floods as Germany counts devastating costhttps://www.nbcnews.com/news/world/almost-200-dead-many-still-missing-after-floods-germany-counts-n1274330: Accessed 06/07/2022.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R. and Williams, J.R. (2005). *Soil and water assessment tool theoretical documentation.* Temple, Texas: USDA-ARS Grassland Soil and Water Research Laboratory, and Texas A&M University, Blackland Research and Extension Center.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R. and Williams, J.R. (2011). *Soil and water assessment tool theoretical documentation version 2009*. College Station, Texas: Texas Water Resources Institute.

Ngaina, J.N. (2014). Flood forecasting over lower Nzoia Sub-Basin in Kenya. *Africa Journal of Physical Sciences ISSN: 2313-3317*, **1** (1): 25-31.

Nguyen, T.G. and Tran, A.P. (2010). Calibration and verification of a hydrological model using event data. *VNU Journal of Science: Earth and Environmental Sciences*, **26** (2): 64-74.

Nicholson, S.E. (2017). Climate and climatic variability of rainfall over eastern Africa. *Reviews of Geophysics*, **55** (3): 590–635. doi: 10.1002/2016RG000544.

Nicholson, S.E. (2018). The ITCZ and the seasonal cycle over equatorial Africa. *Bulletin of the American Meteorological Society*, **99** (2): 337-348.

Nicholson, S.E., Klotter, D., Zhou, L. and Hua, W. (2019). Validation of satellite precipitation estimates over the Congo Basin. *Journal of Hydrometeorology*, **20** (4): 631–656.

Nielsen, S.A. and Hansen, E. (1973). Numerical simulation of the rainfall-runoff process on a daily basis. *Hydrology Research*, **4** (3): 171–190.

Nijssen, B., Lettenmaier, D.P., Liang, X., Wetzel, S.W. and Wood, E.F. (1997). Streamflow simulation for continental-scale river basins. *Water Resources Research*, **33** (4): 711–724.

Nkiaka, E., Nawaz, N.R. and Lovett, J.C. (2017). Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: case study of the Logone catchment, Lake Chad Basin. *Meteorological Applications*, **24** (1): 9–18.

Nka, B.N., Oudin, L., Karambiri, H., Paturel, J.-E. and Ribstein, P. (2015). Trends in floods in West Africa: Analysis based on 11 catchments in the region. *Hydrology and Earth System Sciences*, **19** (11): 4707–4719.

Nobel, A. (2011). Modeling the impact of climate and land use change on discharges in the Citarum river. Bachelor Thesis, Delft University of Technology. http://resolver.tudelft.nl/uuid:eaf4971-67dc-439e-92db-f9e8d4e7342a

Nossent, J., Elsen, P. and Bauwens, W. (2011). Sobol' sensitivity analysis of a complex environmental model. *Environmental Modelling & Software*, **26** (12): 1515– 1525.

Nowak, A. and Hodson, A. (2013). Hydrological response of a High-Arctic catchment to changing climate over the past 35 years: a case study of Bayelva watershed, Svalbard. *Polar Research*, **32** (1): 19691.

Nyakundi, H., Mwanzo, I. and Yitambe, A. (2010). Community perceptions and response to flood risks in Nyando District, Western Kenya. *Jàmbá: Journal of Disaster Risk Studies*, **3** (1): 346-366.

Nyenzi, B. S. (1988). Mechanisms of East African rainfall variability. Doctoral Dissertation. Tallahassee: The Florida State University.

Nyokabi, E.W., Wambua, R.M. and Okwany, R.O. (2021). Assessment of rainfall, streamflow and reservoir level trends for Malewa River Catchment, Naivasha, Kenya. *Journal of Civil, Construction and Environmental Engineering*, **6** (1): 1–8.

O'Connor, K.M. (2005). The Galway Real-Time River Flow Forecasting System (GFFS). In: D.W. Knight and A.Y. Shamseldin (eds) River *Basin Modelling for Flood Risk Mitigation*. 1st ed. Abingdon: CRC Press, pp. 215-234.

OCHA Flash Update #5: Floods in Kenya | 10 May 2018 - Kenya | ReliefWeb (no date). Available at: https://reliefweb.int/report/kenya/ocha-flash-update-5-floods-kenya-10-may-2018 (Accessed: 3 March 2020).

Odiyo, J.O., Phangisa, J.I. and Makungo, R. (2012). Latonyanda River flow contributions to Luvuvhu River downstream of Albasini Dam. *Physics and Chemistry of the Earth, Parts A/B/C*, **50**: 5–13.

Ogallo, L. A. (1984). Temporal fluctuations of seasonal rainfall patterns in East Africa. *Mausam*, **35** (2): 175-180.

Ogallo, L.A. (1993). Dynamics of the East African climate. *Proceedings of the Indian Academy of Sciences; Earth and Planetary Sciences*, **102** (1): 203–217.

Ogallo, L.A. (1988). Relationships between seasonal rainfall in East Africa and the Southern Oscillation. *Journal of Climatology*, **8** (1): 31–43.

Ogwang, B.A., Ongoma, V., Zing, L. and Ogou, K.F. (2015). Influence of Mascarene high and Indian Ocean dipole on East African extreme weather events. *Geographica Pannonica*, **19** (2): 64–72.

Ojara, M.A., Yunsheng, L., Ongoma, V., Mumo, L., Akodi, D., Ayugi, B. and Ogwang, B.A. (2021). Projected changes in East African climate and its impacts on climatic suitability of maize production areas by the mid-twenty-first century. *Environmental Monitoring and Assessment*, **193** (12): 1–24.

Ombajo, L.A., Mutono, N., Sudi, P., Mutua, M., Sood, M., Loo, A.M., Juma, P., Odhiambo, J., Shah, R., Wangai, F. and Maritim, M. (2022). Epidemiological and clinical characteristics of patients hospitalised with COVID-19 in Kenya: a multicentre cohort study. *BMJ open*, **12** (5): p.e049949.

Omondi, P.A., Awange, J.L., Forootan, E., Ogallo, L.A., Barakiza, R., Girmaw, G.B., Fesseha, I., Kululetera, V., Kilembe, C., Mbati, M.M., Kilavi, M., King'uyu, S.M., Omeny, P.A., Njogu, A., Badr, E.M., Musa, T.A., Muchiri, P., Bamanya, D. and Komutunga, E. (2014). Changes in temperature and precipitation extremes over the Greater Horn of Africa region from 1961 to 2010. *International Journal of Climatology*, **34** (4):1262–1277.

Ongoma, V., Guirong, T., Ogwang, B.A, and Ngarukiyimana, J.P. (2015). Diagnosis of seasonal rainfall variability over East Africa: a case study of 2010-2011 drought over Kenya. *Pakistan Journal of Meteorology*, **11** (22): 13-21.

Ongoma, V. and Chen, H. (2017). Temporal and spatial variability of temperature and precipitation over East Africa from 1951 to 2010. *Meteorology and Atmospheric Physics*, **129** (2): 131–144.

Ongoma, V., Chen, H. and Omony, G.W. (2018). Variability of extreme weather events over the equatorial East Africa, a case study of rainfall in Kenya and Uganda. *Theoretical and Applied Climatology*, **131** (1–2): 295–308.

Onyando, J O., Schumann, A.H. and Schultz, G.A. (2003). Simulation of flood hydrographs based on lumped and semi-distributed models for two tropical catchments in Kenya. *Hydrological Sciences Journal*, **48** (4): 511–524.

Onyutha, C. (2016). Variability of seasonal and annual rainfall in the River Nile riparian countries and possible linkages to ocean–atmosphere interactions'.*Hydrology Research*, **47** (1): 171–184.

Opido, G., Odwe, G., Oulu, M. and Omollo, E. (2017). *Migration as adaptation to environmental and climate change: the case of Kenya*. Geneva: International Organization for Migration.

Otiende, B. (2009). *The economic impacts of climate change in Kenya: riparian flood impacts and cost of adaptation*. http://static.weadapt.org/knowledge-base/files/758/4e25a4b8c8bf61C-kenya-riparian-floods-case-study.pdf.

Oudin, L., Andréassian, V. and Perrin, C. (2004). Locating the sources of low-pass behavior within rainfall-runoff models. *Water Resources Research*, **40** (11), W11101, doi:10,1029/2004WR003291.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F. and Loumagne, C., 2005. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of hydrology*, *303*(1-4), pp.290-306.

Pappenberger, F., Thielen, J. and Del Medico, M. (2011). The impact of weather forecast improvements on large scale hydrology: analysing a decade of forecasts of the European Flood Alert System. *Hydrological Processes*, **25** (7):1091–1113. doi: 10.1002/hyp.7772.
Paprotny, D., Sebastian, A., Morales-Nápoles, O. and Jonkman, S. (2018). Trends in flood losses in Europe over the past 150 years. *Nature communications*, **9:** 1985. doi: 10.1038/s41467-018-04253-1.

Parry, J.E,. Echeverria, D., Dekens, J. and Maitima, J. (2012). Climate Risks, Vulnerability and Governance in Kenya: A review. United Nations Development Programme (UNDP); Winnipeg: International Institute for Sustainable Development (IISD). http://www.iisd.org/sites/default/files/pdf/2013/climate_risks_kenya.pdf.

Paschalis, A., Fatichi, S., Molnar, P., Rimkus, S. and Burlando, P. (2014). On the effects of small scale space–time variability of rainfall on basin flood response.*Journal of Hydrology*, **514**: 313–327.

Patil, S.D. and Stieglitz, M. (2015). Comparing spatial and temporal transferability of hydrological model parameters. *Journal of Hydrology*, **525**: 409–417.

Parajka, J., Kohnová, S., Bálint, G., Barbuc, M., Borga, M., Claps, P., Cheval, S., Dumitrescu, A., Gaume, E. and Hlavčová, K. (2010). Seasonal characteristics of flood regimes across the Alpine–Carpathian range. *Journal of Hydrology*, **394** (1–2): 78–89.

Parry, M.L., Canziani, O., Palutikof, J., Van der Linden, P. and Hanson, C. (2007). *Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC*. Cambridge: Cambridge University Press.

Pechlivanidis, I.G. and Arheimer, B. (2015). Large-scale hydrological modelling by using modified PUB recommendations: the India-HYPE case. *Hydrology and Earth System Sciences*, **19** (11): 4559–4579.

Pechlivanidis, I.G., Jackson, B.M., McIntyre, N.R. and Wheather, H.S. (2011). Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST Journal*, **13** (3): 193–214.

Peng, Y. Chu, J., Sun, X., Zhou, H. and Zhang, X. (2016). Flood forecasting that considers the impact of hydraulic projects by an improved TOPMODEL model in the Wudaogou basin, Northeast China. *Water Supply*, **16** (5): 1467–1476.

Perrin, C., Michel, C. and Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, **279** (1–4): 275–289.

Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B. and Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, **79**: 214–232.

Praskievicz, S. and Chang, H. (2009). A review of hydrological modelling of basin-scale climate change and urban development impacts. *Progress in Physical Geography*, **33** (5): 650–671.

Quadro, M.F.L., Berbery, E.H., Silva Dias, M.A.F., Herdies, D.L. and Gonçalves, L.G.G. (2013). The atmospheric water cycle over South America as seen in the new generation of global reanalyses. *AIP Conference Proceedings,* **1531***: 732–735.

Quang, N.H. (2016). Modelling Soil Erosion, Flash Flood Prediction and Evapotranspiration in Northern Vietnam. PhD Thesis, Georg-August-Universität, Göttingen.

Quadro, M.F.L., Berbery, E.H., Silva Dias, M.A.F., Herdies, D.L. and Gonçalves, L.G.G. (2013). The atmospheric water cycle over South America as seen in the new generation of global reanalyses. *American Institute of Physics Conference Proceedings*, **1531** (1); 732–735.

Quinn, P., Beven, K., Chevallier, P. and Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, **5** (1): 59–79.

Qi, W., Zhang, C., Fu, G. and Zhou, H. (2016). Quantifying dynamic sensitivity of optimization algorithm parameters to improve hydrological model calibration. *Journal of Hydrology*, **533**: 213–223.

Rabier, F., Venuti, F., Brown, A., Pappenberger, F., Thepaut, J.-N., Palkovic, M., Neves, L., Bauer, P., Peuch, V.-H., Buontempo, C., Weger, I., Modigliani, U. and DellAcqua, M. (2021). The ECMWF Strategy 2021–30. In *101st American Meteorological Society Annual Meeting*. https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/379171.

Rahman, M.M., Arya, D.S., Goel, N.K. and Dhamy, A.P. (2011). Design flow and stage computations in the Teesta River, Bangladesh, using frequency analysis and MIKE 11 modeling. *Journal of Hydrologic Engineering*, **16** (2): 176–186.

Rafiei Emam, A., Kappas, M., Fassnacht, S., Linh, N.H.K., Emam, A.R., Kappas, M., Fassnacht, S. and Linh, N.H.K. (2018). Uncertainty analysis of hydrological modeling in a tropical area using different algorithms. *Frontiers of Earth Science*, **12** (4): 661–671.

Rapaić, M., Brown, R., Markovic, M. and Chaumont, D. (2015). An evaluation of temperature and precipitation surface-based and reanalysis datasets for the Canadian Arctic, 1950–2010. *Atmosphere-Ocean*, **53** (3): 283–303.

Razad, A.Z.A., Sidek, L.M., Jung, K. and Basri, H. (2018). Reservoir inflow simulation using MIKE NAM rainfall-runoff model: case study of Cameron Highlands. *Journal of Engineering Science and Technology*, **13** (12): 4206–4225.

Refsgaard, J.C. and Storm, B. (1995). MIKE SHE. In: V.P. Singh (ed.) *Computer models of watershed hydrology,* Fort Collins, Colorado: Water Resources Publications, pp. 809–846.

Refsgaard, J.C., van der Sluijs, J.P., Brown, J. and Van der Keur, P.A. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, **29** (11): 1586–1597.

Refsgaard, J.C. and Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, 32 (7): 2189–2202.

Renard, B., Kavetski, D., Thyer, M., Kuczera, G. and Franks, S. (2010). Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors. *Water Resources Research*, **46** (5): W05521. https://doi.org/10.1029/2009WR008328.

Rientjes, T.H.M., Haile, A.T., Kebede, E., Mannaerts, C.M.M., Habib, E. and Steenhuis, T.S. (2011). Changes in land cover, rainfall and stream flow in Upper Gilgel Abbay catchment, Blue Nile basin--Ethiopia. *Hydrology & Earth System Sciences*, **15** (6): 1979-1989.

Rojas, R., Feyen, L., Bianchi, A. and Dosio, A. (2012). Assessment of future flood hazard in Europe using a large ensemble of bias-corrected regional climate simulations. *Journal of Geophysical Research: Atmospheres*, **117** (D17), p. n-a-n/a. doi: 10.1029/2012JD017461.

Romanowicz, R. and Beven, K. (1998). Dynamic real-time prediction of flood inundation probabilities. *Hydrological Sciences Journal*, **43** (2): 181–196.

Rosenbaum, P.R. (2015). Two R packages for sensitivity analysis in observational studies. *Observational Studies*, **1** (2): 1–17.

De Roo, A.P., Thielen, J., Salomon, P., Bogner, K., Nobert, S., Cloke, H., Demeritt, D., Younis, J., Kalas, M., Bódis, K., Muraro, D. and Pappenberger, F. (2011). Quality control, validation and user feedback of the European Flood Alert System (EFAS). *International Journal of Digital Earth*, **4** (SUPPL. 1): 77–90. doi: 10.1080/17538947.2010.510302.

Rosbjerg, D., Blöschl, G., Burn, D.H., Castellarin, A., Croke, B., Baldassarre, G. Di, Iacobellis, V., Kjelsden, T.R., Kuczera, G., Merz, R., Montanari, A., Morris, D., Ouarda, T.B.M.J., Ren, L., Rogger, M., Salinas, J.L., Toth, E. and Viglione, A. (2013). Prediction of floods in ungauged basins. In: G. Blöschl, M. Sivapalan, T. Wagener, A. Viglione and H. Savenije (eds) *Runoff prediction in ungauged basins: Synthesis across processes, places and scales,* Cambridge: Cambridge University Press, pp. 189–225.

Rostami, K.M., Moghadamnia, A., Salmani, H. and Sepahvand, A. (2016). Compare the performance of AWBM, Sacramento, SimHyd, SMAR and Tank. *Natural Ecosystems of Iran*, 7 (1): 47-63.

Rosner, A., Vogel, R.M. and Kirshen, P.H. (2014). A risk-based approach to flood management decisions in a nonstationary world. *Water Resources Research*, **50** (3): 1928–1942.

Ruelland, D., Larrat, V. and Guinot, V. (2010). A comparison of two conceptual models for the simulation of hydro-climatic variability over 50 years in a large Sudano-Sahelian catchment. In: E. Servat, S. Demuth, A. Dezetter, T. Daniell, E. Ferrari, M. Ijjaali, R. Jabrane, H. van Lanen and Y. Huang (eds) *Global Change: Facing Risks and Threats to Water Resources (proceedings of the 6th World FRIEND Conference, Fez, Morocco, Oct. 2010)*, Wallingford: IAHS Publications, pp. 668–678.

Sa'adi, Z., Shahid, S., Ismail, T., Chung, E.S. and Wang, X.J. (2019). Trends analysis of rainfall and rainfall extremes in Sarawak, Malaysia using modified Mann–Kendall test. *Meteorology and Atmospheric Physics*, **131** (3): 263–277. Available at: https://doi.org/10.1007/s00703-017-0564-3.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, X., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. J., Sela, J., Iredel, N., Treadon, R., Kleist, D., van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool. H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G. and Goldberg, M. (2010). The NCEP Climate Forecast System Reanalysis, *Bulletin of the American Meteorological Society*, **91** (8): 1015–1058. doi: 10.1175/2010bams3001.1.

Saji, N.H., Goswami, B.N., Vinayachandran, P.N. and Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, **401** (6751): 360-363.

Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer physics communications*, **145** (2): 280–297.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer physics communications*, **181** (2): 259–270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008). *Global sensitivity analysis: the primer*. 1st ed. Chichester: John Wiley & Sons.

Saltelli, A., Tarantola, S. and Campolongo, F. (2000). Sensitivity analysis as an ingredient of modelling. *Statistical Science*, **15** (4): 377–395.

Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*. 1st ed. Hoboken: John Wiley & Sons. Wiley Online Library.

Salvadore, E., Bronders, J. and Batelaan, O. (2015). Hydrological modelling of urbanized catchments: a review and future directions. *Journal of Hydrology*, **529**: 62–81.

Samaniego, L., Kumar, R. and Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, **46:** WO5523, doi: 10.1029/2008WR007327.

Samani, Z.A. and Pessarakli, M. (1986). Estimating potential crop evapotranspiration with minimum data in Arizona. *Transactions of the ASAE*, **29** (2): 522–524.

Sang, J., Gatheny, M. and Ndegwa, G. (2005). Evaluation of reservoirs as flood mitigation measure in Nyando Basin, Western Kenya Using SWAT. *European Journal of Scientific Research, Editorial Advisory Board e*, **18** (2): 231–239.

Santos, C.A.S., Almeida, C., Ramos, T.B., Rocha, F.A., De Oliveira, R.P. and Neves, R.J.J. (2018). Using a hierarchical approach to calibrate SWAT and predict the semi-arid hydrologic regime of northeastern Brazil', *Water*, **10** (9): 1137.

Savic, D. (2002). Single-objective vs. multiobjective optimisation for integrated decision support. In: *Integrated Assessment and Decision. Proceedings of the First Biennial Meeting of the International Environmental Modelling and Software Society*, pp. 7-12.

Savvidou, E., Efstratiadis, A., Koussis, A.D., Koukouvinos, A. and Skarlatos, D. (2016). A curve number approach to formulate hydrological response units within distributed hydrological modelling. *Hydrology and Earth System Sciences Discussions*, **10** (194):1–32.

Schulzweida, U. (2019). *CDO User Guide*. Climate Data Operators. Verson 1.9.5. doi: 10.5281/zenodo.2558193.

Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, **63** (324): 1379–1389.

Sepulcre-Canto, G., Horion, S., Singleton, A. and Vogt, J. (2012). Development of a Combined Drought Indicator to detect agricultural drought in Europe. *Natural Hazards and Earth System Sciences*, **12** (11): 3519–3531.

Serrat-Capdevila, A., Valdes, J. B. and Stakhiv, E. Z. (2014). Water management applications for satellite precipitation products: synthesis and recommendations. *Journal of the American Water Resources Association*, **50** (2): 509–525.

Setti, S., Maheswaran, R., Sridhar, V., Barik, K.K., Merz, B. and Agarwal, A. (2020). 'Inter-comparison of gauge-based gridded data, reanalysis and satellite precipitation product with an emphasis on hydrological modelling. *Atmosphere*, **11** (11): 1252.

Shamseldin, A.Y., Abdo, G.M. and Elzein, A.S. (1999). Real-Time flood forecasting on the Blue Nile River. *Water International*, **24** (1): 39-45. 8060(1999). doi: 10.1080/02508069908692132.

Shahidian, S., Serralheiro, R., Serrano, J., Teixeira, J., Haie, N. and Santos, F. (2012). *Hargreaves and other reduced-set methods for calculating evapotranspiration*. IntechOpen London. Doi: 10.5772/18059.

Shayeghi, A., Azizian, A. and Brocca, L. (2020). Reliability of reanalysis and remotely sensed precipitation products for hydrological simulation over the Sefidrood River Basin, Iran. *Hydrological Sciences Journal*, **65** (2): 296–310.

Sherlock, E. and Duffy, S. (2019). Establishing the flood forecast centre and expanding Met Éireann's rainfall radar network. Irish Hydrology Conference, 2019, 05.

Shin, M.J., Guillaume, J.H.A. and Croke, B.F.W. (2015). A review of foundational methods for checking the structural identifiability of models: results for rainfall-runoff. *Journal of Hydrology*, **520**: 1–16.

Shin, M.-J. and Kim, C.-S. (2017). Assessment of the suitability of rainfall–runoff models by coupling performance statistics and sensitivity analysis. *Hydrology Research*, **48** (5): 1192–1213.

Singh, A., Singh, S., Nema, A.K., Singh, G. and Gangwar, A. (2014). Rainfall-runoff modeling using MIKE 11 NAM model for Vinayakpur Intercepted Catchment, Chhattisgarh. *Indian Journal of Dryland Agricultural Research and Development*, **29** (2): 01–04.

Singh, R., Subramanian, K. and Refsgaard, J.C. (1999). 'Hydrological modelling of a small watershed using MIKE SHE for irrigation planning', *Agricultural Water Management*, **41** (3): 149–166.

Sivapalan, M., Blöschl, Zhang, L. and Vertessy, R. (2003). Downward approach to hydrological prediction. *Hydrological Processes*, **17** (11): 2101-2111.

Slater, L.J. and Wilby, R.L. (2017). Measuring the changing pulse of rivers. *Science*, **357** (6351): 552.

Smith, A., Sampson, C. and Bates, P. (2015). Regional flood frequency analysis at the global scale. *Water Resources Research*, **51** (1): 539–553.

Smith, R. A. and Kummerow, C.D. (2013). A comparison of in situ, reanalysis, and satellite water budgets over the upper Colorado River basin. *Journal of Hydrometeorology*, **14** (3): 888–905.

Sobol', I.M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, **2** (1): 112–118.

Sobol', I.M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, **55** (1–3): 271–280.

Song, M., Shi, Y., Yao, H. and Zhang, W. (2019). A comparative study of different hydrological model and their application in Bass River Catchment. in*IOP Conference Series: Materials Science and Engineering*, **562** (1): 12116.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M. and Xu, C. (2015). Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. *Journal of Hydrology*, **523** (225): 739–757. Available at: https://doi.org/10.1016/j.jhydrol.2015.02.013.

Sood, A. and Smakhtin, V. (2015). Global hydrological models: a review. *Hydrological Sciences Journal*, **60** (4): 549–565.

Sorooshian, S. and Gupta, V. K. (1995). Model calibration. *Computer Models of Watershed Hydrology*, **1**: 23–68.

Srivastava, P.K., Han, D., Rico-Ramirez, M. and Islam, T. (2014). Sensitivity and uncertainty analysis of mesoscale model downscaled hydro-meteorological variables for discharge prediction. *Hydrological Processes*, **28** (15): 4419–4432.

Stephens, E. and Cloke, H. (2014). Improving flood forecasts for better flood preparedness in the UK (and beyond). *The Geographical Journal*, **180** (4): 310-316.

Stephens, E., Day, J.J., Pappenberger, F. and Cloke, H. (2015). Precipitation and floodiness. *Geophysical Research Letters*, **42** (23): 10–316.

Sun, L., Goetz, E., Kissel, J.S., Betzwieser, J., Karki, S., Viets, A., Wade, M., Bhattacharjee, D., Bossilkov, V., Covas, P.B. and Datrier, L.E. (2020). Characterization of systematic error in Advanced LIGO calibration. *Classical and Quantum Gravity*, **37** (22): 225008.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S. and Hsu, K. (2018). A review of global precipitation data sets: data sources, estimation, and intercomparisons, *Reviews of Geophysics*, **56** (1): 79–107.

Sun, S., Bleck, R., Benjamin, S.G., Green, B.W. and Grell, G.A. (2018). Subseasonal forecasting with an icosahedral, vertically quasi-Lagrangian coupled model. Part I: Model overview and evaluation of systematic errors. *Monthly Weather Review*, **146** (5): 1601-1617.

Sutanudjaja, E.H., Van Beek, R., Wanders, N., Wada, Y., Bosmans, J.H.C., Drost, N., Van Der Ent, R.J., De Graaf, I.E.M., Hoch, J.M., De Jong, K., Karessenberg, D., López, P. L., Pessenteiner, S., Schmitz, O., Straatsma, M.W., Vannamatee, E., Wisser, D. and Bierkens, M.F.P. (2018). 'PCR-GLOBWB 2: A 5 arcmin global hydrological and water resources model. *Geoscientific Model Development,* **11**: 2429–2453. doi: 10.5194/gmd-11-2429-2018.

Svensson, C., Kundzewicz, Z.W. and Maurer, T. (2005). Trend detection in river flow series: 2. Flood and low-flow index series. *Hydrological Sciences Journal*, **50** (5): 811–824. Available at: https://doi.org/10.1623/hysj.2005.50.5.811.

Tamm, O., Luhamaa, A. and Tamm, T. (2016). Modeling future changes in the North-Estonian hydropower production by using SWAT. *Hydrology Research*, **47** (4): 835–846.

Tan, B.Q. and O'Connor, K.M. (1996). Application of an empirical infiltration equation in the SMAR conceptual model. *Journal of Hydrology*, **185** (1–4): 275–295.

Tang, G., Clark, M.P., Papalexiou, S.M., Ma. Z. and Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote sensing of environment*, **240**: 111697.

Tapiador, F.J., Turk, F.J., Petersen, W., Hou, A.Y., Garcia-Ortega, E., Machado, L.A.T., Angelis, C.F., Salio, P., Kidd, C., Huffman, G.J. and De Castro, M. (2012). Global precipitation measurement: methods, datasets and applications. *Atmospheric Research*, **104**: 70–97.

Tarek, M., Brissette, F. and Arsenault, R. (2021). Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies. *Hydrology and Earth System Sciences*, **25** (6): 3331–3350.

Tarek, M., Brissette, F.P. and Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences*, **24** (5): 2527–2544.

Tatsumi, K. and Yamashiki, Y. (2015). Effect of irrigation water withdrawals on water and energy balance in the Mekong River Basin using an improved VIC land surface model with fewer calibration parameters. *Agricultural Water Management*, **159**: 92–106.

Tegegne, G., Park, D.K. and Kim, Y.O. (2017). Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin. *Journal of Hydrology: Regional Studies*, **14:** 49–66. doi: 10.1016/j.ejrh.2017.10.002.

Thiemig, V., Bisselink, B., Pappenberger, F. and Thielen, J. (2015). A pan-African medium-range ensemble flood forecast system'.*Hydrology and Earth System Sciences*, **19** (8): 3365–3385. doi: 10.5194/hess-19-3365-2015.

Thiemig, V., De Roo, A. and Gadain, H. (2011). Current status on flood forecasting and early warning in Africa. *International Journal of River Basin Management*, **9**: 63–78.

Thompson, J.R., Sørenson, H.R., Gavin, H. and Refsgaard, A. (2004). Application of the coupled MIKE SHE/MIKE 11 modelling system to a lowland wet grassland in southeast England. *Journal of Hydrology*, 293 (1–4):151–179.

Thompson, S.A. (1999). *Hydrology for water management*. 1st ed. Abingdon: CRC Press.

Tian, Y., Xu, Y.-P. and Zhang, X.-J. (2013). Assessment of climate change impacts on river high flows through comparative use of GR4J, HBV and Xinanjiang models. *Water Resources Management*, **27** (8): 2871–2888.

Todini, E. (2007). Hydrological catchment modelling: past, present and future. *Hydrology and Earth System Sciences*, **11** (1): 468–482.

Towner, J., Cloke, H.L., Zsoter, E., Flamig, Z., Hoch, J.M., Bazo, J., de Perez, E.C. and Stephens, E.M. (2019). Assessing the performance of global hydrological models for capturing peak river flows in the Amazon Basin. *Hydrology and Earth System Sciences, 23* (7): 3057-3080. doi: 10.5194/hess-2019-44.

Trambauer, P., Maskey, S., Winsemius, H., Werner, M. and Uhlenbrook, S. (2013). A review of continental scale hydrological models and their suitability for drought forecasting in (sub-Saharan) Africa. *Physics and Chemistry of the Earth, Parts A/B/C*, **66**: 16–26.

Tramblay, Y., Mimeau, L., Tramblay, Y., Neppel, L., Vinet, F. and Sauquet, E. (2019). Detection and attribution of flood trends in Mediterranean basins. *Hydrology and Earth System Sciences*, **23** (11): 4419–4431.

Traore, V.B., Sambou, S., Tamba, S., Fall, S., Diaw, A.T. and Cissé, M.T. (2014). Calibrating the rainfall-runoff model GR4J and GR2M on the Koulountou river basin, a tributary of the Gambia river. *American Journal of Environmental Protection*, **3** (1): 36–44.

UNICEF sends supplies to flood-affected children and families in 12 counties in Kenya - Kenya | ReliefWeb (no date). Available at: https://reliefweb.int/report/kenya/unicef-sends-supplies-flood-affected-children-and-families-12-counties-kenya (Accessed: 3 March 2020).

Vansteenkiste, T., Tavakoli, M., Ntegeka, V., De Smedt, F., Batelaan, O., Pereira, F. and Willems, P. (2014). Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections. *Journal of Hydrology*, **519**: 743–755.

Verzano, K. (2009). Climate change impacts on flood related hydrological processes: further development and application of a global scale hydrological model. PhD thesis, International Max Planck Research School on Earth System Modelling, University of Kassel.

Viglione, A., Chirico, G.B., Komma, J., Woods, R., Borga, M. and Blöschl, G. (2010). Quantifying space-time dynamics of flood event types. *Journal of Hydrology*, **394** (1–2): 213–229.

Viney, N.R., Bormann, H., Breuer, L., Bronstert, A., Croke, B.F.W., Frede, H., Gräff, T., Hubrechts, L., Huisman, J.A., Jakeman, A.J., Kite, G.W., Lanini, J., Leavesley, G., Lettenmaier, D.P., Lindström, G., Seibert, J., Sivapalan, M. and Willems, P. (2009). Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) II: ensemble combinations and predictions. *Advances in Water Resources*, **32** (2): 147–158.

Vischel, T., Quantin, G., Lebel, T., Viarre, J., Gosset, M., Cazenave, F. and Panthou, G. (2011). Generation of high-resolution rain fields in West Africa: Evaluation of dynamic interpolation methods. *Journal of Hydrometeorology*, **12** (6): 1465–1482.

Vogel, R.M., Yaindl, C. and Walter, M. (2011). Nonstationarity: Flood magnification and recurrence reduction factors in the United States. *Journal of the American Water Resources Association*, **47** (3): 464–474. Available at: https://doi.org/10.1111/j.1752-1688.2011.00541.x.

Vormoor, K., Lawrence, D., Schlichting, L., Wilson, D. and Wong, W.K. (2016). Evidence for changes in the magnitude and frequency of observed rainfall vs. snowmelt driven floods in Norway. *Journal of Hydrology*, **538**: 33–48.

Von Storch, H. and Navarra, A. (1999). *Analysis of climate variability: applications of statistical techniques*. 2nd ed. Berlin, Heidelberg: Springer Science & Business Media.

Wagener, T., Dadson, S.J., Hannah, D.M., Coxon, G., Beven. K., Bloomfield, J.P., Buytaert, W., Cloke, H., Bates, P., Holden, J., Parry, L., Lamb, R., Chappell, N.A., Fry, M. and Old, G. (2021). Knowledge Gaps in our Perceptual Model of Great Britain's Hydrology. *Hydrological Processes*, **35** (7): [e14288].

Wahren, F.T., Julich, S., Nunes, P.J., Gonzales-Pelayo, O., Hawtree, D., Feger, K.-H. and Keizer, J. J. (2016). Combining digital soil mapping and hydrological modeling in a data scarce watershed in north-central Portugal. *Geoderma*, **264, Part B**: 350–362.

Wainwright, C.M., Finney, D.L., Kilavi, M., Black, E. and Marsham, J.H. (2021). Extreme rainfall in East Africa, October 2019–January 2020 and context under future climate change. *Weather*, **76** (1): 26–31.

Wang, G., Zhang, X. and Zhang, S. (2019). Performance of three reanalysis precipitation datasets over the Qinling-Daba Mountains, eastern fringe of Tibetan Plateau, China. *Advances in Meteorology* 2019: 7698171. https://doi.org/10.1155/2019/7698171.

Wanzala, M. and Cloke H. (2021). What can Africa learn from other countries about combatting severe floods? https://www.anticipation-hub.org/news/what-can-africa-learn-from-other-countries-about-combatting-severe-floods, Accessed 06/07/2022.

Wanzala, M.A., Ficchì, A., Cloke, H.L., Stephens, E.M., Badjana, H.M. and Lavers, D.A. (2022). Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: A case study of Kenya. *Journal of Hydrology: Regional Studies*, **41:** 101105.

Wara, C., Taye, M., Quevauviller, P. and Willems, P. (2015). Climate change impacts on the water cycle and river flow regime of Nyando River catchment in Kenya. Thesis, Katholieke Universiteit Leuven.   doi: 10. 13140/RG.2.2.20641,51043.

Ward, P.J., Jongman, B., Salamon, P., Simpson, A., Bates, P., De Groeve, T., Muis, S., de Perez, E.C., Rudari, R., Trigg, M.A. and Winsemius, H.C. (2015). Usefulness and limitations of global flood risk models. *Nature Climate Change*, **5** (8):712–715.

Warner, K. and Van der Geest, K. (2013). Loss and damage from climate change: local-level evidence from nine vulnerable countries. *International Journal of Global Warming*, **5** (4): 367–386.

Weisheimer, A. and Palmer, T.N. (2014). On the reliability of seasonal climate forecasts. *Journal of the Royal Society Interface*, **11** (96): 20131162.

Werner, M., Reggiani, P., & Weerts, A. H. (2014). Quantifying and Reducing Uncertainties in Operational Forecasting: Examples from the Delft FEWS Forecasting System. In: K. Beven and J. Hall (eds) *Applied Uncertainty Analysis for Flood Risk Management,* London: Imperial College Press, pp. 506-537.

Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., & Heynert, K. (2013). The Delft-FEWS flow forecasting system. *Environmental Modelling & Software*, **40:** 65-77.

Westerberg, I.K., Di Baldassarre, G., Beven, K.J., Coxon, G. and Krueger, T. (2017). Perceptual models of uncertainty for socio-hydrological systems: a flood risk change example. *Hydrological Sciences Journal*, 62 (11): 1705–1713.

Wheater, H., Sorooshian, S. and Sharma, K D. (eds) (2007). *Hydrological modelling in arid and semi-arid areas.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511535734.

Wilcox, C., Vischel, T., Panthou, G., Bodian, A., Blanchet, J., Descroix, L., Quantin, G., Cassé, C., Tanimoun, B. and Kone, S. (2018) 'Trends in hydrological extremes in the Senegal and Niger Rivers', *Journal of Hydrology*, **566**: 531–545.

Willems, P., Mora, D., Vansteenkiste, T., Taye, M.T. and Van Steenbergen, N. (2014). Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes–Part 2: intercomparison of models and calibration approaches. *Journal of Hydrology*, **510**: 591–609.

World Health Organization, 2022. COVID-19 weekly epidemiological update, edition 97, 22 June 2022.

World Meteorological Organization (2022). *State of the Global Climate 2021*.WMO, No. 1290. Geneva: WMO.

Worqlul, A.W., Yen, H., Collick, A.S., Tilahun, S.A., Langan, S. and Steenhuis, T.S. (2017). Evaluation of CFSR, TMPA 3B42 and ground-based rainfall data as input for hydrological models, in data-scarce regions: the upper Blue Nile Basin, Ethiopia. *Catena*. **152**: 242–251.

Wu, W., Chen, J. and Huang, R. (2013). Water budgets of tropical cyclones: three case studies. *Advances in Atmospheric Sciences*, **30** (2): 468–484.

Xing, W., Wang, W., Shao, Q. and Yong, B. (2018). Identification of dominant interactions between climatic seasonality, catchment characteristics and agricultural activities on Budyko-type equation parameter estimation. *Journal of Hydrology*, **556**: 585–599.

Xu, H., Xu, C.-Y., Chen, S. and Chen, H. (2016). Similarity and difference of global reanalysis datasets (WFD and APHRODITE) in driving lumped and distributed hydrological models in a humid region of China. *Journal of Hydrology*, **542:** 343–356. doi: 10.1016/j.jhydrol.2016.09.011.

Xu, L., Chen, N., Moradkhani, H., Zhang, X. and Hu, C, (2020). Improving global monthly and daily precipitation estimation by fusing gauge observations, remote sensing, and reanalysis data sets. *Water Resources Research*, **56** (3): e2019WR026444.

Xu, Y.D., Fu, B.J., He, C.S. and Gao, G.Y. (2012). Watershed discretization based on multiple factors and its application in the Chinese Loess Plateau. *Hydrology and Earth System Sciences*, 16 (1): 59-68.

Yamazaki, D., Kanae, S., Kim, H. and Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, **47** (4). doi: 10.1029/2010WR009726.

Yang, W., Seager, R., Cane, M.A. and Lyon, B. (2015). The annual cycle of East African precipitation. *Journal of Climate*, **28** (6): 2385–2404.

Yapo, P.O., Gupta, H.V. and Sorooshian, S. (1998). Multi-objective global optimization for hydrologic models. *Journal of hydrology*, **204** (1–4): 83–97.

Yue, S. and Wang, C. Y. (2004). The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resources Management*, **18** (3): 201–218. doi: 10.1023/B: WARM.0000043140.61082.60.

Zanchetta, A.D. and Coulibaly, P. (2020). Recent advances in real-time pluvial flash flood forecasting. *Water,* **12** (2): 570.

Zadeh, S.M., Burn, D.H. and O'Brien, N. (2020). Detection of trends in flood magnitude and frequency in Canada. *Journal of Hydrology: Regional Studies*, **28**: 100673.

Zeng, L. Xiong, L., Liu, D., Chen, J. and Kim, J-S. (2019). Improving parameter transferability of GR4J model under changing environments considering nonstationarity.   *Water*, **11** (10). doi: 10.3390/w11102029.

Zhang, X., Hörmann, G., Gao, J. and Fohrer, N. (2011). Structural uncertainty assessment in a discharge simulation model. *Hydrological Sciences Journal*, **56** (5): 854-869.

Zhang, J.Y., O'Connor, K.M. and Liang, G.C. (1994). A software package for river flow forecasting based on the SMAR model. *WIT Transactions on Ecology and the Environment,* 7**:** 163-170. doi: 10.2495/HY940191.

Zhang, X., Srinivasan, R., Zhao, K. and Liew, M. Van (2009). Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. *Hydrological Processes*, **23** (3): 430–441.

Zhao, S., He, W. and Jiang, Y. (2018). Evaluation of NCEP-2 and CFSR reanalysis seasonal temperature data in China using detrended fluctuation analysis. *International Journal of Climatology*, **38** (1): 252–263.

Zhang, X., Hörmann, G., Gao, J. and Fohrer, N. (2011). Structural uncertainty assessment in a discharge simulation model. *Hydrological Sciences Journal*, **56** (5): 854-869.

Appendix

This appendix contains a storyline of how my research evolved, the typeset versions of published chapter 4 and 5 in the international peer reviewed journals and submitted version of chapter 6 presented in this thesis (A1, A2, A3 and A4 respectively).

A5, A6a and A6b present some of the sample results from GR4J model simulation hydrographs for model calibration and validation, Sobol' sensitivity analysis including sensitivity indices and model parameter interactions. A7 presents the paper plan in preparation for the future work covered in Chapter 6, which due to time constraints was not actualised. A8 presents some of the select collaborating author peer and non-peer reviewed publications, media engagements and science communications.

Author contributions statements are provided for A1, A2, and A3 in Chapters 4, 5, and 6 respectively, and are also provided at the end of the typeset versions. All author contributions have been approved by Professor Hannah Cloke, supervisor.

Hannah L. Cloke

A1: Storyline on how my research emerged

This Appendix presents the storyline of my research journey highlighting how my research emerged into a concrete research schedule to address the objectives presented in this thesis.

A2: Hydrological model selection framework for flood applications in Kenya

This paper presents the published version of chapter 4 of this thesis, with the following reference:

Wanzala, M.A., Stephens, E.M., Cloke, H.L. and Ficchì, A., 2022. Hydrological model pre-selection with a filter sequence for the national flood forecasting system in Kenya. *Journal of Flood Risk Management*, https://doi.org/10.1111/jfr3.12846.

A3: **Assessment of reanalysis datasets for flood modelling and applications in Kenya**

This paper presents the published version chapter 5 of this thesis, with the following reference:

Wanzala, M.A., Ficchì, A., Cloke, H.L., Stephens, E.M., Badjana, H.M. and Lavers, D.A., 2022. Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: a case study of Kenya. *Journal of Hydrology: Regional Studies*, *41*, p.101105 https://doi.org/10.1016/j.ejrh.2022.101105.

A4: Detecting trends in flood series and shifts in flood timing across Kenya

This paper presents the submitted version of Chapter 6 of this thesis, currently under first review process, with the following reference:

Wanzala, M.A., Cloke, H.L., Stephens, E.M., Ficchì, A. and Harrigan S. Detecting trends in flood series and shifts in flood timing across Kenya. *Hydrological Sciences.* HSJ-2022-0371*,* Under Review.

A5: GR4J model simulations and observations across some of the Kenyan catchment

These figures provide some of the sample results from the modelling experiment using the GR4J hydrological model in calibration and simulation periods. The sample results form part of the results from the select catchments (Nzioa- Ruambwa, Sio, Yala, Sondu-Miriu and Mara) out of the 19 catchments modelled in the analysis chapter 5.

A6a: Sensitivity analysis sample results: - Sensitivity indices of GR4J model parameters

These figures provide some of the sample results from the Sobol' global sensitivity analysis experiment using the GR4J hydrological model. The figures show First order and Total effect indices of the Sobol' analysis outlined in Chapter 3, Section 3.2.4.3.1, from the 19 catchments modelled in the analysis Chapter 5.

A6b: Sensitivity analysis sample results: - GR4J model parameter interactions

These figures provide some of the sample results from the Sobol' global sensitivity analysis experiment using the GR4J hydrological model. The figures show G44J model parameter interactions from the Sobol' sensitivity analysis outlined in Chapter 3, Section 3.2.4.3.1. The sample results form part of the results from the select catchments (Nzioa-Ruambwa, and Tana-Garissa) out of the 19 catchments modelled in the analysis Chapter 5.

A7: Unveiling flood generating mechanisms in Kenya

This paper presents a version of the paper under preparation which forms the extended part informed by the results of Chapter 6 of this thesis with the following potential reference:

Wanzala, M.A., Cloke, H.L., Stephens, E.M., Ficch-ì, A. and Harrigan S. Unveiling flood generating mechanisms across Kenya. *Journal of Hydrology and Earth Sciences.* Under Preparation.

A8: Collaborative publications, articles, and media interviews

A7 presents some of the collaborative publication papers and articles I have co-authored as well as media interviews and science art communications I was involved in during my PhD.

How my research plan emerged

This appendix material reflectively considers my PhD journey and how the research schedule envisioned at the start of my PhD was transformed into the chapters of this thesis.

The initial plans for my research were geared towards analysis of the skill of extended range flood forecasting for Kenya (e.g., Alfieri *et al.*, 2013; Towner *et al* 2021; Webster *et al.,* 2010). I first considered the potential gaps in flood modelling and forecasting, and in order to understand the in-country perspectives, conducted a preliminary visit to the Kenya Meteorological Department's hydrological forecasting section and met the head of the section.

Figure 0.1:- Preliminary visits in 2019, to Kenya's meteorological department hydrological forecasting sector (top right), the Water Resources Authority (WRA) data management team (lower left) and the WRA Water Security and Resilience project PIs (lower right).

After lengthy discussions it emerged that little research had been undertaken in exploring and characterising hydrometeorological modelling tools in Kenya, particularly hydrological models and data to support and inform operational activities and future modelling developments. When I spoke to Njongu in an interview, he noted that the choice and use of the current model (SMR), was entirely subjective and mainly driven by the project funding following the push to implement a FF system in Nzioa after subsequent destructive flooding events. Additionally, he noted that there was limited documented research on skill assessment to inform the choice of the SMAR model adopted for this purpose: instead, it was chosen for its simplicity and comparatively low data requirement. Moreover, model choice is dependent on the project funds available, and the implementers and collaborators are likely to trial their model of choice according to their own interests and increase the scope of its application, irrespective of the underlying model performance measures. In view of the increasing number

of models within the hydrological community, this makes it difficult to determine which model would be the most suitable to use for these operational purposes. Each model also has a varying performance in terms of its ability to predict correctly the timing and magnitude of river flows: this is dependent on many influences, including hydroclimatic and human induced factors. With so many models available it becomes a very challenging task to apply the right model to different regions.

Models cannot be run in space and time without sufficient input data, which take different forms in terms of spatiotemporal variability and some of which may be statistical in nature (Wahren *et al.,* 2016). The availability of input data is crucial to the modelling community, leading to the need for reliable and accurate data. During the initial stages of data acquisition for my research, however, I realised most of the river catchments in Kenya that I had intended to research in fact had no data because they were ungauged. The small amount of data that was available could only be accessed, even for research purposes, for a prohibitively substantial monetary fee as it is not compiled and stored on a dedicated website, but rather it is maintained by the NHMS. During my conversation with Mr Mwai, the deputy director at the Kenya Meteorological Department (KMD), he noted that "presently, FF in Kenya is limited due to inadequate tools (models), model source codes, data, personnel, forecast lead-time, and inadequate documented research to back up operations", which hinders informed flood preparedness actions. These considerations form the key challenges. The acquisition of data for research necessitates the negotiation of complex administrative processes and requires purchases, which raise three further challenges: (i) they may take a long time, due to organizational bureaucracies and purchasing protocols; (ii) the price of the data quotations is very high, because one is required to pay per station requested, which requires a substantial financial investment; and (iii) the pricing/selling agencies may not be willing to share large quantities of data from their database. In my case, I had to model 19 catchments across the country at 19 discharge stations, which are potential hotspots for floods, and required daily *in situ* rainfall, minimum and maximum temperature, and river discharge data. It was nearly impossible to obtain daily rainfall data across the country for research, due to the reasons stated above, and this slowed down research at the initial stages.

After a period of initial frustration, I was introduced to members of the environmental forecasting team from ECMWF, with whom I shared my research aims and objectives and the aforementioned challenges. After a period of deliberations and useful discussions, I was introduced to David Lavers, who oversees the environmental data retrieval and sharing in the forecasting team. The first attempt was to retrieve available daily rainfall data for Kenya from the synoptic weather stations available at the ECMWF climate data store. There were still massive gaps because of the limited synoptic coverage across Kenya but this did begin to help in solving the data scarcity problem.

This is when the idea of evaluating the freely available precipitation products for hydrological modelling was birthed: this sparked an interest in what their performance might be, especially over Kenyan catchments. There are known variations in the data quality of the different products (such as resolution), and so there was a clear need to assess their performance and carefully quantify the uncertainties associated with each of the products. Since my research was focussed on the needs of Kenya, I opted to evaluate reanalysis precipitation products, because evaluation of reanalysis products for hydrological applications is a research topic of

increasing interest. There is a large and rapidly growing body of academic literature that uses precipitation reanalysis data for hydrological modelling at regional and global scales (e.g., Essou *et al.*, 2016; Beck *et al.*, 2017; Essou, Brissette and Lucas-Picher, 2017; Wang, Zhang and Zhang, 2019; Setti *et al.*, 2020; Tarek, Brissette and Arsenault, 2020; Jiang *et al.*, 2021). This formed the background of my study. Additionally, it is also an exploding research area, not least because of the very strong interest in hydrological modelling with earth system approaches supported by global forecasting centres such as ECMWF (see Harrigan, Cloke and Pappenberger, 2020) which made this evaluation timely, especially for Kenya.

Research findings from evaluation of reanalysis precipitation datasets for hydrological modelling revealed variation in performance of different datasets, especially when model simulations from reanalysis were compared to those from observations, not only because of their inherent properties (e.g., model configurations), but also because of the local and mesoscale features (such as catchment characteristics and seasonal changes). For example, the seasonality of rainfall over Kenya is heavily influenced by weather phenomena such as ENSO and IOD (see Section **Error! Reference source not found.** for detailed discussion) and play a major role in extreme rainfall events and inter-annual variability, which reanalysis may fail to capture. These extreme events have contributed to increased flood events over the recent past across Kenya, but first I thought it would be a good idea to establish how these events have changed in the historical flood series, to help to provide information on future trends.

The next step was to explore the trends in river flow series, given the increasing recent flooding events observed in the country (see Section 1.1). Trends in rainfall and temperature over Kenya and East Africa have been explored in research including Langat *et al*. (2017) and Ayugi *et al. (*2016). However, trends in floods have not been researched or documented, yet this is important in revealing flood characteristics which can inform future planning and adaptation measures and the impacts of flooding. As a result, I chose to explore and establish whether there have been statistically increasing or decreasing trends in high flow indices across Kenya flood prone areas. With the changing climate, the intensity and magnitude of extreme events such as floods are likely to increase, and increase the vulnerability of people (Wainwright *et al.,* 2021). As a result, drier months are becoming wet and wet months seeing more rains than usual. This points to a shift in the climatological seasonality in flooding that has been observed in the recent past over the region. However, conclusions cannot be drawn without in-depth analysis, and in the second part of this research, I looked at the possible shifts in flood seasonality and timing across Kenya.

Therefore, my research considers the following aspects:-(i) Model selection framework for a national flood forecasting centre (Appendix A2), (ii) Assessment of reanalysis precipitation products for hydrological modelling in data-scarce regions (Appendix A3), and (iii) Trend detection in river flow series, flood timing and seasonality ((Appendix A4); these are also outlined in Chapters 4, 5 and 6.

CIWEM Chartered Institution of Water and Environmental Management Journal of **Flood Risk Management** **WILEY**

Hydrological model preselection with a filter sequence for the national flood forecasting system in Kenya

Maureen A. Wanzala[1] [ORCID] | **Elisabeth M. Stephens[1,2,3]** | **Hannah L. Cloke[1,2,4,5]** | **Andrea Ficchi[1,6]**

[1]Department of Geography & Environmental Science, University of Reading, Reading, UK

[2]Department of Meteorology, University of Reading, Reading, UK

[3]Red Cross Red Crescent Climate Centre, The Hague, Netherlands

[4]Department of Earth Sciences, Uppsala University, Uppsala, Sweden

[5]Centre of Natural Hazards and Disaster Science, CNDS, Uppsala, Sweden

[6]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy

Correspondence

Maureen A. Wanzala, Department of Geography & Environmental Science, University of Reading, Reading, UK.
Email: m.a.wanzala@pgr.reading.ac.uk

Abstract

The choice of model for operational flood forecasting is not simple because of different process representations, data scarcity issues, and propagation of errors and uncertainty down the modeling chain. An objective decision needs to be made for the choice of the modeling tools. However, this decision is complex because all parts of the process have inherent uncertainty. This paper provides a model selection with a filter sequence for flood forecasting applications in data scarce regions, using Kenya as an example building on the existing literature, concentrating on six aspects: (i) process representation, (ii) model applicability to different climatic and physiographic settings, (iii) data requirements and model resolution, (iv) ability to be downscaled to smaller scales, (v) availability of model code, and (vi) possibility of adoption of the model into an operation flood forecasting system. In addition, we review potential models based on the proposed criteria and apply a decision tree as a filter sequence to provide insights on the possibility of model applicability. We summarize and tabulate an evaluation of the reviewed models based on the proposed criteria and propose the potential model candidates for flood applications in Kenya. This evaluation serves as an objective model preselection criterion to propose a modeling tool that can be adopted in development and operational flood forecasting to the end-users of an early warning system that can help mitigate the effects of floods in data scarce regions such as Kenya.

KEYWORDS

early warning systems, filter sequence, flood forecasting, hydrological model, Kenyan catchments, model preselection, objective model choice, perceptual model

1 | INTRODUCTION

Hydrological models predict the hydrological variables, particularly river flow. In some cases, where little input and output data exist the model can be used to estimate the runoff and river flow in ungauged catchments (Hrachowitz et al., 2013; Sivapalan, Takeuchi, et al., 2003). Therefore, models are useful in applications

such as short to extended-range flood forecasting (Alfieri et al., 2013; Emerton et al., 2018), climate assessment (Hattermann et al., 2017; Lu et al., 2018; Tamm et al., 2016), hazard and risk-mapping (Artan et al., 2001; Ward et al., 2015), drought prediction (Van Huijgevoort et al., 2014), and water resource assessment (Dessu et al., 2016; Mutie, 2019; Praskievicz & Chang, 2009; Sood & Smakhtin, 2015). However, the scope of application to extract viable information varies across different classes of models at different spatial and temporal scales and the intended purpose.

The choice of model for operational flood forecasting is not simple because of different process representations, data scarcity issues, and propagation of errors and uncertainty down the modeling chain (e.g., Paul et al., 2019; Paul, Gaur, et al., 2020). For example, the practice of choosing a model for an application may be difficult due to several reasons highlighted in Melsen et al. (2019):— (i) Popular models are not tailored to specific climate or circumstances (unless the west European climate counts, implicitly), which makes exclusion on process presentation alone difficult; (ii) Most popular models share the same main properties and the same weaknesses; (iii) The community has failed to create a generalized benchmarking system to rank models and model set-ups, so that suitability has to be ascertained on a case-by-case basis; and (iv) Model evaluation takes primarily place based on streamflow, which in itself is too little to distinguish between models, especially calibrated models. There is need for a modeler to know the perceptual model (Wagener et al., 2021)—quantitative or qualitative description of the existing knowledge and understanding of the catchments (Beven, 2011; Gupta et al., 2008; Westerberg et al., 2017). For instance, Wagener et al. (2021) illustrate a generic perceptual model included in catchment hydrology functions. The processes herein are dynamic and evolve with time in response to changes in water management or land-use, climate conditions and geomorphological changes, thus need to be integrated into the model development. This implies that if such changes are not taken into consideration during and/or model development and upgrade, then the relevant processes will not be presented adequately, thus limiting of the application of a single model over the entire country.

Models are simplifications of reality and thus cannot completely represent every process and aspect of the catchment. The importance and impact of many processes can evolve with time for example in response to changes in water management. In addition, what is the right approach *now* is not necessarily the right approach in the future. Significant buy-in is required to develop operational forecasting capacity with a specific model, and so in recognition of changes in the importance and impact of many processes as a result of land use change, water management etc., it may mean it is more efficient to choose a modeling approach that can represent a larger range of processes. When there are distinct zones of hydro-climatology within a country it could be necessary to adopt different modeling approaches, but this needs to be balanced against the scaling up of the resources required to have human and technical capacity across several different models.

Moreover, data play an important role in hydrological modeling irrespective of the processes represented in a model (Wahren et al., 2016). Many studies point to challenges in modeling due to data scarcity (e.g. Beck et al., 2017; Fuka et al., 2014; Lavers et al., 2012; Najafi et al., 2012; Quadro et al., 2013; Smith & Kummerow, 2013; Wu et al., 2013) which limits the applications of very detailed and complex models due to inherent unquantified uncertainties. Recognizing that data and model are not independent of the errors, for brevity within this paper we describe the aspects and the models herein considering only those uncertainties related to model structure (Pechlivanidis et al., 2011; Smith et al., 2015).

The choice of model depends on the intended purpose, and the modeler needs to objectively select a model based on the end-user needs for more reliable decisions (Parker, 2020; Boelee et al., 2017; Todini, 2007). Various hydrological models exist at different spatial and temporal scales with diverse levels of complexity and data requirements. Additionally, there exists differences between model codes and implemented modeling systems, which may cause difficulties in the choice and application of a particular model. A Multi Criteria Analysis (MCA Sherlock & Duffy, 2019) is recommended to evaluate and grade models from which, a small number of models would be constructed, calibrated, and tested in a real-world context and at the end, a model(s) is chosen to be used in the operational Flood Forecasting Centre (FFC) experiment. However, the proposed MCA relies heavily on evaluation data, is very time consuming for the number of models available hence for data scarce regions, and/or agencies with limited resources, (or in general) an additional decision tree is helpful to trim down the number of options. There is the need to further evaluate the limited selection with for example an MCA and the FFC experiment. To aid this hypothetical modeler there is a clear need for well-conceived and systematic strategies for selecting model structures and establishing data requirements, which forms the novelty of this research.

A plethora of model reviews exist at global and continental scale applications. For example, (Devia et al., 2015; Emerton et al., 2016; Kauffeldt et al., 2016; Pechlivanidis et al., 2011; Salvadore et al., 2015; Sood & Smakhtin, 2015; Trambauer et al., 2013). Most of these reviews highlight and compare existing modeling concepts and gaps but none have focused on model selection

frameworks for final application except for Trambauer et al. (2013) and Kauffeldt et al. (2016). Kauffeldt et al. provide a technical review of large-scale hydrological models for implementation in operational flood forecasting at continental level. Trambauer et al. (2013) review continental scale hydrological models highlighting their suitability for drought forecasting in sub-Saharan Africa. The two cited works look at model review and a selection framework for flood and drought application at continental scales respectively and to the best of my knowledge this is the first model overview and practical objective model selection framework for flood applications at national scale taking into consideration varied catchment characteristics and data scarcity issues.

This paper is to propose a practical approach building on Kauffeldt et al. (2016) and Trambauer et al. (2013) for selecting a model based on a step-by-step filter sequence following objective aspects (such as on the ability to simulate relevant processes to flood applications), as well as considering more practical aspects such as model code availability and ease of use at catchment scale with varied climate characteristics. We follow the filter sequence and develop a Venn diagram to select suitable model candidates. This practical approach is applied to a case study of developing an early warning system that can help mitigate the effects of floods in data scarce regions within Kenya, where there is lack of good observations of climate variables such as precipitation, temperature etc., and this is a limiting factor to properly identify the limitations of model applications at catchment scale.

Our paper is structured as follows. Kenyan hydrology and applications of hydrological models to simulations of flood process is discussed in Section 2. The decision tree is built based on deliberations about Kenyan hydrology and current forecasting experience in Kenya, which outlined in Section 3. The selection of the models based on the decision tree is outlined in Section 4. In Section 5, we focus on specific discussions regarding model selection and how the novelty of the decision tree. The paper then concludes with the key contributions of the suggested pre-selection along with recommendations for next steps to evaluate the models objectively to improve FF in Kenya.

2 | KENYAN HYDROLOGY AND FORECASTING

2.1 | Applying hydrological forecasting models to the simulations of floods in Kenya

It is important to consider the application of the hydrological model when determining which model to use, due to differences in process generations and representations

(Cloke et al., 2011). For example, floods are generated by a range of processes related to extreme rainfall (interception, through-flow), runoff generation process (infiltration, saturation excesses and subsurface storm flow) and runoff routing (Rosbjerg et al., 2013). In addition, floods in snow dominated catchments are regularly caused by snow melt, thus, representation of this process in a hydrological model is crucial, because requires an optimal simulation of the snow related hydrological processes such as snow accumulation and melt (Verzano, 2009). However, this case does not apply to Kenyan catchments.

Moreover, flood formation is a complex combination of extreme precipitation or temperature rise or a combination of both, the retention of the water in different storages and finally the flowing through the river networks. A flood peak caused by extreme rainfall in the upstream part of a catchment, naturally reaches the downstream part of the catchment temporally delayed (Tallaksen & Van Lanen, 2004; Verzano, 2009). Therefore, several effects influence the magnitude of the flood wave in the downstream area, such as tributary contributions and retention in lakes and wetlands. The lateral transport of water through the river network is a particularly important process for the routing of discharge. This applies for average flow conditions as well as for low or high flows. Therefore, it is meaningful to route the water within a hydrological model with a variable flow velocity because the flow velocity varies with the actual river discharge (Verzano, 2009) among other relevant flood generating processes. In many hydrological forecasting systems, the treatment of the rainfall-runoff component (traditionally the core of what is meant by hydrological models) and the routing can be separated. If the routing should be built in, or should be specifically modular, could be another criteria that qualifies the models under consideration. An operational Flood Forecasting System (FFS) aims at producing accurate timely and valuable flood forecast information way in advance to reduce flood-related losses by increasing preparation time. A typical FFS requires a hydrological model, data sources, as well as main processes and fan interactive friendly user interface. For example, Figure 1 shows a simplified conceptual model for a large-scale flood forecasting system, the components required, and the output generated within each component.

2.2 | Subsequent logic for the need of a decision framework with a filter sequence in Kenya

Both the hydroclimate and the human influences create challenges for hydrological modeling and forecasting (Bai
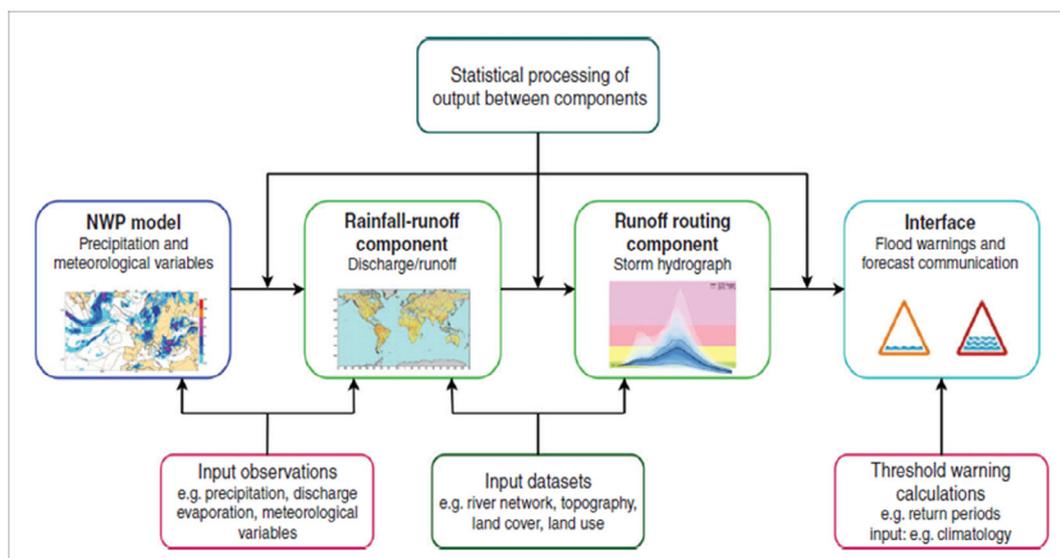
FIGURE 1 Conceptual large-scale hydro-meteorological forecasting system (Emerton et al., 2016)

et al., 2015) because of their massive influence on the catchment processes. For example, Kenya exhibits high variability in physiographic and hydroclimatic conditions (see Figure 2). The highest point is at about 5000 m a.s.l, (mostly areas around central highlands) while the lowest point is about 20 m a.s.l. (mainly around coastal areas). The vegetation cover is mainly a mixed tree cover, grass, and sparse vegetation in most of parts of the country and shrubs and bare land in the arid and semi-arid areas of northern Kenya. As a result, Kenya experiences different climate-related extremes in terms of intensity, magnitude, and timing.

Rainfall pattern follows a bimodal rainfall seasonality (Ongoma & Chen, 2017) with high spatiotemporal variability (Figure 3) (Hession & Moore, 2011). Three seasons are experienced: the "long rains" of March-April-May (MAM), nonrainy months of June-July-August (JJA) and, the "short rains" of October-November-December (OND) (Ogallo, 1988; Ongoma et al., 2015). About 42% of the total annual rainfall is observed during MAM rainfall season (Ongoma & Chen, 2017), with the highest intensity observed near the water bodies of the Indian Ocean, Lake Victoria, and the Kenyan highlands. Freely available packages, proposed models, and inbuilt model functionalities of some of the commonly applied models.

There are five major basins (Marwick et al., 2014) in Kenya (see Figure 4, left panel). These catchments are highly influenced by settlements as well as human activities such as dam constructions and irrigation activities (Figure 4, right panel), which have adverse effects on the catchment response to rainfall runoff processes. At the catchment scale, there is high variability in catchment hydroclimatic characteristics such as surface area and average annual rainfall (Figure 5).

Therefore, it is important to consider the variability in catchment characteristics and the knowledge gaps in the perceptual model (e.g., land cover changes, human activity, data uncertainty and accounting for groundwater fluxes) when selecting a model for application as this may influence the performance of the model. The following section discusses the aspects to consider to objectively preselect a model for application to Kenyan catchments.

3 | MODEL SELECTION FRAMEWORK

Selection framework in this paper follows a selection criterion such as the ability to represent relevant processes, the model structure, flexibility, complexity, availability of the model code and the needs of the user community (Bennett et al., 2013; Kauffeldt et al., 2016), and as such it is more qualitative rather than quantitative. For example, a good model should be able to represent all relevant process such as: gross precipitation (snow, rain), interception storage, evaporation, transpiration, snowpack storage, snowmelt, overland flow, soil storage, recharge to shallow aquifer, capillary rise, intermediate flow, baseflow, leakage to deep aquifer. However, these will require relevant input datasets and more complex models(e.g. fully distributed with numerous parameters) to effectively represent the processes, but worthy to note that increasing the model complexity by incorporating all the above processes does not necessarily increase the model performance (Birkel et al., 2010; Butts et al., 2004). The application and performance of a model may also vary depending on the site (size and characteristics) (Bai
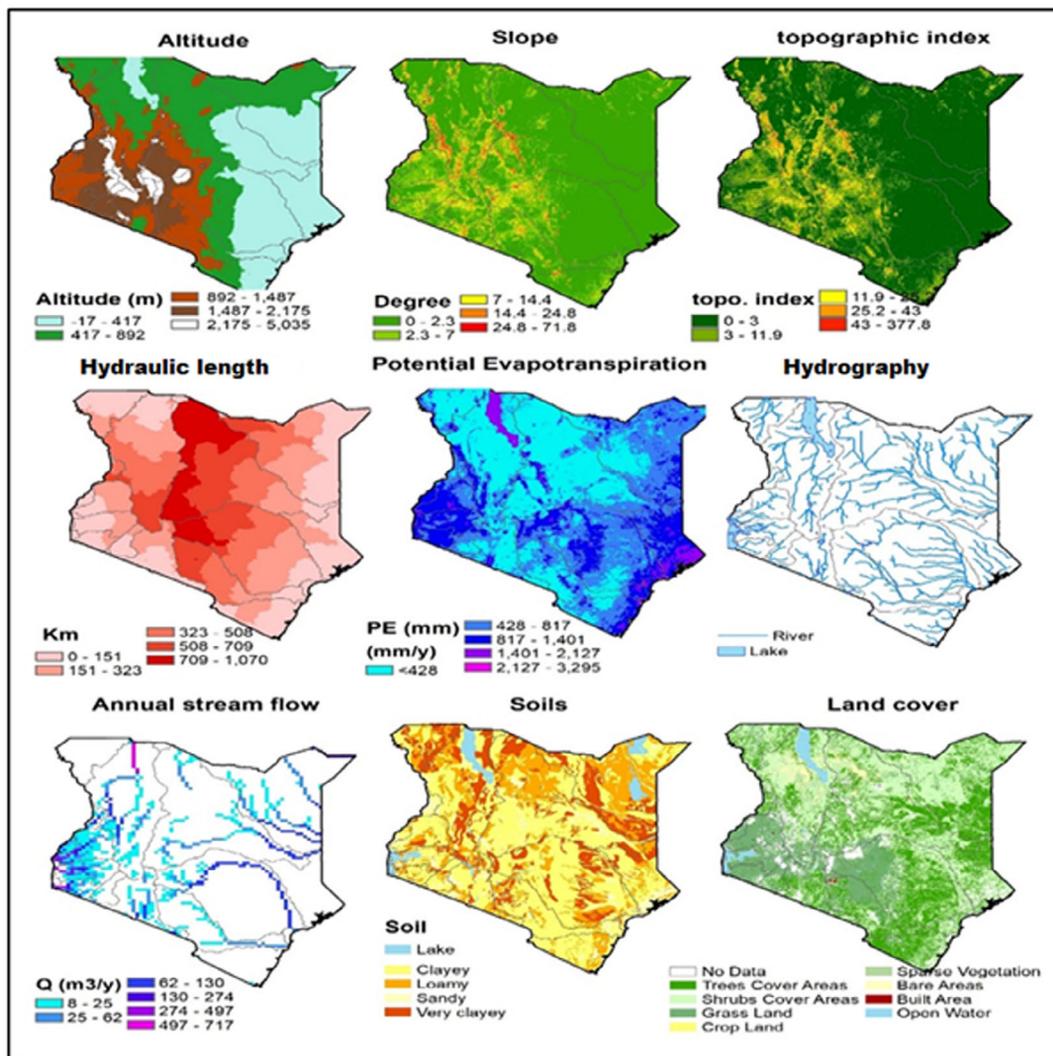
FIGURE 2 Physiographic and hydroclimatic characteristics of Kenya
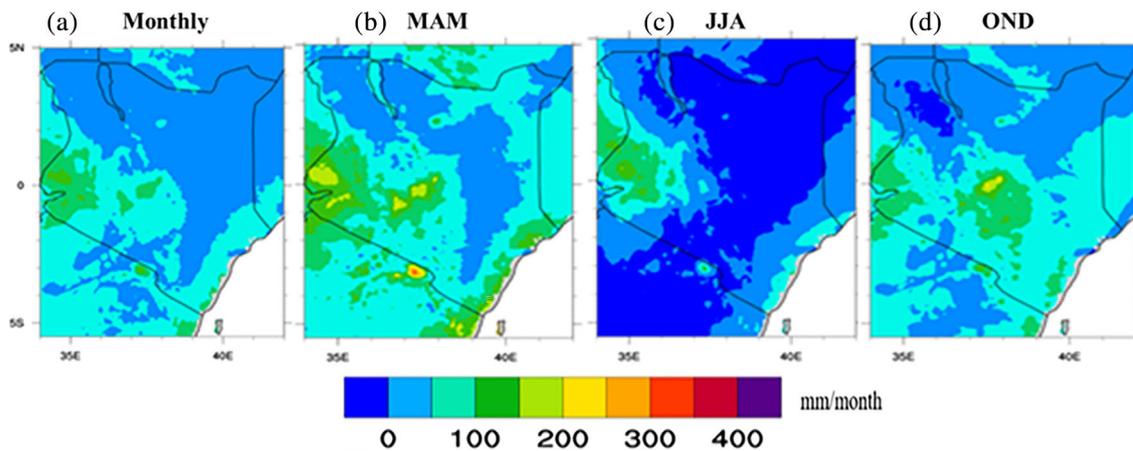
FIGURE 3 Spatial pattern of long-term mean monthly and seasonal rainfall over Kenya (1981–2016) (a) monthly (b) March-April-May (c) June-July-August (d) October-November-December seasons, respectively

et al., 2015; Lanen et al., 2013). Therefore, the following sections summarizes the aspects to aid in objective selection of a hydrological model for flood applications in

Kenyan context, considering Kenya's hydrogeology, physiographic and climatic conditions discussed in Section 1. In total six criteria were found to aid in the decision
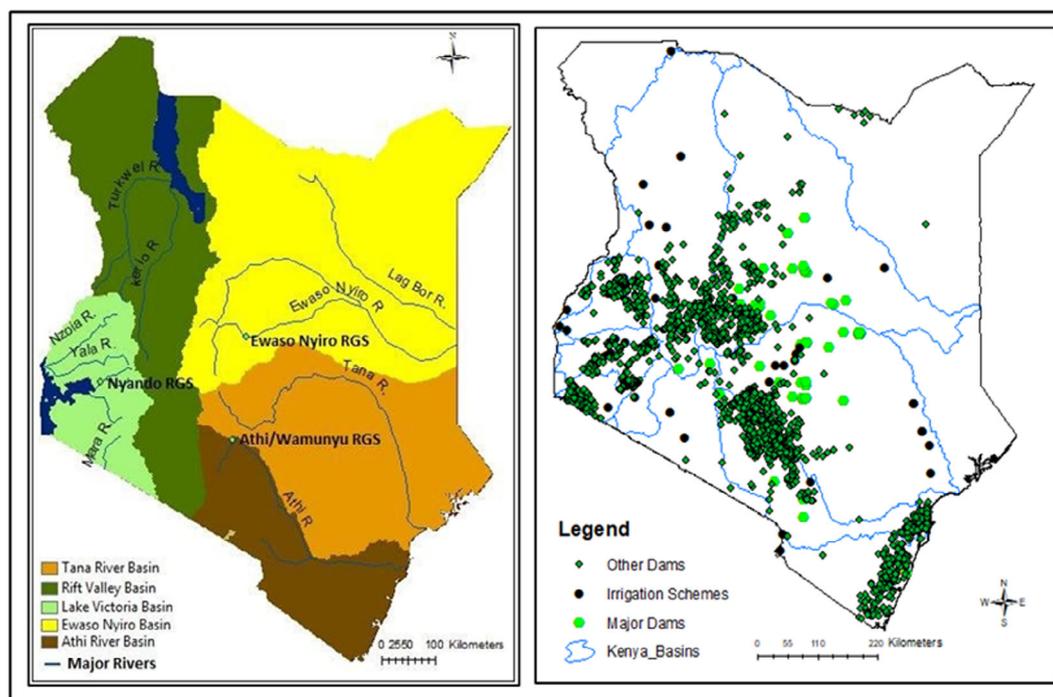
FIGURE 4 Kenya main basins (left panel) and the ongoing human activities (constructed major and small (other) dams and irrigation schemes) in the select catchments (left panel)

making. In the next subsections each of the six criteria is evaluated in relation to Kenya represented processes and fluxes.

A complete hydrological model would represent all the water balance components and fluxes (e.g., as illustrated in Mendoza et al., 2012). The complexity of models often results in many parameters to be determined, which requires more data on hydrogeology (Dobler & Pappenberger, 2013; Muleta & Nicklow, 2005). There needs to be a compromise between model complexity and efficiency for it to work.

More data is needed to make more complex models more accurate. The choice of an appropriate model structure is a crucial step to accurately predict streamflow or other variables, and to understand the dominant physical controls on catchments' responses to climate change (Clark et al., 2008). In Kenya, this requires more data such as groundwater level, which is not readily available.

Some catchments especially those in the arid and semi-arid regions of Kenya have sandy and rocky riverbeds and tend to run dry most of the dry months, for such, the fixed velocity and river channel fields represented in some hydrological models may not apply. This is because of failure to properly represent the roughness index which varies not only with boundary characteristics but also with flow velocity, water depth, and other hydraulic factors (Addy & Wilkinson, 2019; Zhang et al., 2016).

In addition, the more represented process in a model the more the parameterization schemes. For example, a priori estimation requires establishing parameter values from measured physical system properties, presupposing that the model parameters have a sufficiently reliable representation (Beven & Pappenberger, 2003). Therefore, parameter estimation in models of natural systems may require measurements and tests. It then follows that, for effective calibration for such model parameters, it requires more computational power, which may be lacking in the Kenyan operational flood forecasting center.

3.1 | Model applicability to Kenyan hydroclimatic conditions and physiographic settings

Processes that are most relevant for simulating flood conditions in Kenya (see Barasa et al., 2018; Onyando et al., 2003) should be represented in a model. Some extra processes, such as channel losses, evaporation from rivers, wetlands representations, are not considered important in average conditions in some regions due to complexity or lack of interest (Rosbjerg et al., 2013), thus such can be discounted. This is because models incorporating such complex process require more skilled personnel and higher budgets to install and run. This is a challenge in most operational systems in developing
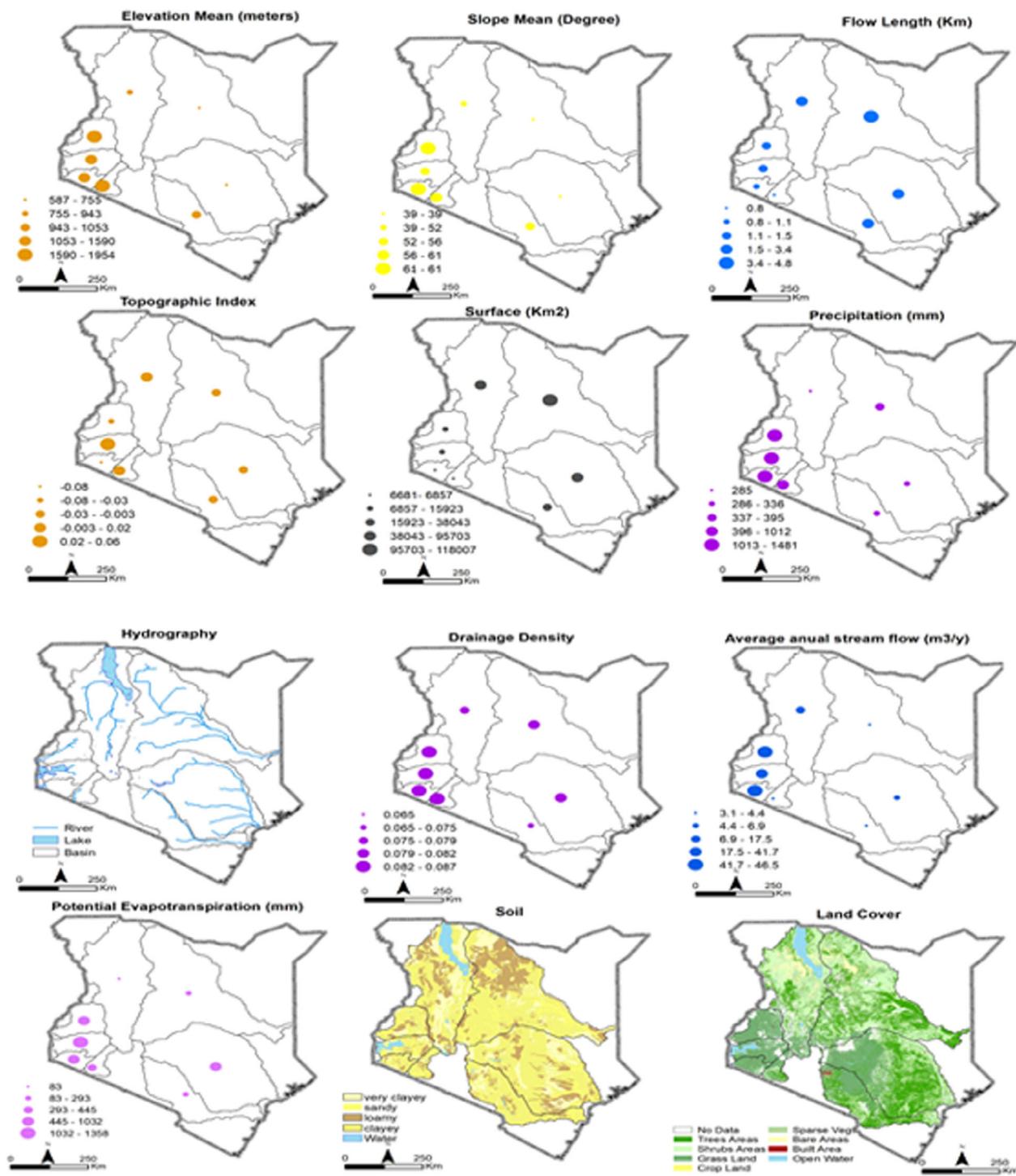
FIGURE 5 -Spatial distribution of the morphological and hydroclimatic characteristics per catchment

countries including Kenya. Temperature plays an important role in river channel and catchment evaporation. In Kenyan case, annual means temperatures range from 15 to 35°C which highly correlates with topography, with the lowest temperature experienced in the central highlands and high temperature in lowlands (Mutimba et al., 2010) and a model incorporating such would be best suited for such place.

Model selection, in dry and wet catchments must be more careful due to the large performance difference in dry catchments (Bai et al., 2015). Wet catchments runoff simulation is significantly better than that in dry catchments, (Haddeland et al., 2011), because of high nonlinearity and heterogeneity of rainfall–runoff processes (Atkinson et al., 2002). In addition, high uncertainty is introduced during model parameter estimation resulting

in significant differences in simulated runoffs behavior (Andersson et al., 2015). Large river basins are often strongly influenced by human activities (e.g. irrigation, reservoirs, and groundwater use) for which information is rarely available (Döll et al., 2009). The Kenyan case where most basins are ungauged may increase such uncertainty (Hrachowitz & Weiler, 2011; Sivapalan, Takeuchi, et al., 2003).

When there are distinct zones of hydroclimatology within a country it could be necessary to adopt different modeling approaches, but this needs to be balanced against the scaling up of the resources required to have human and technical capacity across a number of different models, which is one of the main challenges in the Kenyan case.

3.2 | Data requirements and spatial and temporal resolution of the model

Kenya suffers from lack of good observations of climate and hydrological data. This is a limiting factor to properly identify the limitations of model applications at catchment scale. For example, a detailed representation of groundwater flows and tables and soil moisture content would be very relevant for flood forecasting. However, there is no reliable data (such as ground water and reservoirs) available for research applications, thus limiting the use of model incorporating such kind of data. As a result, a compromise must be reached regarding model spatial variability due to the ungauged status of most Kenyan catchments (Trambauer et al., 2013), and allow the use alternative freely available remote sensing data. Applying a distributed model would require high spatial and temporal resolution data to represent each of the catchment HRUs whereas a lumped conceptual model would represent an entire river basin (Krysanova et al., 1999), but since there are sparsely or no gauging stations in some of the catchments, then this limits the use of most distributed models across Kenyan catchments. However, limiting the models to the type that can only run when directly calibrated on an outlet would be a mistake. This is because there are plenty of ways to discretize in HRUs without individually calibrating each HRU independently. There are ways to calibrate transfer functions to enable modeling and ungauged HRUs (Samaniego et al., 2010). There are model setups that do not rely on calibration as a first principle (such as wflow-sbm, Imhoff et al., 2020) and based on globally available data. The challenge here is the transferability of the model to suite Kenyan catchment and operations and represents the catchment processes adequately because it needs to be as simple as possible.

Moreover, modeling experiments on Kenyan catchments may yield more plausible results if data at high frequency time steps are used as it contains more information (Ficchì et al., 2016). This is because the better modeling of the rainfall–runoff relationship is highly affected by subhourly dynamics of precipitation (Paschalis et al., 2014) due to nonlinear nature of infiltration process (Blöschl & Sivapalan, 1995), such as the peak discharge value (Gabellani et al., 2007) and runoff volume (Viglione et al., 2010). In Kenya, the temporal resolution of the available reliable data may be limited to a higher time steps (such as monthly and yearly) and this may limit the application of a model on a subdaily/hourly timesteps. Models incorporating higher timesteps data such as daily and monthly are easily applicable in Kenyan case as compared to those limited to hourly or subdaily timesteps.

3.3 | Capability of the model to be downscaled to a river basin scale

The issue of scale problem in hydrological models is highlighted in Beven (1995), where the aggregation approach toward macroscale hydrologic modeling is an inadequate approach to the scale problem. For semi-distributed and distributed models, grid size selection is intricately linked to the spatial scale at which the model will be applied. Also, when lumped approaches are applied to considerably larger basins the integration of the processes will naturally occur over a greater area, and thus any differences in small-scale processes within the basin will not be well considered. Due to lack of locally developed models, the continental models are applied at catchment scale, thus the need to be downscaled to suit the grid size under application. However, for larger grids, processes that are only important at the local scale (such as overland flow) may not be considered in the model structure but only if there is an extensive change in the model grid width and this may at times introduce structural uncertainties. Some models may not be easily downscaled to Kenyan river basins with varying spatial scales (see Table 1) without making significant changes in the structure of the model.

3.4 | Operational model for flood early warning system at large scales with potential adoption to local scale

With the increase in flood events in Africa in the recent past, Thiemig et al. (2015) proposed a FFS for Africa hereafter referred to AFFS. Following the illustration in

TABLE 1 Freely available packages, proposed models, and inbuilt model functionalities of some of the commonly applied models (*Source:* Astagneau et al., 2021)

Package	Repository	Proposed model	Package functionalities						
			Data processing	Criteria	Data transformation	Automatic calibration	Plot function	Graphical user Interface	Independent snow function
airGR	(logo)	GR models	√	√	√	√	√	√	√
dynatopmodel	(logo)	Dynamic TOPMODEL	√	√	×	×	√	×	×
HBV.IANIGLA	(logo)	HBV	√	×	×	×	×	×	√
hydromad	(logo)	IHACRES AWBM GR4J Sacramento	√	√	√	√	√	×	√
sacsmar	(logo)	Sacramento	×	×	×	×	×	×	√
topmodel	(logo)	TOPMODEL 1995	√	√	×	×	×	×	×
TUWmodel	(logo)	Modified HBV	×	√/×	×	√/×	×	×	×
WALRUS	(logo)	WALRUS	√	√	√	√/×	√	√/×	√

Figure 8, LISFLOOD, physical-based hydrological model is selected for AFFS, which relies on historical hydrological observations, historical as well as near real-time meteorological observations, real-time meteorological forecasts, and an African GIS dataset. The four main processes AFFS runs are: the calculation of hydrological thresholds, the computation of the initial hydrological conditions, of the computation of the ensemble hydrological predictions, and the identification of flood events. Figure 8 shows a schematic overview of AFFS with all the components and processes. This was developed as a prototype for Africa but never taken forward to operations and since then no literature or research on the skill or applicability of this system has been documented.

Also, Princeton University has developed African Flood and Monitor (AFDM) tool (Sheffield et al., 2014). The aim is to demonstrate the potential for tracking drought conditions across Africa using available satellite products and modeling in data scarce region. The system provides daily updates in near real-time (2–3 days lag) of surface hydrology, streamflow and vegetation stress, short-term hydrological forecasts for flooding, and seasonal forecasts for drought and agricultural impacts as demonstrated in Figure 9 (https://platform. princetonclimate.com/platform-ng/pca/products).

The system has been installed at regional centers in Africa more notable in West Africa (ACMAD), where it is operational for the Niger basin using the Hype-Niger model and the World-Hype applied to the whole West Africa region. A schematic illustration of the FFS for the Niger Basin in West Africa is shown in Figure 6.

Narrowing down to Kenya, the Kenya Meteorological Department runs an operational flood forecast system in Nzioa basin (Personal communication from Andrew Njogu) with plans underway to upscale to other nine additional flood prone areas spread across the other seven basins (Athi, Galana, Sabaki, Nyando, Tana, Sondu and Ewaso Ngiro etc.). A schematic representation of the

FFS in River Nzioa Basin in Kenya and the steps involved is illustrated in Figure 7. The models adopted for this system is the Soil Moisture Accounting and Routing Model (SMAR) incorporated in the Galway Flow Forecasting System (GFFS) (O'Connor, 2005). The GFFS is a suite of models developed at the department of engineering hydrology national university of Ireland, Galway, Ireland. The five models embedded in the software are system theoretic models; simple linear model (SLM), linear perturbation model (LPM), linearly varying gain factor model (LVGF), and artificial neural network (ANN) and one conceptual model which is SMAR model. Ordinary least square solution for (SLM, LPM, and LVGF), conjugate gradient algorithm for ANN and Rosenbrock, simple search and genetic optimization methods for (SMAR) are used for calibration of the model parameters (O'Connor, 2005).

Speaking to Njongu in an interview, he noted that the choice and use of the SMR model was entirely subjective mainly driven by the project funding following the push to implement a FF system in Nzioa after subsequent destructive flooding events. Additionally, he noted that there is limited documented research on skill assessment inform the choice of the SMAR model adopted for this cause, but rather due to its simplicity and less data requirement. Moreover, model choice is dependent on the project funds available, and the implementers and collaborators are likely to trial out their model of choice based on their interests and advance the application scope, irrespective of the underlying model performance measures. It then follows that the choice and application of the SMAR model in the Kenyan FFS was due to the above reason.

With the current developments, there has been ongoing initiatives spearheaded by the Kenya Water Resources Authority—a parastatal mandated to set and manage the water resources rules and hydrological data. Under the ongoing project—Kenya Water Security and
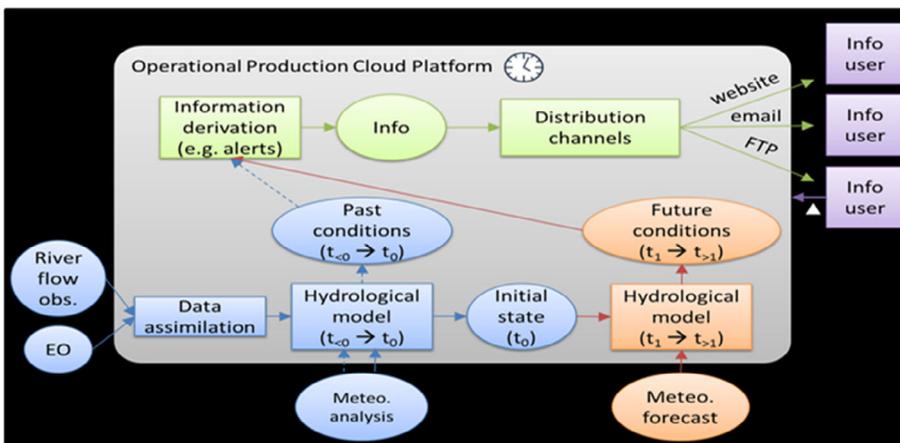
FIGURE 6 Niger HYPE-model for Niger basin and Hype world for the rest of West Africa (https://fanfar.eu/ production/)
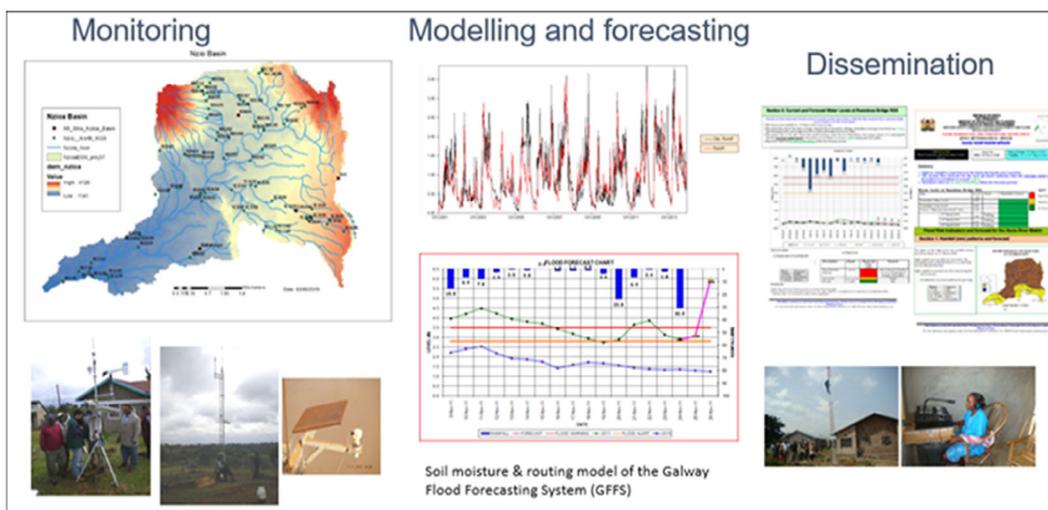
FIGURE 7 Overview of the flood monitoring, modeling, forecasting and dissemination for the operational flood forecasting in Nzioa basin, Kenya (*Source*: Kenya Meteorological Department)

FIGURE 8 A representation of the soil moisture evaluation in soil moisture accounting and routing model (yellow), GR4J (blue) and MIKE NAM (green) over the Nzioa basin in Kenya (*Source*: Kenya Meteorological Department)
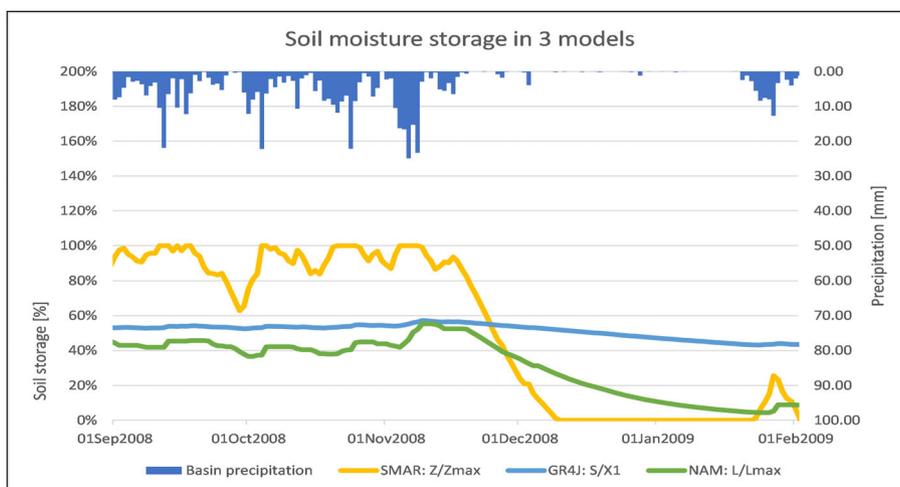
Climate Resilient Project, WRA is in the process of trialing out three hydrological models (SMAR, NAM, GR4J) in Nzoia to be incorporated into the FFEWS under development (WRA reports). For example, an initial assessment for model performance in Nzioa basin has been started. Figure 8 shows soil moisture representation in SMAR, GR4J and MIKE NAM over the basin.

The above highlights point to fact that a model needs to be able to be incorporated into an operational (up and running) system, if the main aim of the model selection is to provide a tool for the end-users of an early warning system that can help mitigate the effects of floods. In this respect, a model that can easily be implemented in a forecasting environment is preferred. Hence, the model should be stable, have reliable error and inconsistency checks, be able to flag off missing data (e.g., when input sources fail), be able to fit into an operational environment and should preferably be user friendly.

3.5 | Availability of model code and model run-time

Code must be available for use (open source or through agreements) with possibilities of adaptation to specific purposes (e.g., possibility to change the represented processes, ingested time-step and/or catchment discretization). These adaptions are possible but not existent in most of the freely available model codes. Code must be actively used and developed with core developers identified to ensure that proper support can be given in initial phases. Executable code is not enough, since changes, for instance, reading of input data will be necessary (Paul, Gaur, et al., 2020). Forecast deliveries run the risk of being delayed if bug fixes or updates cannot quickly be incorporated in the model. Key aspects are the service level agreement struck between the model and the forecasting system provider, outlining a clear overview of

which parts are maintained locally, and which parts are outsourced. In addition, codes available only through purchases may limit the use of models especially for research and operational purposes, thus model should be open source but then not all open-source licenses are the same.

Some modeling communities have availed accessible packages for some select models with dedicated functions such as input data preparations, data processing and transformation, calibration etc. These packages may have all or limited functionalities for the application under consideration, thus limiting its use. For example, Table 1 shows some of the freely available packages, the proposed models to run and the package functionalities that can be executed.

The model run-time (Central Processing Unit)—computational time to run a simulation from model spin up varies with different models and area of application. For example, Astagneau et al. (2021) show how different models and implementations can differ an order of magnitude in required calculation time for the same set of catchments (Figure 5). The computational power lacks in many of the African National Meteorological and Hydrological Services (NMHS), especially if ensemble simulations, data assimilation methods, and further computational intensive uncertainty estimation methods are to be applied, and Kenya is not an exception.

4 | APPLICATION OF THE SELECTION FRAMEWORK TO KENYA′S CHOSEN CATCHMENTS

The above section outlines the aspects to consider when selecting a suitable model for national flood forecasting and application in Kenya. The application of selection framework to Kenya based on the above proposed selection criteria is outlined in Table 2. There are marked differences from catchment to catchment, which point to the fact that a single model single initialization with all the same parameters cannot be suitably applied at country level but rather at catchment scale, thus the need to objectively select a model based on the user needs and catchment processes.

4.1 | Application of decision tree to Kenyan catchments

To assess the suitability of hydrological models with focus to flood applications in Kenya, considering the aspects described, Figure 9 shows a flow diagram of the filter sequence in the selection criteria in defining model suitability to this application which may suffice as a decision

tree. At the top of the decision tree are all the processes that are deemed important in a model for effective flood applications in Kenya. Firstly, Kenya has a large distinction in terms climates, some areas are Arid and Semi-Arid (ASALs) for example Eastern and North-eastern parts, whereas others are wetlands (e.g., Western and Central highlands) (see Figure 2). Therefore, a distinction is made in the second step for processes that are important to the different climatic zones. Secondly, Kenya is currently facing data scarcity due to ungauged nature of many catchments. This, however, should not be a setback to hydrological studies and as a result we filter the model based on the input data availability and possibility to use alternative data. In the third step, we explore the availability of the model code to a wider user community. Here the concept of code executability and online updating, accessibility and the computational run time are explored. At the fourth stage the ability of the model to be downscaled to catchment local scale is considered. Fixed grid sizes and limitations of applicability to certain basin sizes are mainly considered here. Finally, we explore the preferences of the model based on their ease to be implemented in the forecasting system environment. However, this piece of work does not involve the actual analysis of the models under consideration, and it is based on the elimination method following previous studies on the performances of the models over the region. As a result, we present a yes/no decision tree which has a potential implicit weighting factors of "0" or "1" based if the model meets a certain criteria or not from the MCA perspective. The above aspects in the selection framework form the basis of this model overview and selection sections. In this study, a combination of conceptual and process-based lumped and distributed hydrological models are considered for further evaluation to establish if they fit in the above aspects. The hydrological model should be suitable to evaluate the spatial and temporal occurrence of floods based on a defined indicator. Therefore, the models considered (and described in Appendix S1) range from the few applied or under consideration for the Kenyan setup as well as the other widely used models in studies across the African continent for FF that in our opinion would be applicable to Kenyan case. A total of 12 rainfall-runoff models were initially listed as potential candidates for small-scale operational flood forecasting (see Table 3 for main references). LISFLOOD and HYPE are included in this review despite being developed for large-scale applications because they were adopted for the prototype in the AFFS and West Africa, respectively. The models were chosen mainly based on the existing literature reviews and application studies particularly to Africa and Kenya. Table 4 provides a summary of the evaluation of all the 12 models based on the explained criteria herein.

TABLE 2 Summary of catchment-by-catchment evaluation based on the proposed framework

Catchment by catchment evaluation based on the proposed framework									
Catchment name	EWASO NG'IRO	TANA RIVER	ATHI	NZIOA	YALA	NYANDO	TURKWEL	GUCHA	MARA
Catchment Area (km²)	30,000	96,000	46,600	12,800	2777	5110	9,303	6310	13,750
Dominant hydrological processes fluxes present									
Precipitation	√(unimodal)	√(bimodal)	√(bimodal)	√(bimodal)	√(bimodal)	√(bimodal)	√(unimodal)	√(bimodal)	√(bimodal)
Infiltration	√	√(good upstream, poor in delta)	√(good upstream, poor in delta)	√(good upstream, poor in floodplain)	√	√*	√	√	√
Interception	√*	√	√	√	√	√	√*	√	√*
Evapotranspiration	√(low)	√(low upstream, high downstream) Hargreaves	√(low)	√(high)	√(high)	√(high)	√(low)	√(high)	√(low)
Snow	×	×	×	×	×	×	×	×	×
Soil storage	√*	√*	√*	√	√	√	√*	√	√*
Ground water storage	×	√ Shallow and deep	√ Shallow and deep	× ×	×	×	√ Shallow and deep	×	×
Lake and reservoirs	×	√ Linear res.	√ Linear res.	√	√	√	√ Linear res.	√	×
Runoff	√infiltration excess	√	√	√saturation excess/	√saturation excess	√infiltration excess overland	√infiltration excess	√	√
Ground water recharge	×	√	√	√	√	√	×	√	√
Hydroclimatic and physiographic setting									
Arid	√	×	×	×	×	×	√	×	×
Semi-arid	√	√(upstream)	√(upstream)	×	×	×	√	√	√
Wetland	×	√(wet delta)	√(wet delta)	√	√	√	×	√	×
Data availability and resolution									
Observed meteorological data (precipitation, temperature, wind, etc.)	√*Daily √monthly	√*Daily √monthly	√*Daily √monthly	√*Daily √monthly	√*Daily √monthly	√*Daily √monthly	√*Daily √monthly	√*Daily √monthly	√Daily √monthly
Hydrological data (discharge, ground water reservoir levels)	√* discharge for some stations	* Ground water and reservoir levels √* discharge for some stations	* Ground water and reservoir levels √* discharge for some stations	*Lake reservoir levels √* discharge for Ruambwa and missing years for other stations	√* discharge for some years	√* discharge for some years	* Ground water and reservoir levels √*discharge for some stations	√*discharge for some years	√* discharge for some stations and years
Catchment characteristics									
Size (Sq.km)	30,000	96,000	46,600	12,700	3200	3400	28,000	5100	13,750
Human influence	√Irrigation, dams	√Irrigation, dams	√Irrigation, dams	*√Irrigation	*√Irrigation	*√Irrigation	√Irrigation, dams	*√Irrigation	*√Ground water abstraction
Soil type	Sandy/loamy	Sandy/loamy	Sandy/loamy	Loamy/clay	Loamy/Clay	Loamy clay	Sand/loamy	Loamy/clay	Sand/loamy
Vegetation cover	Bare land/Sparse vegetation	Bare land/Sparse vegetation	Bare land/Sparse vegetation	Grassland/Trees/Croplands	Grassland/Trees/croplands	Grassland/Trees/Croplands	Bare land/Sparse vegetation	Grassland/Trees/Croplands	Bare land/Sparse vegetation
Some models calibrated over these catchments									
Model applied to the catchment	√*	×	×	√	×	√	√	√	√

Abbreviations: √, Present; ×, Not Present; ×√, Partially Present.
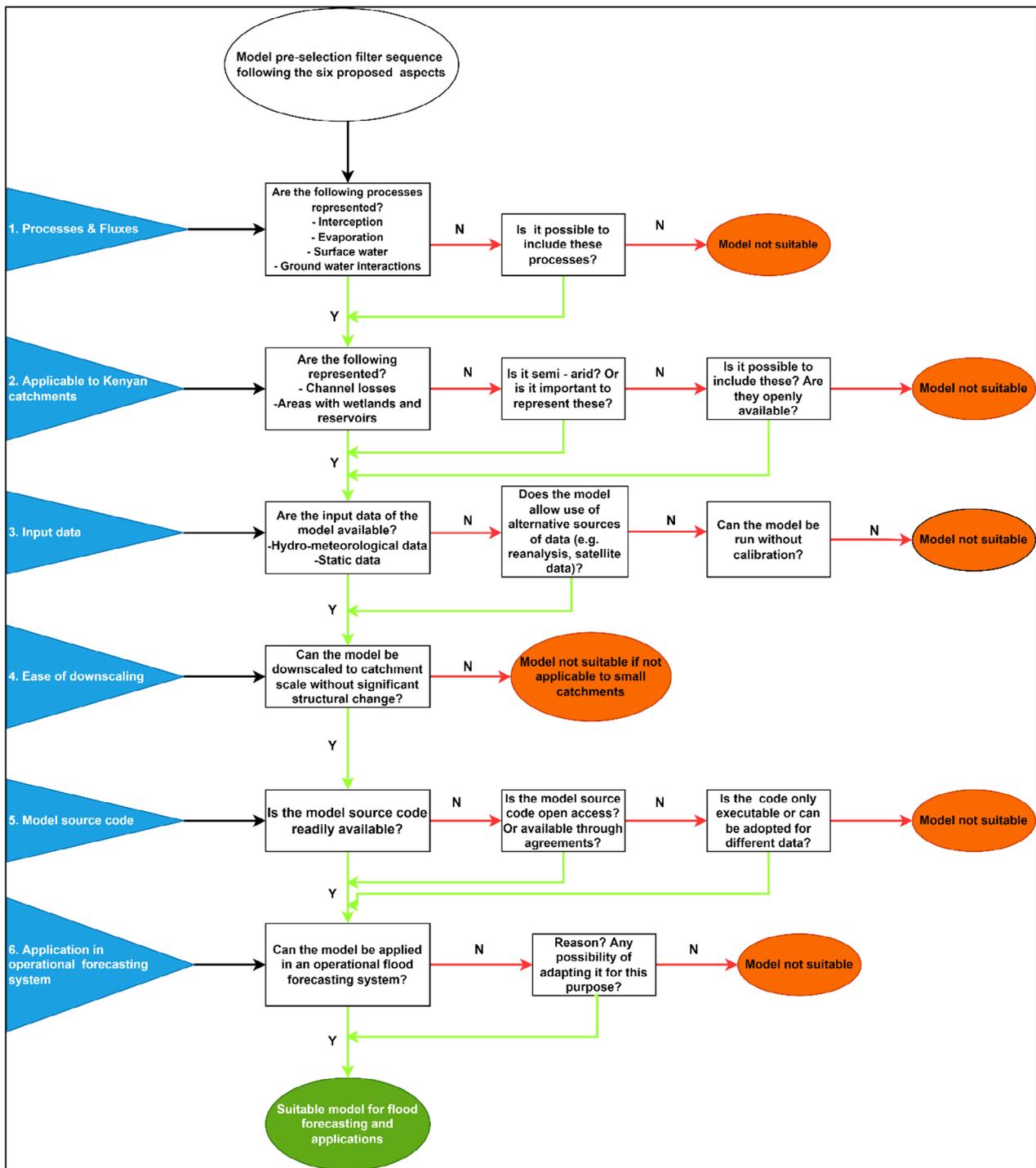
FIGURE 9 Flow sequence to serve as a decision tree for evaluating and selecting a suitable hydrological model for flood forecasting in Kenya, based on the proposed criteria

4.2 │ Actual model selection based on decision tree

The Venn diagram (Figure 10) presents model selection following a comprehensive evaluation carried out in Table 4. All the models under consideration are described and summarized in Appendix S1. Following the filter sequence presented in Figure 4, each model is evaluated on step by step then potential models summarized in actual selection presented in Figure 10 shows a Venn diagram following the framework presented in Figure 9 for the models described and summarized (Appendix S1) and the evaluation information presented in Table 4. Table 4 evaluates all the 12 models based on the

TABLE 3 Twelve rainfall-runoff models listed as potential candidates for small-scale flood applications with their main technical references

Model	Main references
GR4J (modele du Genie Rurala 4 parametres au pas de temps Journalier	Technical (Perrin et al., 2003)
NAM (Nedbør-Afstrømnings-Model)	Technical (Nielsen & Hansen, 1973)
SMAR (Soil Moisture Accounting and Routing)	Technical O'Connor, 2005;
PDM (Probability Distribution Model)	Technical (Goswami & O'Connor 2010; Moore, 2007)
SWAT (Soil Water Assessment Tool)	Technical (Arnold et al., 1998; Neitsch et al., 2005)
MIKE SHE (MIKE Système Hydrologique Européen	Technical (Abbott et al., 1986; Ma et al., 2016)
HBV-96 (Hydrologiska Byråns Vattenbalansavdelning)	Technical (Lindström et al., 1997)
TOPMODEL (TOPography based hydrological)	Technical (Beven and Kirby 1979; Beven et al., 1984)
GeoSFM (Geospatial Streamflow Simulation Model)	(Artan et al., 2001, 2004; Asante et al., 2008)
VIC (Variable Infiltration Capacity)	Technical: (Gao et al., 2010; Lohmann et al., 1996)
LISFLOOD	Technical: (Burek, 2013; van der Knijff et al., 2010)
HYPE (European Hydrological Predictions for the Environment	Technical: (Lindström et al., 2010) http://hypecode.smhi.se

framework presented in Section 2. This provides the summary statistics of each of the models based on the process representation, data input requirements, model code availability, ease of downscaling to Kenyan catchments, and application of models to operational flood forecasting. Out of the 12 models, only VIC and TOPMODEL do not represent important processes for flood generation unique to Kenyan catchments. VIC and TOPmodel were eliminated because it could not represent groundwater processes and requires the calibration of all the parameters which in turn means that the calibration data must be available, which is hardly the case in most of the Kenyan catchments. As a result, they were excluded in the final selection presented in the Venn diagram in Figure 10. The figure shows the flow diagram of the filter sequence which is used to filter out the 12 models to those deemed appropriate candidates for flood applications in Kenya (Figure 10).

From the 12 models reviewed, five are considered suitable candidates for flood applications in Kenyan (Figure 8). The outermost circle (A) presents the 10 models under consideration excluding VIC and TOP-MODEL. VIC and TOPMODEL were not at this point because lack of representing important process such as ground water (see Table 4). In addition, this category includes all the models which can be applied to the study catchments due to reasonable data input requirements, model code availability, ease of downscaling to Kenyan catchments, both in drylands—semi arid—and wetlands, and application of models to operational flood forecasting.

Circle B represents model selection based on data input requirements and the number of calibrated parameters. At this stage, we eliminate LISFLOOD, HBV-96, PDM, GeoSFM and MIKE SHE. LISFLOOD, MIKE SHE, and HBV 96 and GeoSFM are fully and semi-distributed models, respectively, with very many parameters to be calibrated (Berglöv et al., 2009; Ma et al., 2016; van der Knijff et al., 2010). In addition, they are run on hourly timestep with very many data input requirements. The calibration of many parameters will also require intensive computer run time which may be a challenge in many NMHS (Vema & Sudheer, 2020). The ungauged nature of the most of operational centers in Kenya may not have reliable data at high frequency (e.g., at hourly or even daily timesteps). However, circle B is white area because there is the option of alternative remotely sensed data. These models with high data requirements in data scarce areas, there are alternative sources of satellite and reanalyses datasets that are effectively utilized to force the model with caution. This is because the datasets come with their own uncertainties, including random and systematic errors (Fortin et al., 2015; Sun et al., 2018). Inherent input uncertainties will affect the performance of models for a given catchment, and as a result, we eliminated LISFLOOD, HBV-96 and MIKE SHE at this stage. PDM is also eliminated at this point because model configuration comprises of a probability-distributed soil moisture storage, a surface storage, and a groundwater storage components (Moore, 2007). The latter is hardly available input as there is no data on reservoirs and ground water storage in Kenya's NMHS.

Circle (C) represents models which their code is easily available as free open source. This category is meant to rule out models whose codes are available but only in executable format as changes for instance reading of input data may be necessary and is not provided for in executable model codes. The candidate models filtered through to this step HYPE, SWAT, and SMAR have freely available open-source codes (Paul, Gaur, et al., 2020). GR4J and NAM source codes are available through open

TABLE 4 Evaluation of the 12 models based on the proposed framework

Model evaluation

Model name/ criteria	VIC	SWAT	GeoSFM	HBV-96	MIKE-SHE	TOPMODEL
Represented processes and fluxes						
Interception	√ f(LAI)	√(LAI)	√f(forest/land)	*√ (Modified)	√	*√(Modified)
Evaporation	√ Penman-Monteith	√ P-M/P-T/ Hargreaves	√ Penman- Monteith	√	√	√
Snow	√ Energy balance	√Degree day	√ Degree day	√ Degree day	√	√Degree day
Soil	√ 2 or 3 layers	√ ≤ 10 layers	√ 2 layers	√ 2 layers	√ 3 layers	√ 2 layers
Ground water	×	√ Shallow and deep	√ Shallow and deep	×	√ Shallow and deep	√ subsurface and base
Lake, reservoirs	×	√ Linear res.	√ Linear res.	√	√	√
Runoff	√ saturation excess/ function	√ (SCS-CN)	√ (SCS-CN)	√ saturation excess/	√	√ infiltration excess and over land
Routing	√ Linear transfer function	√Muskingum- Cunge	√Muskingum- Cunge/ Diffusion Analogue	√Muskingum- Cunge	√	√
Calibration parameters	*√ Several	√ Several	×	√ Several	√	√ Several
Energy balance	√	×	√	√	√	√
Water use	×	√	√	√	×	√
Data requirement and resolution of the model						
Input Meteorological data	Daily or sub-daily precipitation, air temperature and wind speed	Daily precipitation, minimum and maximum temperature	Daily precipitation, potential evaporation	Daily precipitation, temperature and estimates of potential evaporation	precipitation, air temperature and solar radiation	Precipitation
Model spatial resolution	0.5°	Sub-basins	Semi-distributed	Semi-distributed	Sub-grids	Distributed
Model temporal resolution	Daily	Daily	Daily	Daily	Daily	Hourly/daily
Model code availability						
Open source	×	√	×	√ Executable	×	√ (as an R code called top model
Only Executable			√		√	
Model applicability to Kenyan catchments						
Geographically	√	√	√	√	√	*√ sensitive to grid size (≤50 recommended)
Climatic conditions	* in semi-arid catchments √ in humid catchments	* in semi-arid catchments √in humid catchments	* in semi-arid catchments √ in humid catchments	* in semi-arid catchments √ in humid catchments	* in semi-arid catchments √ in humid catchments	√ in semi-arid catchments √ in humid catchments
The ease of model to be downscaled to river basin scale						
Ease of downscaling without model structure modification	√	√	√	√	√	*√
Models have been calibrated over some Kenyan catchments and applied to operational FF						
Model applied to Kenyan catchment	×	*√	*√	×	*√	×

TABLE 4 (Continued)

	Soil moisture accounting and routing model — PDM	NAM	HYPE	GR4J	LISFLOOD
Represented processes and fluxes					
	√f(canopy)	√f(LAI)	√f(LAI)	√f(LAI)	√f(LAI)
	√	√	√	√	√input
	√Degree day	√Degree day	√Degree day	√Degree day	√Degree day
	√2 layers	√2 layers	√2 layers	√2 layers	√2 layers
	√subsurface and base	√Shallow and deep	√subsurface and base	√	√2 parallel surface
	√Linear res.	√	√Linear res	√	√linear reservoir
	√	√	√	√	√infiltration excess
	√cubic nonlinear	√	√	√	√Kinematic Wave Appr
	√Several	*√9 parameters	√Several	√4 parameters	×
	×	×	×	×	×
	√	×	√	×	×
Data requirement and resolution of the model					
	Daily rainfall and potential evapotranspiration	Daily rainfall, potential evapotranspiration and temperature	Daily precipitation, estimates of potential evaporative	Daily precipitation, estimates of potential evaporative	Daily rainfall, potential evaporative and daily mean air temperature
	0.5°	Lumped	Sub-basins	Lumped	100 m and larger
	Hourly/Daily	Daily	Daily	Daily	Hourly/Daily
Model code availability					
	√	√	√open source	√R-package called airGR	×
			√can be run on subgrid		√
Model applicability to Kenyan catchments					
	√	√	√	√	√
	*in semi-arid catchments √in humid catchments	√in semi-arid catchments √in humid catchments	√in semi-arid catchments √in humid catchments	√in semi-arid catchments √in humid catchments	√in semi-arid catchments √in humid catchments
The ease of model to be downscaled to river basin scale					
	√	√	√	√	*√
	×	*√	*√	*√	×

Models have been calibrated over some Kenyan catchments and applied to operational FF

FIGURE 10 Venn diagram following the model selection procedure, starting with all the all models under consideration in circle A resulting with the selected models in innermost circle D.

A: All models considered suitable for FF

B: Processes of simulating flood and applicability to Kenyan catchments

C: Downscalable to Kenyan catchments based on data and model code availability

D: Models considered for FF in Kenyan catchments

collaborations (Humphrey et al., 2016). The innermost green circle represents models that can be applied easily to Kenyan catchments through simplistic downscaling and suitable for flood forecasting in different Kenyan catchments. Regarding the last criterion as to whether the model is suited for operational purposes, all models reviewed are continuous simulation models and no model is rejected at this step because we assume that, if necessary, they can be modified to be suitable for use in an operational environment.

5 | DISCUSSION

We provide an insight into the need to understanding of the quantitative or qualitative description of the existing knowledge and understanding of the catchments and how this would influence the choice of the modeling tools at catchment scale, acknowledging the gaps and challenges. Models used for different applications in different parts of the world are reviewed based on the six aspects, which builds on the previous works of Kauffeldt et al. (2016) and Trambauer et al. (2013), with the aim of assessing their suitability for flood applications in Kenya. The two foundational works provide a technical review of large-scale hydrological models for implementation in operational flood forecasting highlighting their suitability for drought forecasting at continental level, specifically in sub-Saharan Africa. They are important and provide a comprehensive model review and a selection framework for flood and drought application at continental scales, respectively. However, these studies are applied at a larger scale (continental), yet models simulate process

differently in different hydroclimatic conditions thus, the need to link the process at catchment scale to model specifications and applications.

It can be noted that not all models are good at capturing and or representing the important processes relevant to flood generations (e.g., as transmission losses along the river channel, re-infiltration, and subsequent evaporation of surface) both in wetland and ASALs of Kenya as summarized in Table 2. It should be noted that, with the current data scarcity, most modeling frameworks incorporate satellite and reanalyses data. These products have a coarse resolution and high uncertainty in their estimations at catchment scale, which in turn impacts the model performance. Thus, the way forward for objective choice of modeling tools should ensure that the models are stable, have reliable error and inconsistency checks, be able to flag missing data errors (e.g., when input sources fail), be able to fit into an operational environment and should preferably be user friendly. Considering the data scarcity issues, most models can be implemented as the redundancy related to missing data can be incorporated in the preprocessing. Therefore, if a model can run with missing data, it is a requirement that the run is clearly flagged as having missing data. Model stability can be tested by looking at the distributions of parameters where they became remarkably well-behaved and near-elliptic when numerical error control is implemented in the model (Kavetski et al., 2006). However, since the properties of parameter distributions are dependent on (i) the data, (ii) the model, and (iii) the objective function, testing model stability before application may not be achieved. A sensitivity and uncertainty analysis of model parameters is run to establish model errors, which

should be reliable (Song et al., 2015), but this requires more computational power, which is missing in Kenya.

The practical proposed and presented model preselection with a filter sequence for flood applications was used to filter out models to a subset considered suitable for Kenyan catchment types. Through the filter sequence presented, possible adaptation assumptions are considered in some cases. The filter sequence criteria to assess model suitability including the representation of important processes, availability of the model code, existing user community, input data requirements, possibility of calibration, model resolution and data assimilation and operational implementation into a flood forecasting system. Out of the 12 models, only 5: SWAT, SMAR, GR4J, NAM, HYPE were considered suitable candidates for catchment scale flood forecasting by local authorities in Kenya. The above preselection process forms initial steps and criterion in the choice of a modeling tool to the end-users of to effectively be used both at catchment scale modeling and potentially adopted in an operational early warning system to help mitigate the effects of floods in data scarce regions such as Kenya.

This work does not look at direct analysis of each of the proposed model to evaluate its performance based on some past events. As a starting point, this work provides background of hydrological models and the Kenyan set up to inform a criteria of model preselection for flood applications at national level. The modelers and users of the models can then use the information and arrive at models to apply for some select events. A MCA (Sherlock & Duffy, 2019) forms the basis of these initial steps. The whole process of an MCA is to assess multiple alternatives based on a mix of quantitative and mostly qualitative information from multiple sources. However, the proposed MCA relies heavily on evaluation data, is very time consuming for the number of models available hence for data scarce regions, and/or agencies with limited resources, (or in general) an additional decision tree is helpful to trim down the number of options. There is the need to further evaluate the limited selection with for example an MCA and the FFC experiment. This is mainly because within the same catchment, inhomogeneities of the physical and hydroclimatic properties is a complex issue that is essential in deciding which model to use, thus the importance of the selection criteria.

6 | CONCLUDING REMARKS

There are some challenges that are inherent when applying the above decision framework not only to data scarce regions but also to a wider global scale. For example, with the advancement in research, there is an increasing number of models and none of them is error free, mainly

due to a compromise reached when considering model complexity and computational run time, which is a major challenge (McMillan et al., 2011). Also, it is difficult to balance complexity of model structure, the parameterization and input data requirements, because complex models do not guarantee reliable results (Paul, Gaur, et al., 2020; Trambauer et al., 2013). The use of certain models depends on the computational capabilities (skills) of the individuals as well as the NMHS in general. As a result, model selection may be biased based on the easy of applications depending on the skills of the modeler. In addition, there is no documented research outlining the pros and cons of each of models in a single platform in which a potential model user may easily use to identify which model is suitable (Mannschatz et al., 2016).

To address the highlighted challenges modeling communities in developing countries, Paul, Zhang, et al. (2020) and Souffront Alcantara et al. (2019) suggest some of the way forward. For example, developing countries should consider working on developing their own models. The current models tailored to catchment scale or geographical locations, developed with nicely all year round flowing rivers in relatively wet catchments and the inclusion of a variety of hydrologists and model developers with different needs and perspectives is most welcome and needed to produce hydrological models for a wider range of environments. This may take a long time due to inadequate technological capacities but will suffice as a milestone to addressing some of the challenges associated with model selection. A well-prepared and comprehensive database platform with useful information pooled together, such as:—different input information, advantages and disadvantages of different models is important in providing initial information to judge by eye on which model would work best. This is also likely to facilitate easy model selection alongside frequent webinars by model developers to enhance the skill of modelers in developing countries.

This research provides initial steps to inform the choice of modeling tools in data-scarce region. There is need for further analysis of the proposed model to Kenyan catchments, to assess their skills in simulating the past events. This will provide additional and useful information in the choice and application of these models at catchment scale with varied hydroclimatic characteristics. We acknowledge that it has not proven that the criteria "suffice" as the selection procedure leads because it leads to multiple models and no follow-up strategy is presented here as these forms the basis of future work. Additionally, the filter steps are not operationalized to the level where it can be said to be objective. For example, a model may be excluded on "many parameters" and that the preselection criteria presented here follows a flow chart which may be subjective. First, we make a

preselection based on expert judgment and link to models that have been applied to diversified environments that deem suitable candidates to Kenyan setup.

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT
The data availability statement is not applicable for this study.

ORCID
Maureen A. Wanzala https://orcid.org/0000-0002-9613-1878

REFERENCES
Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System—Systeme Hydrologique Europeen,"SHE", 2: Structure of a physically-based, distributed modelling system. *Journal of Hydrology*, *87*(1-2), 61–77.

Addy, S., & Wilkinson, M. E. (2019). Representing natural and artificial in-channel large wood in numerical hydraulic and hydrological models. *WIREs Water*, *6*(6), e1389.

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013). GloFAS-global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, *17*(3), 1161–1175. https://doi.org/10.5194/hess-17-1161-2013

Andersson, J., Pechlivanidis, I., Gustafsson, D., Donnelly, C., & Arheimer, B. (2015). Key factors for improving large-scale hydrological model performance. *European Water*, *49*, 77–88.

Artan, G., Verdin, J., & Asante, K. (2001). *A wide-area flood risk monitoring model* [Conference presentation]. Fifth International Workshop on Application of Remote Sensing in Hydrology, Montpellier, France (Vol. 2).

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment part I: model development 1. *JAWRA Journal of the American Water Resources Association*, *34*(1), 73–89.

Artan, G. A., Asante, K. O. & Verdin, J. P. (2004). Continental scale flood hazard monitoring system. *AGU Spring Meeting Abstracts*, *2004*, H51D-08.

Asante, K. O., Artan, G. A., Pervez, S., Bandaragoda, C., & Verdin, J. P. (2008). Technical manual for the geospatial stream flow model (GeoSFM). *World Wide Web*, *605*, 594–6151.

Astgneau, P. C., Thirel, G., Delaigue, O., Guillaume, J. H., Parajka, J., Brauer, C. C., Viglione, A., Buytaert, W., & Beven, K. J. (2021). Hydrology modelling R packages–A unified analysis of models and practicalities from a user perspective. *Hydrology and Earth System Sciences*, *25*(7), 3937–3973.

Atkinson, S. E., Woods, R. A., & Sivapalan, M. (2002). Climate and landscape controls on water balance model complexity over changing timescales. *Water Resources Research*, *38*(12), 50–51.

Bai, P., Liu, X., Liang, K., & Liu, C. (2015). Comparison of performance of twelve monthly water balance models in different climatic catchments of China. *Journal of Hydrology*, *529*, 1030–1040.

Barasa, B. N., & Pradeep Perera, E. D. (2018). Analysis of land use change impacts on flash flood occurrences in the Sosiani River basin Kenya. *International Journal of River Basin Management*, *16*(2), 179–188.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., & Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, *21*(12), 6201–6217.

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T. H., Norton, J. P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A. A., Fathi, B. D., & Perrin, C. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, *40*, 1–20.

Berglöv, Gitte, German, Jonas, Gustavsson, Hanna, Harbman, Ulrika, & Johansson, Barbro. (2009). *Improvement HBV Model Rhine in FEWS: Final Report*.

Beven, K. J., Kirkby, M. J., Schofield, N., & Tagg, A. F. (1984). Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. *Journal of Hydrology*, *69*(1-4), 119–143.

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, *24*(1), 43–69.

Beven, K. (1995). Linking parameters across scales: Subgrid parameterizations and scale dependent hydrological models. *Hydrological Processes*, *9*(5–6), 507–525.

Beven, K. J. (2011). *Rainfall-runoff modelling: The primer*. John Wiley & Sons.

Beven, K., & Pappenberger, F. N. (2003). Towards the hydraulics of the hydroinformatics era: By MB ABBOTT, VM BABOVIC and JA CUNGE. *Journal of Hydraulic Research*, *39*(4), 339–349.

Birkel, C., Tetzlaff, D., Dunn, S. M., & Soulsby, C. (2010). Towards a simple dynamic process conceptualization in rainfall–runoff models using multi-criteria calibration and tracers in temperate, upland catchments. *Hydrological Processes: An International Journal*, *24*(3), 260–275.

Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: A review. *Hydrological Processes*, *9*(3–4), 251–290.

Boelee, L., Zsoter, E., Lumbroso, D. M., Samuels, P. G., Bazo, J., & Cloke, H. L. (2017). Estimation of uncertainty in flood forecasts—A comparison of methods. *Journal of Flood Risk Management*, *12*, e12516.

Butts, M. B., Payne, J. T., Kristensen, M., & Madsen, H. (2004). An evaluation of the impact of model structure on hydrological

modelling uncertainty for streamflow simulation. *Journal of Hydrology*, *298*(1–4), 242–266.

Burek, P. (2013). *LISFLOOD, distributed water balance and flood simulation model revised user manual 2013*. JRC Publications Repository.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., & Hay, L. E. (2008). Framework for understanding structural errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2007WR006735

Cloke, H. L., Hannah, D. M., Pappenberger, F., Prudhomme, C., Balsamo, G., Fortin, V., & Sauquet, E. (2011). Large-scale hydrology-advances in understanding processes, dynamics and models from beyond river basin to global scale. *Hydrological Processes*, *25*(7), 991–1200.

Dessu, S. B., Seid, A., Abiy, A. Z., & Melesse, A. (2016). Flood forecasting and stream flow simulation of the upper Awash River basin, Ethiopia using geospatial stream flow model (GeoSFM). In A. Melesse & W. Abtew (Eds.), *Landscape dynamics, soils and hydrological processes in varied climates* (pp. 367–384). Springer.

Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A review on hydrological models. *Aquatic Procedia*, *4*(Icwrcoe), 1001–1007. https://doi.org/10.1016/j.aqpro.2015.02.126

Dobler, C., & Pappenberger, F. (2013). Global sensitivity analyses for a complex hydrological model applied in an alpine watershed. *Hydrological Processes*, *27*(26), 3922–3940.

Döll, P., Fiedler, K., & Zhang, J. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrology and Earth System Sciences*, *13*(12), 2413–2432.

Emerton, R. E., Stephens, E. M., Pappenberger, F., Pagano, T. C., Weerts, A. H., Wood, A. W., Salamon, P., Brown, J. D., Hjerdt, N., Donnelly, C., Baugh, C. A., & Cloke, H. L. (2016). Continental and global scale flood forecasting systems. *WIREs Water*, *3*(3), 391–418. https://doi.org/10.1002/wat2.1137

Emerton, R., Zsoter, E., Arnal, L., Cloke, H. L., Muraro, D., Prudhomme, C., Stephens, E. M., Salamon, P., & Pappenberger, F. (2018). Developing a global operational seasonal hydro-meteorological forecasting system: GloFAS-seasonal v1. 0. *Geoscientific Model Development*, *11*, 3327–3346.

Ficchì, A., Perrin, C., & Andréassian, V. (2016). Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events. *Journal of Hydrology*, *538*, 454–470.

Fortin, V., Roy, G., Donaldson, N., & Mahidjiba, A. (2015). Assimilation of radar quantitative precipitation estimations in the Canadian precipitation analysis (CaPA). *Journal of Hydrology*, *531*, 296–307.

Fuka, D. R., Walter, M. T., MacAlister, C., Degaetano, A. T., Steenhuis, T. S., & Easton, Z. M. (2014). Using the climate forecast system reanalysis as weather input data for watershed models. *Hydrological Processes*, *28*(22), 5613–5623.

Gabellani, S., Boni, G., Ferraris, L., Von Hardenberg, J., & Provenzale, A. (2007). Propagation of uncertainty from rainfall to runoff: A case study with a stochastic rainfall generator. *Advances in Water Resources*, *30*(10), 2061–2071.

Gao, H., Tang, Q., Shi, X., Zhu, C., Bohn, T., Fengge, S., Pan, M., Sheffield, J., Lettenmaier, D., & Wood, E. (2010). Water budget record from Variable Infiltration Capacity (VIC) model. *Environmental Science*, 120–173.

Goswami, M., & O'Connor, K. M. (2010). A "monster" that made the SMAR conceptual model "right for the wrong reasons". *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, *55*(6), 913–927.

Gupta, H. V., Wagener, T., & Liu, Y. (2008). Reconciling theory with observations: Elements of a diagnostic approach to model evaluation. *Hydrological Processes: An International Journal*, *22*(18), 3802–3813.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., … Gerten, D. (2011). Multimodel estimate of the global terrestrial water balance: Setup and first results. *Journal of Hydrometeorology*, *12*(5), 869–884.

Hattermann, F. F., Krysanova, V., Gosling, S. N., Dankers, R., Daggupati, P., Donnelly, C., Flörke, M., Huang, S., Motovilov, Y., Buda, S., Yang, T., Müller, C., Leng, G., Tang, Q., Portmann, F. T., Hagemann, S., Gerten, D., Wada, Y., Masaki, Y., … Samaniego, L. (2017). Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. *Climatic Change*, *141*(3), 561–576.

Hession, S. L., & Moore, N. (2011). A spatial regression analysis of the influence of topography on monthly rainfall in East Africa. *International Journal of Climatology*, *31*(10), 1440–1456.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., … Cudennec, C. (2013). A decade of predictions in ungauged basins (PUB)—A review. *Hydrological Sciences Journal*, *58*(6), 1198–1255.

Hrachowitz, M., & Weiler, M. (2011). Uncertainty of precipitation estimates caused by sparse gauging networks in a small, mountainous watershed. *Journal of Hydrologic Engineering*, *16*(5), 460–471.

Humphrey, G. B., Gibbs, M. S., Dandy, G. C., & Maier, H. R. (2016). A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology*, *540*, 623–640.

Imhoff, R. O., Van Verseveld, W. J., Van Osnabrugge, B., & Weerts, A. H. (2020). Scaling point-scale (Pedo) transfer functions to seamless large-domain parameter estimates for high-resolution distributed hydrologic modeling: An example for the Rhine River. *Water Resources Research*, *56*(4), e2019WR026807.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., & Thielen, J. (2016). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software*, *75*, 68–76.

Kavetski, D., Kuczera, G., & Franks, S. W. (2006). Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis. *Journal of Hydrology*, *320*(1–2), 187–201.

Krysanova, V., Bronstert, A., & Müller-Wohlfeil, D. I. (1999). Modelling river discharge for large drainage basins: From lumped to distributed approach. *Hydrological Sciences Journal*, *44*(2), 313–331.

Lanen, V., Henny, A. J., Wanders, N., Tallaksen, L. M., Loon, V., & Anne, F. (2013). Hydrological drought across the world: Impact of climate and physical catchment structure. *Hydrology and Earth System Sciences*, *17*, 1715–1732.

Lavers, D. A., Villarini, G., Allan, R. P., Wood, E. F., & Wade, A. J. (2012). The detection of atmospheric rivers in atmospheric reanalyses and their links to British winter floods and the large-scale climatic circulation. *Journal of Geophysical Research: Atmospheres*, *117*(D20). https://doi.org/10.1029/2012JD018027

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, *201*(1-4), 272–288.

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., & Arheimer, B. (2010). Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research*, *41*(3-4), 295–319.

Lohmann, D. A. G., Nolte-Holube, R., & Raschke, E. (1996). A large-scale horizontal routing model to be coupled to land surface parametrization schemes. *Tellus A*, *48*(5), 708–721.

Lu, W., Wang, W., Shao, Q., Yu, Z., Hao, Z., Xing, W., Yong, B., & Li, J. (2018). Hydrological projections of future climate change over the source region of Yellow River and Yangtze River in the Tibetan plateau: A comprehensive assessment by coupling RegCM4 and VIC model. *Hydrological Processes*, *32*(13), 2096–2117.

Ma, L., He, C., Bian, H., & Sheng, L. (2016). MIKE SHE modeling of ecohydrological processes: Merits, applications, and challenges. *Ecological Engineering*, *96*, 137–149. https://doi.org/10.1016/j.ecoleng.2016.01.008

Mannschatz, T., Wolf, T., & Hülsmann, S. (2016). Nexus tools platform: Web-based comparison of modelling tools for analysis of water-soil-waste nexus. *Environmental Modelling & Software*, *76*, 137–153.

Marwick, T. R., Tamooh, F., Ogwoka, B., Teodoru, C., Borges, A. V., Darchambeau, F., & Bouillon, S. (2014). Dynamic seasonal nitrogen cycling in response to anthropogenic N loading in a tropical catchment, Athi–Galana–Sabaki River, Kenya. *Biogeosciences*, *11*(2), 443–460.

McMillan, H., Jackson, B., Clark, M., Kavetski, D., & Woods, R. (2011). Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. *Journal of Hydrology*, *400*(1–2), 83–94.

Melsen, L. A., Teuling, A. J., Torfs, P. J., Zappa, M., Mizukami, N., Mendoza, P. A., Clark, M. P., & Uijlenhoet, R. (2019). Subjective modeling decisions can significantly impact the simulation of flood and drought events. *Journal of Hydrology*, *568*, 1093–1104.

Mendoza, P. A., McPhee, J., & Vargas, X. (2012). Uncertainty in flood forecasting: A distributed modeling approach in a sparse data catchment. *Water Resources Research*, *48*(9), 22–31. https://doi.org/10.1029/2011WR011089

Moore, R. J. (2007). The PDM rainfall-runoff model. *Hydrology and Earth System Sciences Discussions*, *11*(1), 483–499.

Muleta, M. K., & Nicklow, J. W. (2005). Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of Hydrology*, *306*(1–4), 127–145.

Mutie, S. M. (2019). *Modelling the influence of land use change on the hydrology of Mara River using USGS geospatial stream flow model*. JKUAT-COETEC. http://hdl.handle.net/123456789/4986

Mutimba, S., Mayieko, S., Olum, P., & Wanyama, K. (2010). *Climate change vulnerability and adaptation preparedness in Kenya*. Heinrich Böll Stiftung.

Najafi, M. R., Moradkhani, H., & Piechota, T. C. (2012). Ensemble streamflow prediction: Climate signal weighting methods vs. climate forecast system reanalysis. *Journal of Hydrology*, *442*, 105–116.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Williams, J. R., & King, K. W. (2005). *Soil and water assessment tool theoretical documentation version 2005* (Vol. 1). Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Blackland Research Center, Texas Agricultural Experiment Station, Texas, 1.

Nielsen, S. A., & Hansen, E. (1973). Numerical simulation of the rainfall-runoff process on a daily basis. *Hydrology Research*, *4*(3), 171–190.

O'Connor, K. M. (2005). The Galway real-time river flow forecasting system (GFFS). In D. Knight, D. W. Knight, & A. Shamseldin (Eds.), *River Basin modelling for flood risk mitigation* (p. 215). Taylor & Francis.

Ogallo, L. J. (1988). Relationships between seasonal rainfall in East Africa and the southern oscillation. *Journal of Climatology*, *8*(1), 31–43.

Ongoma, V., & Chen, H. (2017). Temporal and spatial variability of temperature and precipitation over East Africa from 1951 to 2010. *Meteorology and Atmospheric Physics*, *129*(2), 131–144.

Ongoma, V., Guirong, T., Ogwang, B. A., & Ngarukiyimana, J. P. (2015). Diagnosis of seasonal rainfall variability over East Africa: A case study of 2010-2011 drought over Kenya. Pakistan. *Journal of Meteorology*, *11*(22), 13–21.

Onyando, J. O., Schumann, A. H., & Schultz, G. A. (2003). Simulation of flood hydrographs based on lumped and semi-distributed models for two tropical catchments in Kenya. *Hydrological Sciences Journal*, *48*(4), 511–524.

Paschalis, A., Fatichi, S., Molnar, P., Rimkus, S., & Burlando, P. (2014). On the effects of small scale space–time variability of rainfall on basin flood response. *Journal of Hydrology*, *514*, 313–327.

Paul, P. K., Gaur, S., Kumari, B., Mishra, A., Panigrahy, N., & Singh, R. (2020). Application of a newly developed large-scale conceptual hydrological model in simulating streamflow for credibility testing in data scarce condition. *Natural Resource Modeling*, *33*(4), e12283.

Paul, P. K., Gaur, S., Kumari, B., Panigrahy, N., Mishra, A., & Singh, R. (2019). Diagnosing credibility of a large-scale conceptual hydrological model in simulating streamflow. *Journal of Hydrologic Engineering, ASCE*, *24*(4), 4019004.

Paul, P. K., Zhang, Y., Ma, N., Mishra, A., Panigrahy, N., & Singh, R. (2020). Selection of hydrological model (s) for study in developing countries: Perspective of global, continental and country scale model. *Journal of Hydrology*, *600*, 126561.

Parker, W. S. (2020). Model evaluation: An adequacy-for-purpose view. *Philosophy of Science*, *87*(3), 457–477.

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflowsimulation. *Journal of Hydrology*, *279*(1–4), 275–289.

Pechlivanidis, I. G., Jackson, B. M., McIntyre, N. R., & Wheater, H. S. (2011). Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST Journal*, *13*(3), 193–214.

Praskievicz, S., & Chang, H. (2009). A review of hydrological modelling of basin-scale climate change and urban development impacts. *Progress in Physical Geography*, *33*(5), 650–671. https://doi.org/10.1177/0309133309348098

Quadro, M. F. L., Berbery, E. H., Dias, S., Maria, A. F., Herdies, D. L., & Gonçalves, L. G. G. (2013). The atmospheric water cycle over South America as seen in the new generation of global reanalyses. In *AIP conference proceedings* (Vol. 1531, pp. 732–735). American Institute of Physics.

Rosbjerg, D., Bloschl, G., Burn, D., Castellarin, A., Croke, B., Di Baldassarre, G., Iacobellis, V., Kjeldsen, T. R., Kuczera, G., Merz, R., Montanari, A., Morris, D., Ouarda, T. B. M. J., Ren, L., Rogger, M., Salinas, J. L., Toth, E., & Viglione, A. (2013). Prediction of floods in ungauged basins. In G. Blöschl, M. Sivapalan, T. Wagener, A. Viglione, & H. Savenije (Eds.), *Runoff prediction in ungauged basins: Synthesis across processes, places and scales* (pp. 189–225). Cambridge University Press.

Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-basedhydrologic model at the mesoscale. *Water Resources Research*, *46*, WO5523. https://doi.org/10.1029/2008WR007327

Salvadore, E., Bronders, J., & Batelaan, O. (2015). Hydrological modelling of urbanized catchments: A review and future directions. *Journal of Hydrology*, *529*, 62–81.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., & Ogallo, L. (2014). A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bulletin of the American Meteorological Society*, *95*(6), 861–882.

Sherlock, E., & Duffy, S. (2019). *Establishing the flood forecast centre and expanding met Eireann's rainfall radar network*. http://hydrologyireland.ie/wp-content/uploads/2019/11/05-Eoin-Sherlock-Establishing-the-Flood-Forecast-Centre-and-Expanding-Met-Eireanns-Radar.pdf

Souffront Alcantara, M. A., Nelson, E. J., Shakya, K., Edwards, C., Roberts, W., Krewson, C., Ames, D. P., Jones, N. L., & Gutierrez, A. (2019). Hydrologic modeling as a service (HMaaS): A new approach to address hydroinformatic challenges in developing countries. *Frontiers in Environmental Science*, *7*, 158.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., & Zehe, E. (2003). IAHS decade on predictions in ungauged basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, *48*(6), 857–880.

Smith, A., Sampson, C., & Bates, P. (2015). Regional flood frequency analysis at the global scale. *Water Resources Research*, *51*(1), 539–553.

Smith, R. A., & Kummerow, C. D. (2013). A comparison of in situ, reanalysis, and satellite water budgets over the upper Colorado River basin. *Journal of Hydrometeorology*, *14*(3), 888–905.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., & Xu, C. (2015). Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. *Journal of Hydrology*, *523*, 739–757.

Sood, A., & Smakhtin, V. (2015). Global hydrological models: A review. *Hydrological Sciences Journal*, *60*(4), 549–565.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107.

Tamm, O., Luhamaa, A., & Tamm, T. (2016). Modeling future changes in the north-Estonian hydropower production by using SWAT. *Hydrology Research*, *47*(4), 835–846.

Tallaksen, L. M., & Van Lanen, H. A. (Eds.). (2004). *Hydrological drought: processes and estimation methods for streamflow and groundwater*. Elsevier.

Thiemig, V., Bisselink, B., Pappenberger, F., & Thielen, J. (2015). A pan-African medium-range ensemble flood forecast system. *Hydrology and Earth System Sciences*, *19*(8), 3365–3385. https://doi.org/10.5194/hess-19-3365-2015

Todini, E. (2007). Hydrological catchment modelling: Past, present and future. *Hydrology and Earth System Sciences*, *11*(1), 468–482.

Trambauer, P., Maskey, S., Winsemius, H., Werner, M., & Uhlenbrook, S. (2013). A review of continental scale hydrological models and their suitability for drought forecasting in (sub-Saharan) Africa. *Physics and Chemistry of the Earth, Parts A/B/C*, *66*, 16–26.

van der Knijff, J. M., Younis, J., & de Roo, A. P. J. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, *24*(2), 189–212. https://doi.org/10.1080/13658810802549154

Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Teuling, A. J., & Uijlenhoet, R. (2014). Identification of changes in hydrological drought characteristics from a multi-GCM driven ensemble constrained by observed discharge. *Journal of Hydrology*, *512*, 421–434.

Vema, V. K., & Sudheer, K. P. (2020). Towards quick parameter estimation of hydrological models with large number of computational units. *Journal of Hydrology*, *587*, 124983.

Verzano, K. (2009). *Climate change impacts on flood related hydrological processes: Further development and application of a global scale hydrological model*. University of Kassel Kassel.

Viglione, A., Chirico, G. B., Komma, J., Woods, R., Borga, M., & Blöschl, G. (2010). Quantifying space-time dynamics of flood event types. *Journal of Hydrology*, *394*(1–2), 213–229.

Wagener, T., Dadson, S. J., Hannah, D. M., Coxon, G., Beven, K., Bloomfield, J. P., Buytaert, W., Cloke, H., Bates, P., Holden, J., Parry, L., Lamb, R., Chappell, N. A., Fry, M., Old, G., & Holden, J. (2021). Knowledge gaps in our perceptual model of Great Britain's hydrology. *Hydrological Processes*, *35*, e14288.

Wahren, F. T., Julich, S., Nunes, J. P., Gonzalez-Pelayo, O., Hawtree, D., Feger, K.-H., & Keizer, J. J. (2016). Combining digital soil mapping and hydrological modeling in a data scarce watershed in north-Central Portugal. *Geoderma*, *264*, 350–362.

Ward, P. J., Jongman, B., Salamon, P., Simpson, A., Bates, P., De Groeve, T., Muis, S., de Perez, E. C., Rudari, R., Trigg, M. A., & Winsemius, H. C. (2015). Usefulness and limitations of global flood risk models. *Nature Climate Change*, *5*(8), 712–715.

Westerberg, I. K., Baldassarre, D., Giuliano, B., Keith, J., Coxon, G., & Krueger, T. (2017). Perceptual models of uncertainty for socio-hydrological systems: A flood risk change example. *Hydrological Sciences Journal*, *62*(11), 1705–1713.

Wu, W., Chen, J., & Huang, R. (2013). Water budgets of tropical cyclones: Three case studies. *Advances in Atmospheric Sciences*, *30*(2), 468–484.

Zhang, S. T., Liu, Y., Li, M. M., & Liang, B. (2016). Distributed hydrological models for addressing effects of spatial variability of roughness on overland flow. *Water Science and Engineering*, *9*(3), 249–255.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wanzala, M. A., Stephens, E. M., Cloke, H. L., & Ficchi, A. (2022). Hydrological model preselection with a filter sequence for the national flood forecasting system in Kenya. *Journal of Flood Risk Management*, 1–24. https://doi.org/10.1111/jfr3.12846
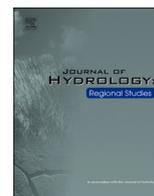
Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: A case study of Kenya

Maureen A. Wanzala [a,*], Andrea Ficchi [a,b], Hannah L. Cloke [a,c,d,e,f], Elisabeth M. Stephens [c,g], Heou M. Badjana [a], David A. Lavers [e,h]

[a] Department of Geography & Environmental Science, University of Reading, Reading, UK
[b] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy
[c] Department of Meteorology, University of Reading, Reading, UK
[d] Department of Earth Sciences, Uppsala University, Uppsala, Sweden
[e] European Centre for Medium-Range Weather Forecasts, Reading, UK
[f] Centre of Natural Hazards and Disaster Science, CNDS, Uppsala, Sweden
[g] Red Cross Red Crescent Climate Centre, The Hague, The Netherlands
[h] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

ARTICLE INFO

ABSTRACT

Study region: 19 flood prone catchments in Kenya, Eastern Africa

Study focus: Flooding is a major natural hazard especially in developing countries, and the need for timely, reliable, and actionable hydrological forecasts is paramount. Hydrological modelling is essential to produce forecasts but is a challenging task, especially in poorly gauged catchments, because of the inadequate temporal and spatial coverage of hydro-meteorological observations. Open access global meteorological reanalysis datasets can fill in this gap, however they have significant errors. This study assesses the performance of four reanalysis datasets (ERA5, ERA-Interim, CFSR and JRA55) over Kenya for the period 1981–2016 on daily, monthly, seasonal, and annual timescales. We firstly evaluate the reanalysis datasets by comparing them against observations from the Climate Hazards group Infrared Precipitation with Station. Secondly, we evaluate the ability of these reanalysis datasets to simulate streamflow using GR4J model considering both model performance and parameters sensitivity and identifiability.

New hydrological insights for the region: While ERA5 is the best performing dataset overall, performance varies by season, and catchment and therefore there are marked differences in the suitability of reanalysis products for forcing hydrological models. Overall, wetland catchments in the western regions and highlands of Kenya obtained relatively better scores compared to those in the semi-arid regions, this can inform future applications of reanalysis products for setting up hydrological models that can be used for flood forecasting, early warning, and early action in data scarce regions, such as Kenya.

1. Introduction

Precipitation is arguably the most important driver of catchment hydrological response (e.g. MacLeod et al., 2021), but it is

challenging to get accurate information on the amount, duration, and intensity of rainfall events (Beck et al., 2017a; Tapiador et al., 2012), due to the high spatio-temporal variability (Nicholson et al., 2019; Vischel et al., 2011). This is compounded by a low spatial coverage and a net decline in the number of ground gauge stations in the historical climatological observation network, especially in developing countries such as Kenya (Menne et al., 2018; Tarek et al., 2020, 2021; Zaitchik et al., 2011). Unreliable or incomplete datasets are unable to correctly identify seasonal or short-range temporal patterns (e.g., Gosset et al., 2013; Le et al., 2017).

Other sources of precipitation data such as those from satellite remote sensing are now available, but they come with their own errors, including random and systematic (see Beck et al., 2021; Beck et al., 2017a; Beck et al., 2017b; Fortin et al., 2015; Sun et al., 2018). Another freely available source of precipitation data are meteorological reanalysis products which are becoming increasingly promising due to upgrades in their spatial resolution and improved representation of atmospheric processes in global models (Hersbach, 2018). Reanalysis data combine a wide range of remotely sensed observations with a dynamical–physical coupled numerical model to produce the best estimate of the state of the atmosphere. Reanalysis is not reliant on the density of surface observational networks and can give surface variables in locations with little to no surface coverage. As a result, they can generate several variables both at the land surface and on vertical atmospheric levels, and hence have been applied in several studies both for climatological and hydrological purposes across the world (e.g., Beck et al., 2017a; Chen et al., 2018; Emerton et al., 2017; Essou et al., 2017). Several different reanalysis products exist but they are known to vary in quality with recurrent upgrades. It is important to evaluate them carefully both to inform the users and the developers of the datasets. The developers of these products can only work on improving their updates when there is a complete feedback loop between applications and developments Therefore, ground validation of reanalysis precipitation is very important but very challenging, particularly where the rain gauge networks are sparse.

Several studies attempt to quantify and account for the sampling errors comparing reanalysis data with observations in different parts of the world (e.g., Guo et al., 2018; Tang et al., 2020; Xu et al., 2020; Zaitchik et al., 2011), at a global scale (Beck et al., 2017a; 2017b), at regional or basin scale (e.g., Acharya et al., 2019; Nkiaka et al., 2017; Tarek et al., 2020) and at a national scale (e.g. Arshad et al., 2021; Gleixner et al., 2020; Koukoula et al., 2020; Lakew et al., 2020; Shayeghi et al., 2020; Tesfaye et al., 2017). However, the findings of these studies were mixed. Differences in approaches, regions, and time scales resulted in inconsistency in product performance, implying that site-specific performance evaluation may be required. Existing studies also aimed at analysing a single product or a few products for short periods of time, thus their estimated errors may not reflect long-term behaviour.

Additionally, the temporal dynamics of rainfall are very important as they play an important role in the total accumulated rainfall on daily and monthly timescales (Ficchì et al., 2016), thus influencing the bimodal seasonality observed over Kenya. Also the highly-variable temporal dynamics are key in explaining the nonlinear nature of infiltration process (Blöschl and Sivapalan, 1995), such as the peak discharge value (Gabellani et al., 2007) and runoff volume (Viglione et al., 2010) in hydrological modelling. Thus, the above highlights the need to consider different temporal scales, when evaluating the reanalysis precipitation relative to observation.

In Kenya, there were 20 major floods from 1964 to 2020 which were driven by precipitation falling in the seasonal rains. More than 160,000 people were displaced countrywide by floods in October 2019 (ReliefWeb, 2019a; 2019b; Opere, 2013). Annual average economic loss from flooding is estimated to be 5.5% of gross domestic product (Njogu, 2021). Thus, understanding the best representation of precipitation in flood models which can be used for forecasting or risk analysis is of great societal importance. Kenya has a widely varying physical geography resulting in great variability of river catchment characteristics across the country. Thus, it is essential not only to understand the representation of precipitation at a country scale, but also on a catchment-by-catchment basis (Golian et al., 2021; Meresa et al., 2021). Previous evaluation of reanalysis products in capturing Kenyan rainfall show varied levels of agreement in spatio-temporal variability relative to observations (e.g., Alemayehu et al., 2018; Dile and Srinivasan, 2014; Gleixner et al., 2020; Khan et al., 2011). Moreover, studies employing hydrological modelling generally used discharge observations from a small number of catchments (e.g., Alemayehu et al., 2018; Bitew et al., 2012; Langat et al., 2017; Le et al., 2017; Worqlul et al., 2017) and did not quantify uncertainties associated with each reanalysis (e.g., Alemayehu et al., 2018), leading to combined rainfall and model uncertainty that is not easily interpreted. Hence, there is a notable gap in the literature associated with evaluating the accuracy of multiple reanalysis products across different catchments, accounting for both model and input errors, especially in data-scarce regions like in Kenya, and this gap was an important motivation for the present study. This paper evaluates four reanalysis precipitation products with respect to observations and assesses their suitability for use in hydrological modelling in 19 Kenyan catchments. We assess their performance in reproducing the most important features of rainfall events and regimes, and in simulating catchment streamflow through answering the following research questions:

- How well do the precipitation datasets compare in terms of temporal dynamics at the basin scale? Which product is the most accurate compared to observations?
- How well do precipitation datasets compare in terms of spatial patterns? Which product shows consistency in spatial heterogeneity compared to observations?
- How does the general hydrological model performance vary with different datasets?
- How does the sensitivity of a rainfall runoff model (GR4J) vary with alternative rainfall forcing?

We consider both model performance and parameter uncertainty and compute a Model Suitability Index (MSI) by coupling the results of model performance statistics and Global Sensitivity Analysis. We compare four reanalysis datasets using the GR4J model across 19 Kenyan catchments with varied climate and morphological characteristics, to investigate which input data is suitable or require caution, when used in the place of observation dataset in different regions. This work is a steppingstone and an essential guide for hydrological applications of global reanalysis datasets because it compares several reanalysis products to observations on daily, monthly, and seasonal scales, and unveils the propagation of uncertainty from different reanalysis when used as model inputs. All the

above reviewed studies looked at the performance of at most one reanalysis dataset in simulating streamflow and only over one catchment; but none looked at such a country-scale performance. To our knowledge, this is the very first evaluation of the different reanalysis products over Kenya for simulating streamflow coupled with sensitivity analysis.

2. Study area and catchment characteristics

The study is undertaken in 19 Kenyan catchments (Fig. 1) with varying characteristics (Table 1). These were selected due to the frequency and magnitude of the impacts of floods, as well as the availability of river flow observations (Table 1). Kenya mainly experiences a bimodal rainfall pattern, occurring in the seasons of March - April - May (MAM) and October -November - December (OND) (Ayugi et al., 2016; Yang et al., 2015), which are commonly known as the 'long' and 'short' rains respectively. The rainfall seasonality and the migration of the precipitation zone is mainly influenced by the north-south movement of the inter-tropical convergence zone (ITCZ) (Black et al., 2003; Ongoma et al., 2015). The rainfall season migrates northward at a slower rate than it migrates southward, hence the two different names – 'long rains' and 'short rains' respectively (Nyenzi, 1988). The rainfall exhibits high spatiotemporal and interannual variability (Ongoma and Chen, 2017) and is strongly influenced by perturbations in global Sea Surface Temperature (SSTs) especially in the Pacific and Indian Oceans with the El-Niño Southern Oscillation (ENSO) (Black et al., 2003; Ogallo, 1993) and the
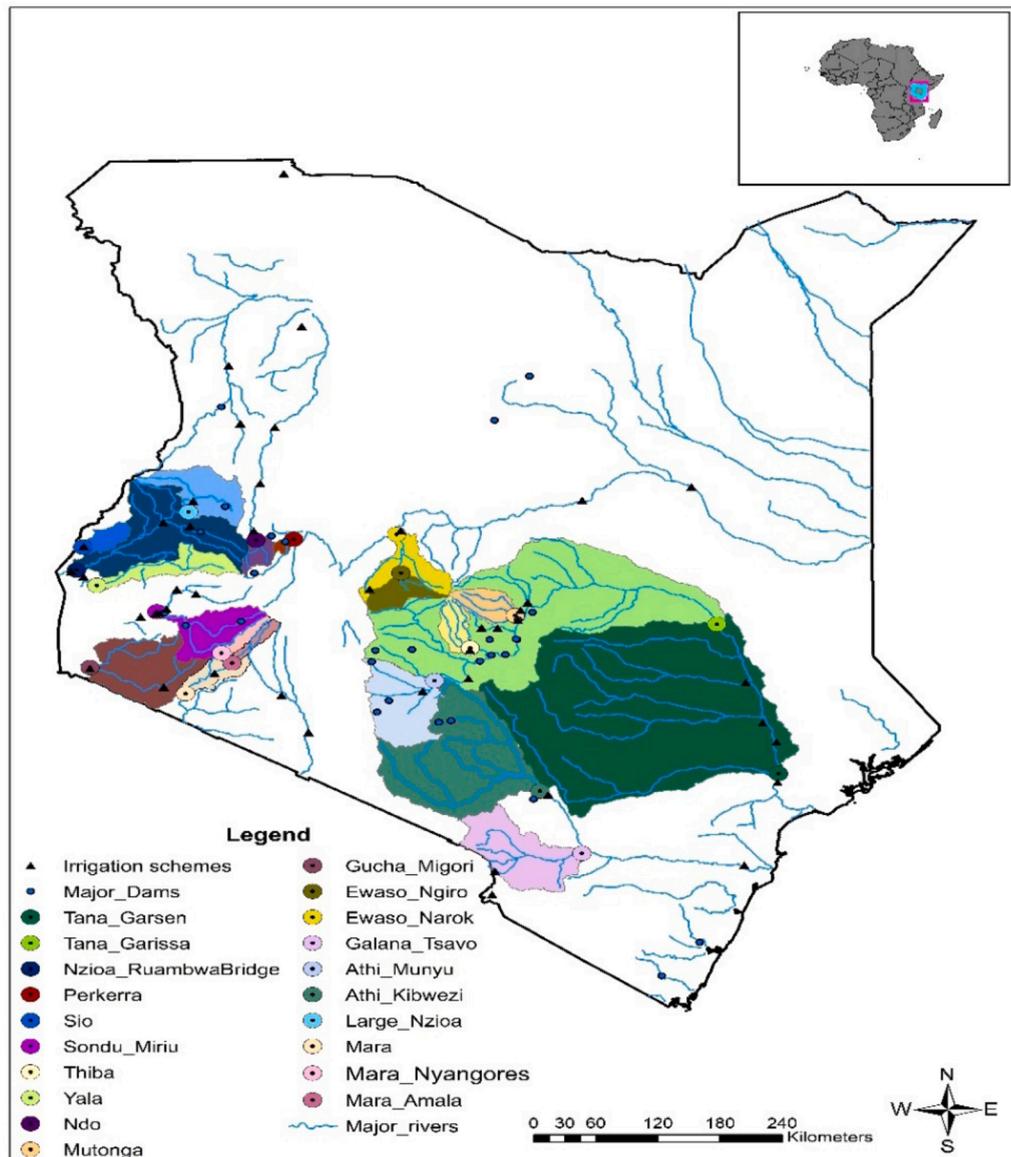
Fig. 1. Study catchments across Kenya, with locations of the outlet river gauges (show by cirloured circled dots) used in this study and the main irrigation schemes (black trianges) and major dams (blue circles) provided by Kenya Water Authrority (WRA).

Table 1
Summary of the catchments considered, their characteristics, and the main human influences, including number of dams and water abstraction activities (Source: WRA-K).

River Name	Catchment Outlet point	Station ID	Lon	Lat	Drainage Area (km²)	Mean Elevation (m.a.s.l)	Mean Annual Rainfall (mm)	Annual Discharge (m³s⁻¹)	Catchment Characteristics	Human Influence		First & Last year of record	Record length (years)	Amount missing (%)
										Dams	Irrigation schemes			
Tana	Tana Garsen	4G02	40.11	-2.28	80 760	720	672	135.8	Semi-arid plains	9	11	1981–2016	36	58.2
Tana	Tana Garissa	4G01	39.7	-0.45	32,695	870	868	169.3	Highlands on the upstream & semi-arid plains in the lowland	8	7	1981–2018	38	14.2
Nzioa	RuambwaBridge	1EF01	34.09	0.12	12,643	1740	1488	151.2	Dense forest cover (highlands) & low trees & bushes (lower reaches)	2	4	1981–2018	38	13.6
Galana	Galana Tsavo	3G02	38.47	-2.99	6560	930	628	3.3	Semi-arid savannah plains	3	1	1981–2015	35	59.6
Gucha	Gucha Migori	1KB05	34.21	-0.95	6 310	1650	1435	45.0	Eastern lowlands with dense vegetation cover	0	2	1981–2015	35	47.8
Athi	Athi Munyu	3DA02	37.19	-1.09	5689	1730	822	18.8	Highlands and forest cover	3	1	1981–2017	37	21.6
Nzioa	Large Nzioa	1BD02	35.06	0.76	3878	1720	1267	15.3	Dense forest cover	1	1	1981–2011	31	28.8
Sondu	Sondu Miriu	1JG04	34.80	-0.33	3444	2017	1614	53.9	Low lying plains (western) & highland (Eastern)	2	2	1981–2018	38	64.4
Mara	Mara	1LA04	35.04	-1.23	2977	2100	1262	11.8	Low lying shrubs, semi-arid	0	1	1981–2015	35	77.7
Yala	Yala	1FG02	34.27	0.04	2700		1696	40.8	Swampy	0	0	1981–2019	39	59.6
Ewaso	Ewaso Narok	5AC10	36.73	0.43	2597	1600	880	5.3	Low lying shrubs & mainly semi-arid	0	2	1981–2018	38	26.5
Tana	Mutonga	4EA07	37.89	-0.38	1867	1830	1427	35.5	Highlands and forest cover	0	1	1981–2016	36	44.2
Ewaso	Ewaso Ngiro	5BC04	36.91	0.09	1837	1700	972	20.6	Low lying shrubs & mainly semi-arid	0	0	1981–2019	39	35.0
Sio	Sio	2EE07A	34.14	0.39	1011	1650	1822	15.5	Low trees & bushes & swampy in lower reaches	0	1	1981–2018	38	18.1
Turkwel	Ndo	2C07	35.65	0.45	897	1133	1371	9.1	Extensive palaeo-floodplain & arid conditions	0	1	1981–1993	13	47.2
Mara	Amala	1LB02	35.44	-0.89	695	2100	1377	6.8	Low lying shrubs, semi-arid	0	0	1981–2017	37	25.6
Mara	Nyangores	1LA03	35.35	-0.79	692	2008	1262	11.8	Semi-arid savannah plains, low lying shrubs, semi-arid	0	0	1981–2017	37	15.5
Turkwel	Perkerra	2EE07A	35.97	0.46	371	1023	832	5.7	Extensive palaeo-floodplain and arid conditions	1	1	1985–2005	21	50.1

Table 2

Overview of the global reanalysis and the blended (Satellite and observation) Chirps precipitation dataset(s) used in the study.

Short Name	Full Name and details	Data sources (*)	Spatial resolution	Spatial coverage	Temporal coverage	Temporal resolution	Reference
CHIRPS V2.0	Climate Hazards group Infrared Precipitation (CHIRP) V2.0 (http://chg.ucsb.edu/data/chirps/)	G, S, R *	0.05^o	Land, < 50	Daily	1981–NRT*	(Funk et al., 2015)
ERA5	European Centre for Medium-Range Weather Forecasts Reanalysis (https://www.ecmwf.int/en/research/climate-reanalysis/era-5)	R *	~ 0.31^o	Global	Hourly	1979 – NRT*	(Hersbach, 2018)
ERA-Interim	European Centre for Medium-Range Weather Forecasts ReAnalysis Interim (https://www.ecmwf.int/en/research/climate-reanalysis/era-interim)	R *	~ 0.75^o	Global	3-Hourly	1979–2019	(Dee et al., 2011)
JRA-55	Japanese 55-year Reanalysis (JRA-55) (https://jra.kishou.go.jp/JRA-55; https://data.diasjp.net/dl/storages/filelist/dataset:204)	R *	~ 0.56^o	Global	3-Hourly	1951-NRT*	(Kobayashi et al., 2015)
NCEP-CFSR	National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR; http://cfs.ncep.noaa.gov/cfsr/)	R *	~ 0.38^o	Global	Hourly	1979– 2010	(Saha et al., 2010)

$^{(*)}$ NRT= Near Real Time with a delay of several days, G = Gauge, S = Satellite, R = Reanalysis

Indian Ocean Dipole (IOD) (Blau et al., 2020; Owiti et al., 2008) being the most important modes. Other systems that influence rainfall variability include the high pressure systems (e.g. the Mascarene and the Arabian) (Ogwang et al., 2015), the Quasi-Biennial Oscillation (QBO) (Collier et al., 2016; Indeje and Semazzi, 2000), the Madden-Julian Oscillation (MJO) (Kilavi et al., 2018), Tropical cyclones (Finney et al., 2020; Wainwright et al., 2021) and jet streams, eg., the Turkana jet (Hartman, 2018; Kinuthia, 1992). The country has complex topography with the lowest altitudes along the coastline and Lake Victoria basin which are particularly prone to floods, while in the highlands, frequent thunderstorms and lightning threaten life.

3. Data and methodology

3.1. Datasets

3.1.1. Reanalysis and observational data

Four reanalysis products, namely ERA5, ERA-Interim (hereafter ERAI), Climate Forecast System Reanalysis (CFSR), and the Japanese 55-year Reanalysis (JRA55), and a gridded observational dataset, the Climate Hazards group Infrared Precipitation with Station (CHIRPS), were used in this study (see Table 2). We used the daily precipitation, maximum and minimum temperature variables from the reanalysis products for the study.

ERA5 is the latest global atmospheric reanalysis product from European Centre for Medium-Range Weather Forecasts (ECMWF) which spans the modern observing period from 1950 onward (Hersbach, 2018). In this study, 3-hourly ERA5 was obtained from ECMWF on a fixed grid of $0.31° \times 0.31°$. ERAI is the previous global reanalysis product created by ECMWF (Dee et al., 2011). Daily ERAI was obtained from ECMWF on a fixed grid of $0.75° \times 0.75°$. JRA55 is a global reanalysis dataset constructed by the Japan Meteorological Agency (JMA) (Kobayashi et al., 2015). Daily JRA55 was obtained from National Center for Atmospheric Research (NCAR) climate data guide at a fixed gird of $0.56° \times 0.56°$. CFSR is a global reanalysis dataset of atmosphere fields produced by the National Centers for Environmental Prediction and for Atmospheric Research (NCEP/NCAR) (Saha et al., 2010). The CHIRPS dataset was used as a benchmark observation dataset since it has been used in several studies showing good results compared to observations over eastern Africa (Dinku et al., 2018). CHIRPS is a quasi-global, high resolution, daily, pentad, and monthly precipitation dataset (Funk et al., 2015). Based on infrared Cold Cloud Duration (CCD) data, CHIRPS has a long enough history of precipitation data. The algorithm is based on (i) a 5 km climatology that uses satellite data to represent sparsely gauged locations, ii) includes daily, pentadal, and monthly 5 km CCD-based precipitation estimates from 1981 to the present, iii) combines station data to generate tentative information product with a latency of about 2 days and a final product with an average latency of about 3 weeks, and iv) interpolation weights are assigned based on a novel blending method which uses the spatial correlation structure of CCD estimates. This makes it comparatively an alternative in data scarce regions. We opted for the gridded observations as the daily observed gauge datasets were not available for the catchments of study and are known to be very sparse and present large data gaps (Dinku et. al, 2019; 2018; Le et. al, 2017).

3.1.2. Observed river discharge and potential evapotranspiration

River discharge datasets at daily time step for the period 1981–2016 were provided by the Kenya Water Resource Authority (WRA) for the selected catchments across the country, summarized in (Table 2). The potential evapotranspiration (PET) required for the catchment modelling was estimated from the average daily temperature of the four reanalysis products and CHIRTS_daily data from the Climate Hazard Centre (CHC). As temperature readings were the readily available meteorological data that relates to PET, for this study, temperature-based methods were used to estimate the PET (Hargreaves and Samani, 1985). For this study, the Hamon method

(Hamon, 1960) was used to estimate PET daily averages for different datasets.

3.2. Modelling experiment methodology

To obtain the monthly and annual totals for observations and reanalysis datasets, the daily values were accumulated. The seasonal total precipitation was calculated by summing monthly precipitation for three seasons: (i) the March-April-May, hereafter referred to as MAM, (ii) the June-July-August, hereafter JJA, and (iii) the October-November-December, hereafter OND. All datasets were converted to the same units for consistency (e.g., JRA55 and CFSR were converted from $kg/m^2/s$ to mm/d). ERA5, ERAI, JRA-55 and CFSR were regridded by first-order conservative interpolations to a horizontal grid of 0.5^o x 0.5^o (Schulzweida, 2019).

We first qualitatively evaluate the performance statistics of the reanalysis datasets in terms of temporal dynamics and biases with respect to precipitation observations (CHIRPS), considering the following metrics: Pearson Linear Correlation Coefficient (CC), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Error (ME), long-term relative bias (BIAS) and annual number of dry days calculated on monthly, annual, and seasonal scales. We produce spatial maps for the standardized precipitation anomalies, bias and annual number of dry days, to assess their consistency compared to observations, and tabulate the other statistics to show the aggregate performance across the different datasets.

Second, we calibrate the GR4J (Perrin et al., 2003) rainfall runoff model. In the study, we used five different inputs sources (for both precipitation and PET) into GR4J model from CHIRPS, ERA5, ERAI, CFSR and JRA55. We calibrate the model with each of the input source at a time and compute the KGE score and compare how this varies across the four different datasets relative to observations. The GR4J model is a simple daily lumped rainfall-runoff model belonging to the family of soil moisture accounting models. There are four main parameters (Fig. 2) to be calibrated in GR4J model, namely: (1) the maximum capacity of the production store (X1, mm), (2) the groundwater exchange coefficient (X2, mm), (3) the maximum capacity of the non-linear routing store (X3, mm), and (4) the time base of the unit hydrograph (X4, days). There are also a few fixed parameters, whose values were set by Perrin et al. (2003). All four free parameters are real numbers, X1 and X3 are positive, X4 is greater than 0.5 and X2 can be either positive, zero or negative. The typical inputs of GR4J are the areal precipitation depth (P, mm) and the potential evapotranspiration (PE, mm) estimate over the catchment. Most optimization algorithms used to calibrate the model parameters require knowledge of an initial parameter set. Given the small number of model parameters, simple optimization algorithms are generally capable of identifying parameter values yielding satisfactory results. The choice of an objective function depends on the objectives of model user. The choice and use of the GR4J model is mainly due to its simple and relatively quick to calibrate structure, ensuring high-levels of performance and robustness (Ficchì et al.,
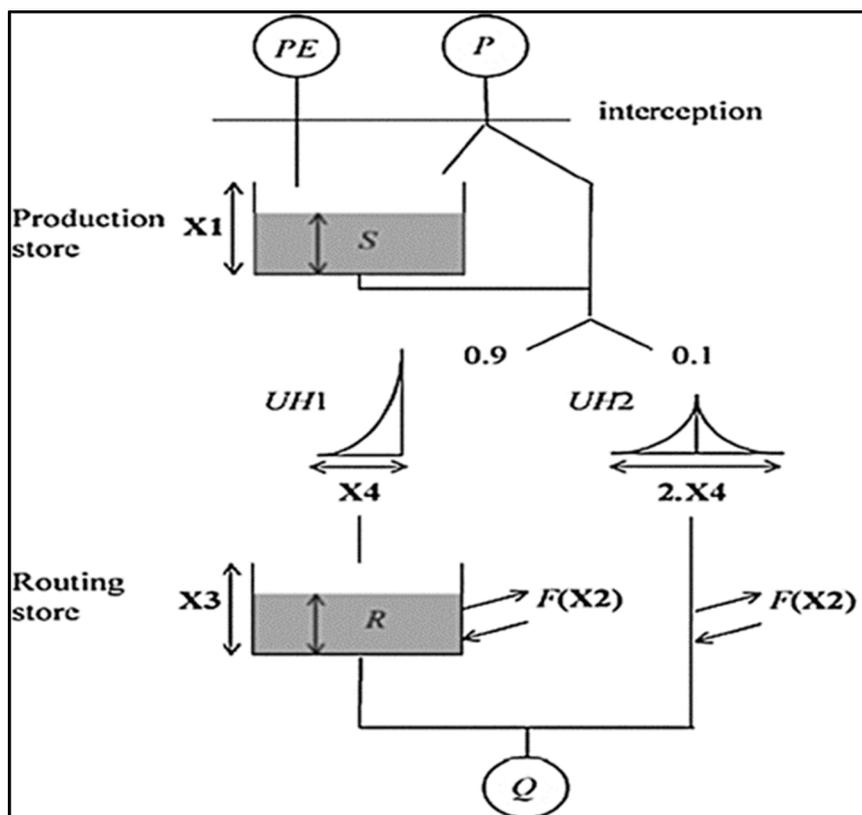
Fig. 2. The GR4J rainfall-runoff model (Source (Perrin et al., 2003)). P is the rainfall depth and E is the potential evapotranspiration (PE) averaged over the basin at a daily time step. X1, X2, X3 and X4 are the model parameters.

2019; Mostafaie et al., 2018; Oudin et al., 2004; Van Esse et al., 2013; Pushpalatha et al., 2011).

The four free parameters of the GR4J model were calibrated using the default optimisation algorithm provided in the airGR package (Coron et al., 2019; Delaigue et al., 2019). This simple optimization algorithm, mainly based on a local optimisation, proved to be equally efficient to locate a robust optimum compared to more complex global search algorithms (Coron et al., 2019) and proved effective in terms of the number of model runs required for convergence (Mathevet et al., 2006). The Michel method (Michel, 1983) is based on two steps:

I. A systematic inspection of the global parameter space is performed to determine the most likely zone of convergence. In our study, this is done by direct grid-screening.
II. A steepest descent local search procedure is carried out to find an estimate of the optimum parameter set starting from the best parameter set from step 1.

The four free model parameters were calibrated by applying the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009) as the objective function and the daily observed river discharge data of the selected catchments as reference. We use different inputs (precipitation datasets) from CHIRPS, ERA5, ERAI, CFSR and JRA55 to calibrate the GR4J model. The KGE was used also for evaluating the performance of the GR4J model when forced with different reanalysis data. The KGE objective function represents a weighting of three components that correspond to bias, correlation, and variability, ensuring that KGE is sensitive to errors in the overall distribution of streamflow (Adeyeri et al., 2020; Kling et al., 2012). We therefore calculated the hydrological model performance statistics for the calibration and validation periods and compared across the different reanalysis datasets to investigate the overall suitability of the different reanalysis as input data to simulate river flows. We adopted a threshold of model performance in the range $-0.41 < \text{KGE} \leq 1$ as reasonable, following Knoben et al., (2019) being -0.41 the KGE value corresponding to a mean flow benchmark.

A split-sample validation technique (Klemeš, 1986) was used to test model performance beyond the calibration period. For this study, 36 years (1981–2016) of streamflow data for each catchment were available, so we split into two equal 18-year Split-Sample Testing (SST) periods hereafter referred to as SST1 and SST2.

Third, we perform Sensitivity Analysis by applying the global Sobol sensitivity method for the GR4J model parameters using the KGE as our target function and the daily observed data of the 19 catchments as reference. We adopt the Sobol method because it estimates the relative contribution of individual model parameters and their interactions through the decomposition of model output variance (Nossent et al., 2011). A sensitivity analysis allows a reduction of the number of parameters incorporated in the optimization by determining the most influential parameters of a model and their identifiability (Saltelli et al., 2000). As no prior information is available on the parameters, the input parameter values for the Sensitivity Analysis are sampled from a uniform distribution (Nossent et al., 2011). The different parameter ranges are scaled between 0 and 1 with a linear transformation. Then, we obtain one value of the Sensitivity Indices (SI) per parameter, and we investigate the relative role of each parameter in explaining the output variance and assess possible over-parameterization issues by counting the number of sensitive parameters. The value of the objective function for the calibration of parameters can be used as the model performance statistics for sensitivity analysis, as such we adopted the KGE.

Last, we assess the overall suitability of the rainfall–runoff model when forced with different meteorological inputs by calculating the Model Suitability Index (MSI). We compare the performance of the four reanalysis datasets across the 19 catchments and investigate which of the input dataset is suitable and which require caution, because of low model performance and possible parameter identifiability or over-parameterization problems. The well-known problem of over-parameterisation due to insensitive parameters in models with large number of parameters (van Griensven et al., 2006) makes sensitivity and performance statistics important. This may result in uncertain model simulations arising from equifinality in model calibration but yielding unequifinal model simulations in
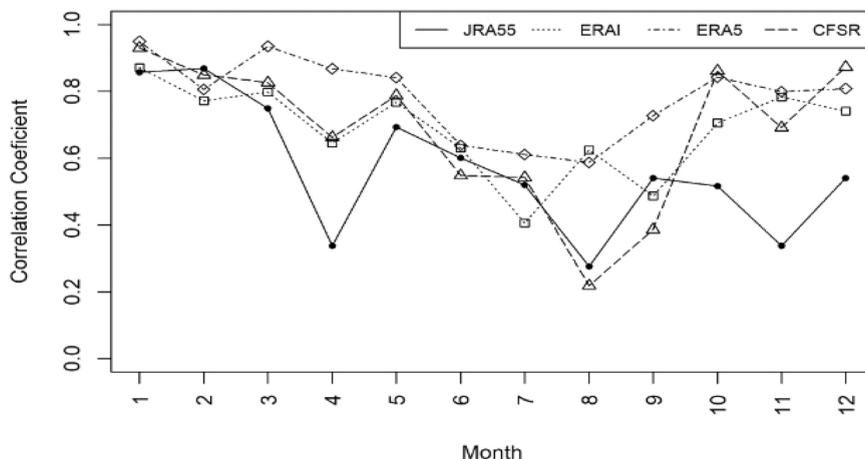
Fig. 3. Line graph of correlation coefficients (CC) between monthly observations and ERA5, ERAI, JRA55 and CFSR precipitation for the period 1981 – 2016 on average across the 19 study catchments.

validation (Beven, 2012). This mostly arises due to application of calibrated multiple optimal parameters sets with significantly variable parameter values (Shin et al., 2015; Shin and Kim, 2017). Therefore, in most cases the problem arising from prediction uncertainty may pose problems to modellers when it comes to decision making. By applying the quantitative method of Sobol's SA, this enabled us to couple the results with the performance statistics. The MSI aggregates both sensitivity indices & performance statistics (Shin and Kim, 2017), providing a clear index to judge the relative global performance of the reanalysis products with respect to observations. The computed MSI can be used in comparison studies with any catchment data. If all the model parameters are sensitive, this would yield a MSI of 1 and a perfectly matched hydrograph between the simulations and observations.

We adapt Shin' and Kim (2017) Model Suitability Index (MSI), which is a combined measure of performance statistics and Sensitivity Analysis results. The MSI can be expressed as:

$$MSI = 0.5 \times \left(\frac{1}{n} \sum_{i=1}^{n} SR_i \right) + 0.5 \times \left(\frac{1}{m} \sum_{j=1}^{m} PS_j \right) \tag{1}$$

Where the SR is the sensitivity ratio (i.e., the ratio of the number of sensitive parameters out of the total number of model parameters) ranging from [0, 1] and PS is the performance statistics, n is the number of years over which the sensitivity analysis is run, and m is the number of split sample periods in model calibration. It is necessary to set a sensitivity threshold to ascertain the sensitive parameters, hence we adopted a minimum value of 0.2 for the TSI of a sensitive parameter. This value has been suggested and used in some past studies (e.g (van Werkhoven, 2009; Van Werkhoven et al., 2009 & Shin et al., 2013). It is worth noting that this is an arbitrary value, thus we acknowledge the need to practice caution when the parameters' TSI values are nearing the threshold. PS is computed by obtaining the average value of all the periods considered (i.e., two split sample periods). To calculate the average PS, we considered the calibration and validation performance statistics (KGE). As both measures are equally important, we gave the same averaged weight to PS and SR in calculating the MSI.

4. Results and discussion

4.1. Results

4.1.1. Overall performance evaluation using observations

The performance of ERA5, ERAI, JRA55 and CFSR on monthly, seasonal, and annual scales is presented in this section. We used the monthly scale as a base time scale and calculated CC, RMSE, MAE and ME for all the four reanalysis products.

4.1.1.1. Performance on monthly scale. ERA5, ERAI, JRA55 and CFSR were first evaluated on a monthly timescale with respect to observations at the country level. All the datasets passed the significance test of the correlation coefficient at the 99% confidence level and to eliminate the influence of the seasonal cycle on the values, each Correlation Coefficient was calculated per month as shown in Fig. 3. ERA5 shows the highest average correlation coefficient value of 0.71 on monthly timescale compared to observations and is consistently higher across all months (Fig. 3, and Table 3) than the other reanalysis products. ERAI and CFSR have good average correlation but show larger drops in some months (especially in the drier month of August). JRA55 obtained a poor correlation coefficient of 0.46 on average. In general, ERA5, ERAI and CFSR show higher correlations to observations in rainy months (March-April-May and October- November-December) and lower in the dry months (June-July- August), whereas JRA55 shows worst correlations during both rainy seasons.

The average twelve months evaluation indices for each of the reanalysis product is shown in Table 3. Overall, ERA5, ERAI and CFSR show a similar good ability to simulate the precipitation for all the indices under consideration. ERA5 has a better CC, BIAS and RMSE whereas JRA55 has the lowest CC and the largest BIAS and RMSE suggesting that JRA55 is worst performing reanalysis dataset over the Kenyan catchments.

4.1.1.2. Performance on seasonal and annual timescales. The overall performance of the four reanalyses (ERA5, ERAI, CFSR and JRA55) were evaluated on seasonal and annual timescales to explain the propagation of errors at these timescales. The results of the different performance indices are shown in Table 4.

The overall correlation coefficients on seasonal and annual timescale are shown in Fig. 4. Higher CC across all the datasets were obtained in the wet seasons of MAM and OND, whereas lower CC were obtained in the dry season of JJA, with the performance index

Table 3
Average CC, BIAS, RMSE, MAE and ME between the four reanalysis precipitation datasets and observations on monthly timescale for the period 1981 – 2016 over all the study catchments.

Index	ERA5	ERA-Interim	JRA55	CFSR
CC	0.71	0.63	0.46	0.68
BIAS (%)	49.72	-26.97	146.66	-76.68
RMSE (mm)	31.59	43.73	115.12	67.36
MAE (mm)	25.11	37.02	79.92	59.95
ME (mm)	1.64	-30.00	71.52	-59.91

Table 4
CC, BIAS, RMSE, MAE and ME between the reanalyses and observation precipitation data at a seasonal and annual timescale averaged over the 19 study catchments in Kenya.

Season	Dataset	CC	BIAS (%)	RMSE (mm)	MAE (mm)	ME (mm)
MAM	JRA55	0.34	155	965.10	898.12	898.12
	ERAI	0.59	-28.1	193.87	168.09	-160.36
	ERA5	0.88	47.6	322.44	282.42	275.60
	CFSR	0.78	-84.6	481.03	475.14	-475.14
JJA	JRA55	0.52	67.4	361.72	290.77	267.76
	ERAI	0.24	-39.5	182.05	157.12	-154.93
	ERA5	0.25	-1.4	85.66	70.90	-5.446
	CFSR	0.22	-84.6	332.89	327.95	-327.95
OND	JRA55	0.44	271.7	1036.4	983.25	983.25
	ERAI	0.81	-27.9	123.34	111.09	-100.56
	ERA5	0.52	96.6	397.00	349.70	349.70
	CFSR	0.84	-84.3	307.76	296.47	-296.47
ANNUAL	JRA55	0.25	171	2902.21	2760.63	2760.63
	ERAI	0.46	-26.3	462.513	421.14	-421.14
	ERA5	0.52	44.7	801.75	728.57	720.70
	CFSR	0.60	-85.2	1354.36	1349.21	-1349.21

Fig. 4. Boxplots of the seasonal (MAM, JJA and OND) and annual Correlation Coefficients (CC) for four reanalysis CFSR (pink), ERA5(green), ERAI (blue) and JRA55 (purple) across the 19 catchments. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.
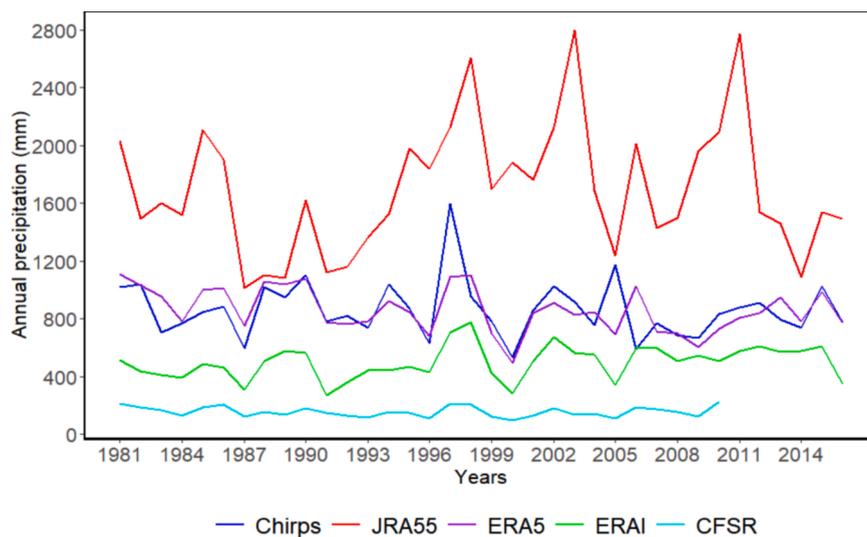
Fig. 5. Areal average annual precipitation from the observations and the reanalysis datasets, averaged across the 19 study catchments.
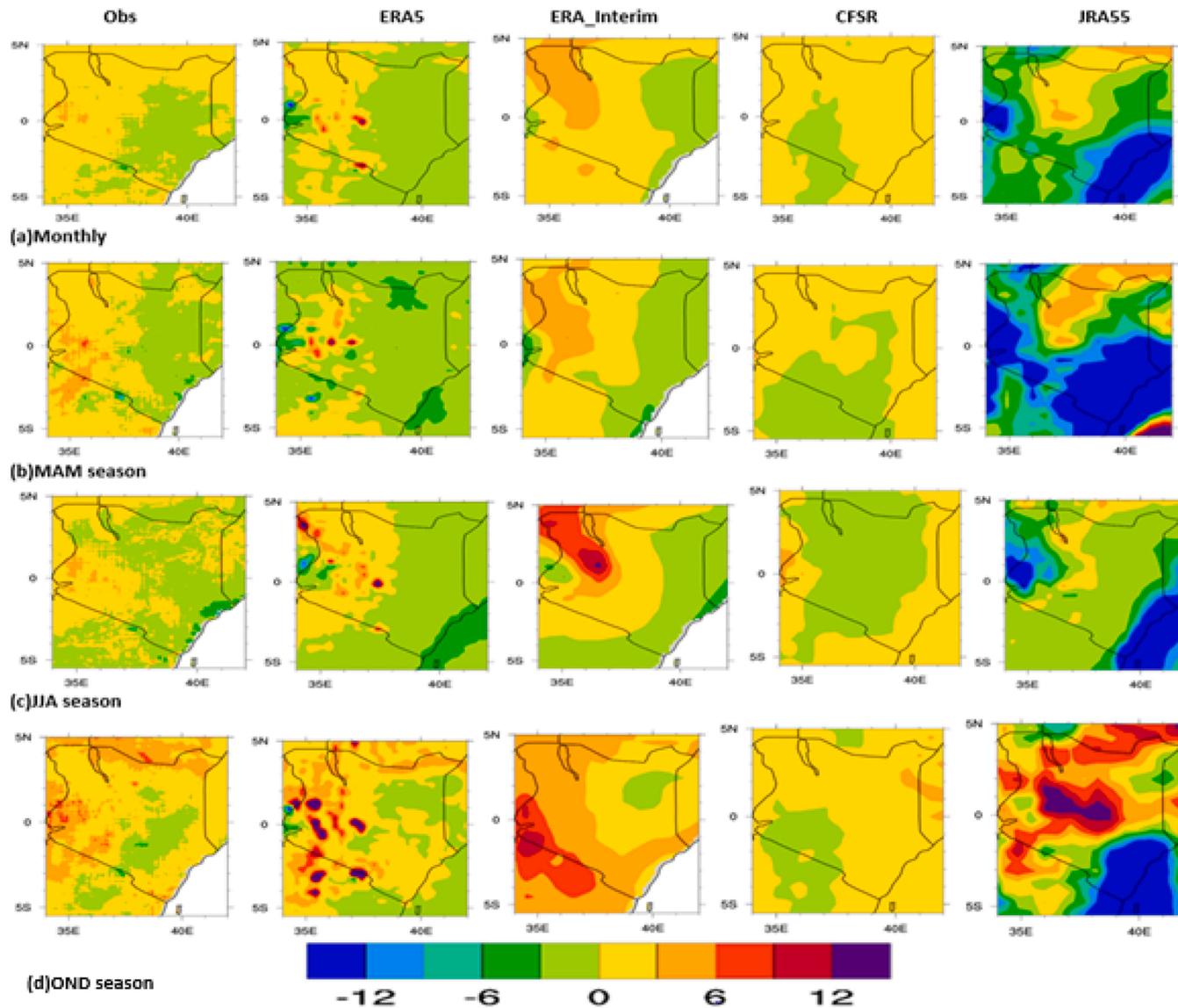
Fig. 6. Mean monthly and seasonal standardized precipitation anomalies in the four reanalysis products for 1981–2016 period: (a) Monthly anomalies, (b) MAM season, (c) JJA and (d) OND anomaly index in seasonal precipitation, for ERA5 (2nd column), ERA-I (3rd column), CFSR (4th column) and JRA-55 (5th column). Column 1 shows the observations (CHIRPS).
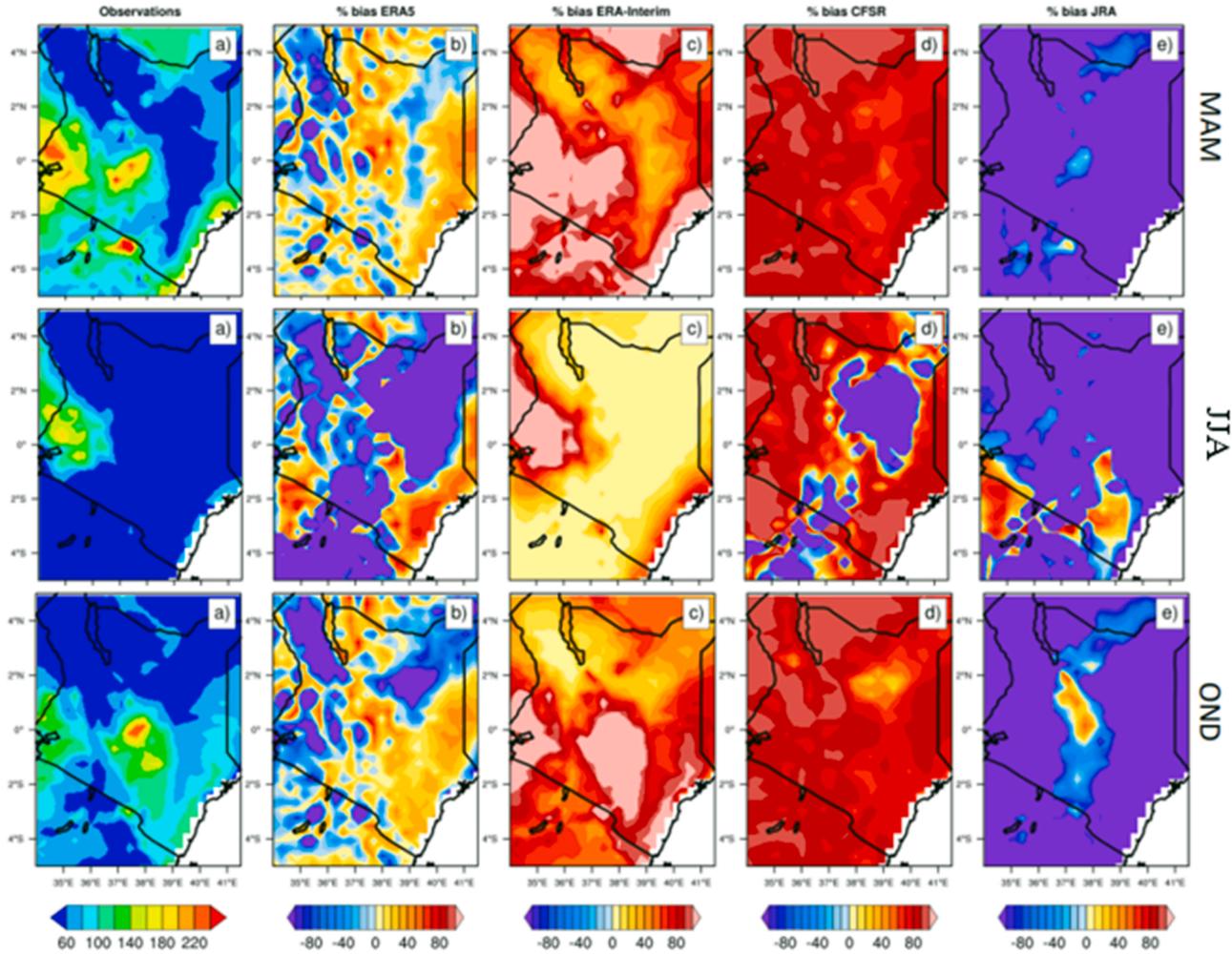
Fig. 7. Seasonal observed precipitation (mm) and mean bias (%) of the extreme rainy days at 95th percentile in the four reanalysis products for 1981–2016 period: (a) Long-term observed (OBS) seasonal average precipitation (mm) from CHIRPS, (b– e) mean relative bias (%), in seasonal precipitation in ERA5 (b), ERA-Interim (c), CFSR (d), JRA-55 (e), with respect to CHIRPS. MAM season (top), JJA (middle) and OND (bottom panel).
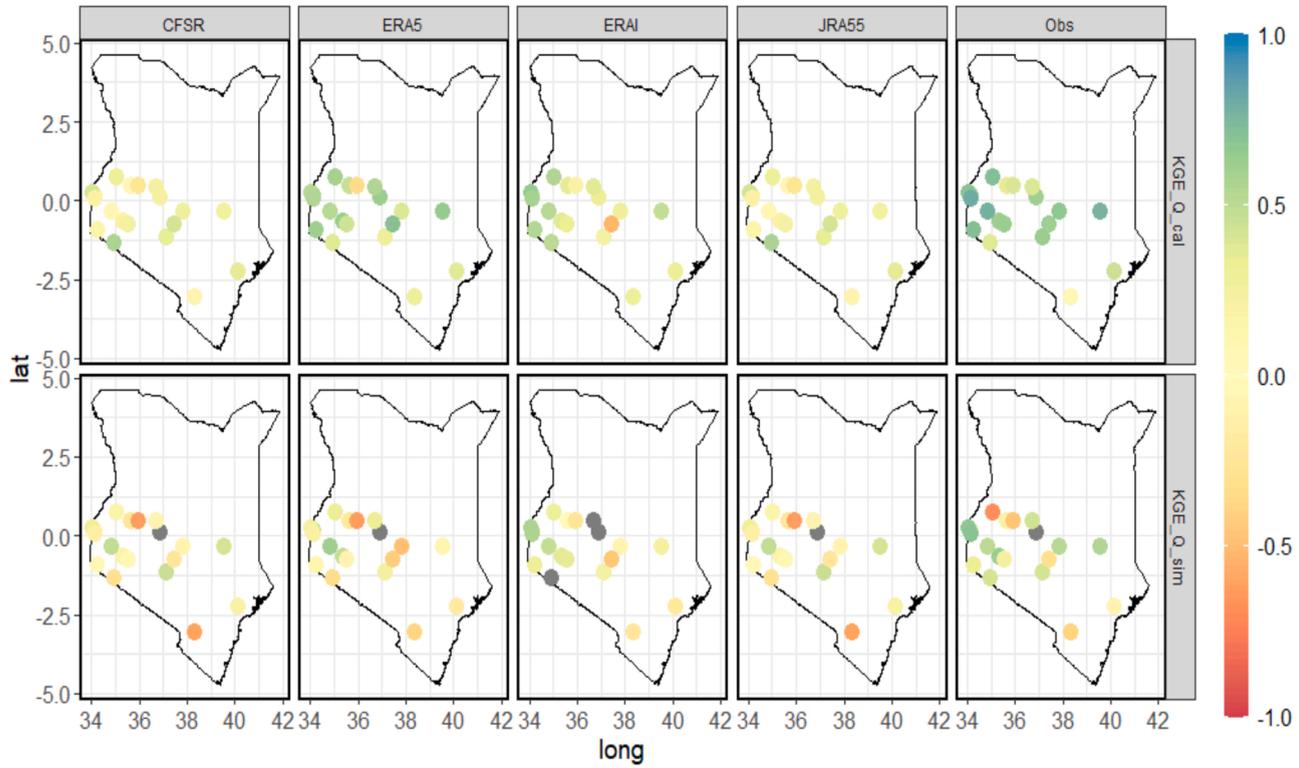
Fig. 8. GR4J model performance (KGE) in calibration (top panel) and validation (bottom panel) across the 19 catchments for different input datasets, Pan.1 (CFSR), Pan.2 (ERA5), Pan.3 (ERAI), Pan. 4 (JRA55), Pan.6 (CHIRPS).

higher in OND than in MAM. ERA5 obtained the highest CC (0.88) in MAM, whereas CFSR the highest (0.84) in OND. JRA55 showed lower CC of 0.34 and 0.44 in the two seasons respectively and a CC of 0.52 in the dry season, depicting a tendency of a wet bias over the dry months. On average, the variability in the CC index across the four datasets was relatively lower in the OND season and higher in the MAM season. The BIAS across the four datasets was lower in the dry season (JJA) and higher in the wet seasons (MAM &OND) with JRA55 showing a higher positive BIAS across all the seasons. There are large values in the RMSE and the MAE across the four datasets in the two wet seasons and this may be linked to the high precipitation concentrations during those seasons across most of the catchments. Generally, it can be noted that JRA55 shows the worst performance in comparison to observation especially in the wet seasons of MAM and OND but obtained relatively better scores in the dry season of JJA. ERA5 shows better agreement with observations across the three seasons thus may be an appropriate option for simulating precipitation over the Kenyan catchments.

On annual timescale, the average annual precipitation of CFSR, ERA5, ERAI and JRA55 was computed and compared with the observation (CHIRPS) (Fig. 5). ERA5, ERAI and CFSR show similar trend compared to observations across all the years, with CFSR and ERAI underestimating the precipitation. JRA55 shows a higher tendency of overestimating the annual precipitation over the study catchments. In terms of the performance indices, CFSR, ERAI and ERA5 showed better CC indices of 0.60, 0.46, 0.52 respectively whereas JRA obtained lower CC of 0.25 (Fig. 4). The variability in the CC was higher in JRA55 (Fig. 4). ERA5 and JRA55 show a positive bias of 45% and 171% respectively, whereas ERAI and CFSR show negative bias of − 26 and − 85%. ERA5 has a lower RMSE and ME whereas JRA55 has the highest. These results show that ERA5 is the best performing reanalysis dataset compared to observations on annual timescales whereas JRA55 is the worst performing.

The mean monthly and seasonal standardized precipitation anomalies in the four-reanalysis precipitation for a base climatological period 1981–2016 is shown in Fig. 6. On monthly timescale, the observations show a positive anomaly over the central highland and the western parts of Kenya (Fig. 6, pan1). The arid and semi-arid parts in the eastern and coastal lowlands show a negative anomaly (dry bias). This pattern is also captured in ERA5, ERAI and JRA55 although JRA55 has too high and widespread negative anomalies compared to the former two. On seasonal timescales, ERA5, ERAI and CFSR show positive anomalies in Western and central highland in the three seasons except for JRA55 which has a stronger negative and positive anomaly in MAM and OND seasons respectively.

An evaluation of the extreme precipitation in the four reanalysis was also performed (Fig. 7). For this case, we focused on the 95th percentile of rainy days for MAM, JJA and OND season during the period 1981–2016. A rainy day represents a day for which the recorded precipitation amount is greater than or equal to 1 mm (Gudoshava et al., 2020). The observed extreme precipitation varied between 60 mm to more than 240 mm across western, central highlands and coastal catchments for the rainy seasons (MAM and OND). The observed extreme precipitation during the dry season JJA varied between 100 mm and 160 mm across the western catchments
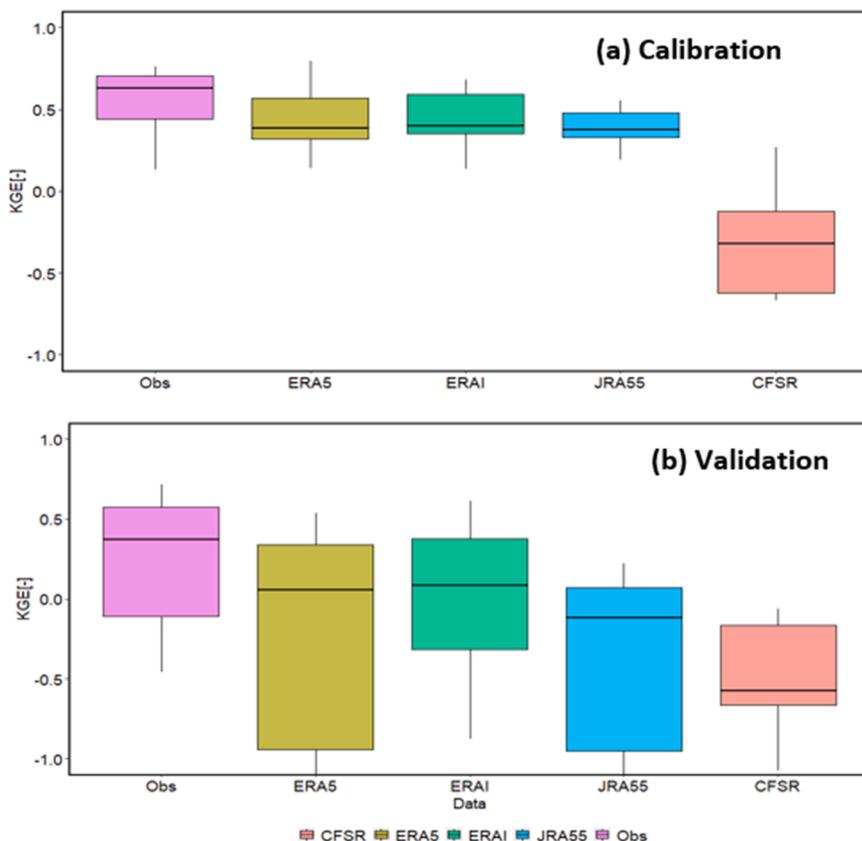
Fig. 9. Boxplots of the overall GR4J model performance (KGE [-]) in (a) calibration and (b) validation mode over the 19 catchments.
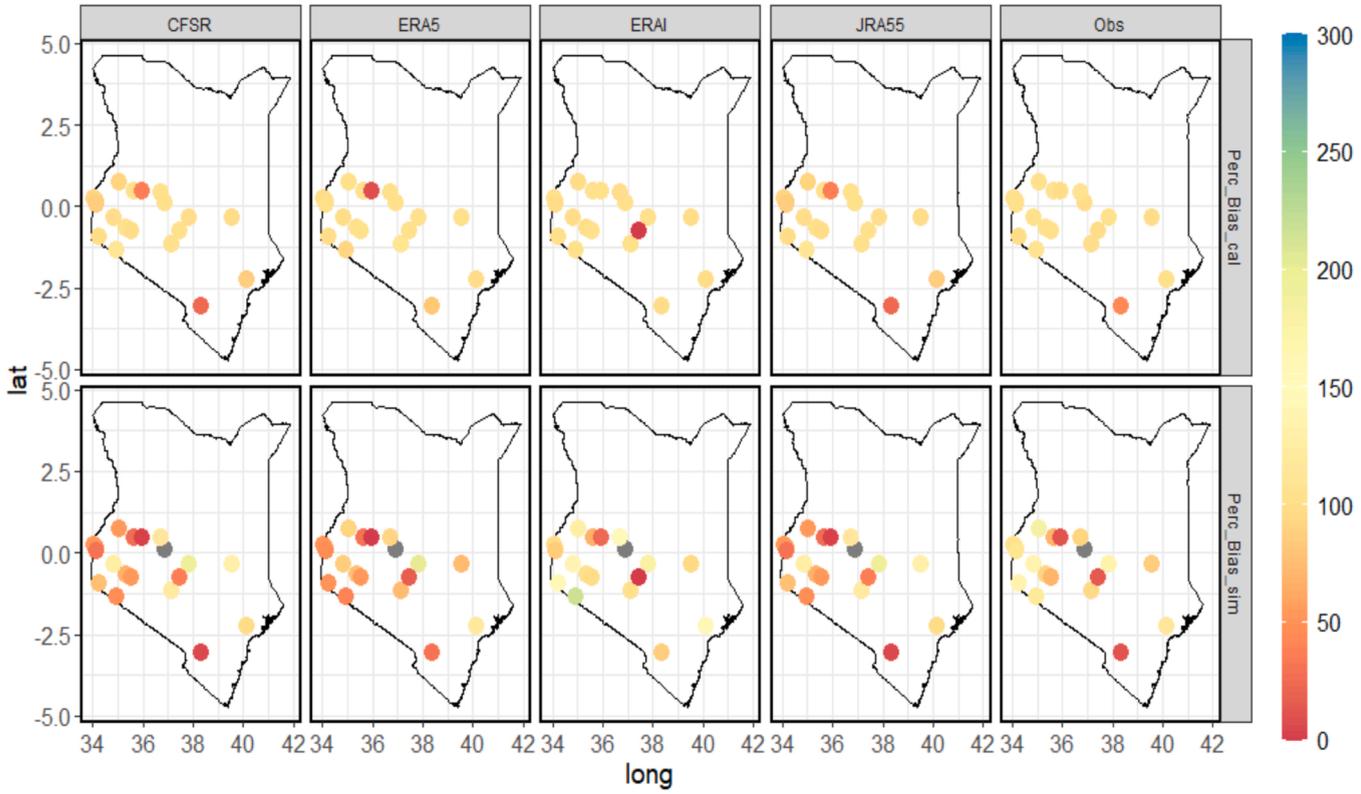
Fig. 10. GR4J model performance (Percentage KGE- Bias) in calibration (top panel) and validation (bottom panel) across the 19 catchments.
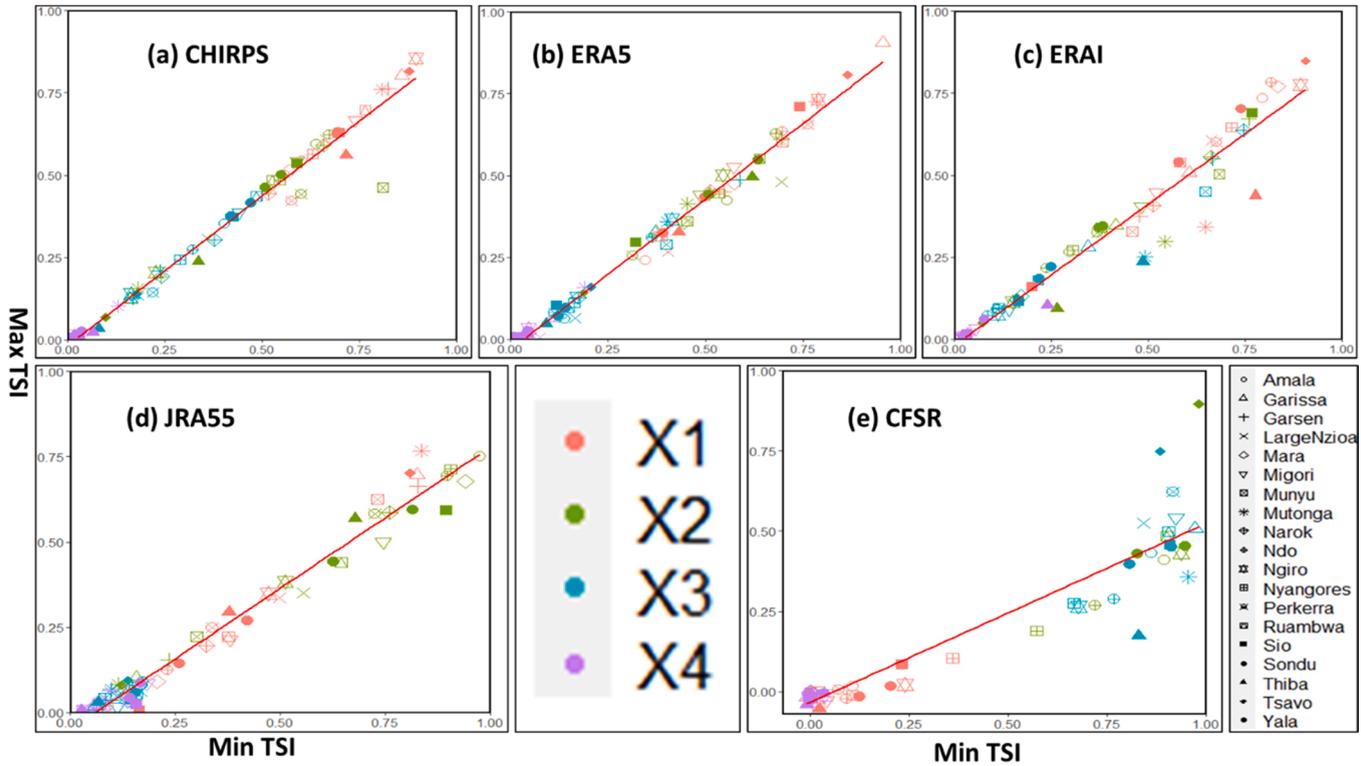
Fig. 11. Scatter plot of Sobol Total Sensitivity indices (TSI) for the different reanalysis datasets and the GR4J model parameters for the nineteen catchments. Minimum and maximum TSI were calculated for the whole data period. (a) Chirps, (b) ERA5, (c) ERAI, (d) JRA55 and (e) CFSR. Diagonal line is one-to-one line.

only whereas in the rest of the other regions the observed precipitation was less than 60 mm. Our results show that CFSR and ERAI show a positive bias for the extreme precipitation across most parts of the country in all the three seasons, like results in (Garibay et al., 2021). JRA55 has an enhanced negative bias of the extreme precipitation in most parts of the country except for isolated positive bias in the central highlands' region in JJA and OND season. ERA5 has a positive bias in MAM and OND in most parts of the country with some patches of negative bias in the western and central highlands catchments. It has an enhanced negative bias in JJA season with a positive bias in the western and coastal strip. We conclude that ERA5 outperforms other reanalysis products as it captures the wet extremes over the regions in which observations show enhanced precipitation in the respective seasons. The results are consistent with the findings in Gleixner et al. (2020) which showed both ERA5 and ERAI to have the capability to capture wet extremes in the dry seasons with ERA5 matching more closely to observations than the too wet ERA-interim. A promising performance in the ERA5 to simulate wet extremes can be attributed to improved bias correction method which incorporates aircraft measurements, satellite radiances, radiosonde measurements and surface pressure (Probst and Mauser, 2022). In addition, better performance in the central highlands can be attributed to the improved horizontal resolution in ERA5, which results in better estimates in orographic precipitation.

4.1.2. Evaluation of the reanalyses as inputs for hydrological modelling

4.1.2.1. Assessment of the overall model performance using different reanalysis. The performance of the four reanalysis datasets were evaluated using the GR4J model in the 19 catchments for the period spanning $1981 - 2016$. The KGE in calibration (top panel) and validation (bottom panel) scores obtained using different datasets for each of the catchments are represented in Fig. 8. Overall, wetland catchments in the western and highlands of Kenya obtained relatively better calibration scores compared to those in the semi-arid regions, with Yala, Sio, Nzioa and Gucha (wetland catchments) performing best and Perkerra, Ndo, Tsavo, Thiba and Tana (semi-arid catchments) performing worst. For each of the catchments, ERA5 showed better calibrated KGE scores compared to observations while CFSR and JRA55 obtained poorer KGE scores. However, we take caution in the interpretation of our results in terms of performance criteria because these catchments have a high influence of human activities such as irrigation schemes and dams. As such the low performance in some catchments may not be solely due to uncertainty in the input data.

The overall variability in GR4J model KGE scores across the four reanalyses are shown in Figure10. There are overall high-performance scores (KGE>0.5) in calibration mode in about half of the catchments for all datasets except CFSR, which suggests problems in using CFSR as hydrological model inputs in the region that cannot be solved or compensated by calibration. In Fig. 9a, ERA5, ERAI and JRA55 show similar overall performance compared to observation. The range for the performance statistic is narrower in the ERA5, ERAI and JRA55, indicating a more stable model performance in the region, while is wider in the CFSR data (Fig. 9a). In validation mode, the performance markedly decreases, as expected, for all datasets (Fig. 9b): ERAI, ERA5 and JRA55 have the highest median KGE value (just above or about 0) whereas CFSR has the lowest median values (KGE<−0.5). The range of KGE values is relatively larger compared to observations; thus, a relatively unstable prediction ability is expected for streamflow in reanalysis in the region. The range of performances is more variable in ERA5 and JRA55 and less variable in ERAI. Overall, the variability in KGE values is highest in validation than in calibration across all the reanalysis compared to observations, as expected. Fig. 10 shows the percentage
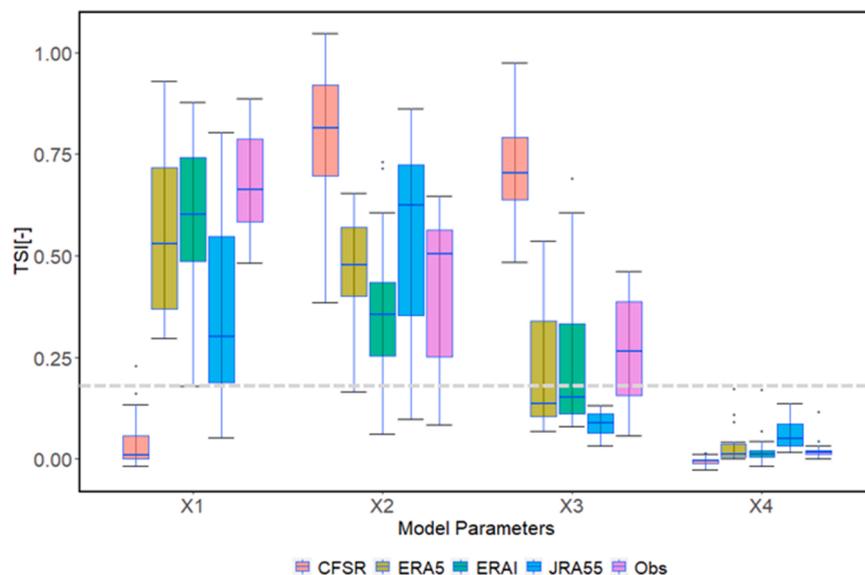
Fig. 12. Boxplots of the Sobol Total Sensitivity Indices (TSI) for the GR4J parameters for Obs. (pink), CFSR (orange), JRA55(blue), ERAI (green) and ERA5(forest-green) over the nineteen catchments. Dashed grey line represents the sensitivity threshold. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.

bias of the KGE component in each catchment in calibration (top panel) and in validation (bottom panel). The bias in all the four reanalysis is higher in calibration, whereas in validation most catchments exhibit lower biases except for Perkerra.

4.1.3. Sensitivity analysis results

4.1.3.1. Variability in sensitivity of model parameters. The GR4J's model maximum and minimum TSIs for the four reanalysis datasets is illustrated in Fig. 11. The maximum and minimum TSIs represent the variability of parameter sensitivity within the catchment with respect to KGE over the sampling periods and the variation across the four-reanalysis relative to observations. If the maximum and minimum TSIs for a parameter are equal (on the one-to-one line), that parameter has the same TSI for the sampling period, implying that the parameter is more stable across time, and would be expected to vary depending on the catchment characteristics and input data as well. In all the four datasets, the routing parameter (X4) related to the unit hydrograph is evidently the least sensitive as it is way below the threshold, followed by the capacity of the routing store (X3), whereas the two parameters governing the water-balance, i.e. the soil moisture accounting store (X1) and the groundwater exchange coefficient (X2), are the most sensitive across the datasets in most of the catchments, except for CFSR, where X1 is less sensitive.

Observations show more stability for the parameters for all catchments except six (Munyu, Thiba, Ndo, Ewaso Ngiro, Perkerra and Tsavo) with respect to reanalysis datasets (Fig. 11). In ERA5, most of the catchments showed stability in parameters except in Thiba, Tsavo, Large Nzioa and Ewaso Ngiro catchments (Fig. 11b). In ERAI, there is high variability in model parameter stability with less stability for some catchments such as Thiba, Munyu, Mutonga and Tsavo catchments (Fig. 11c). In JRA55 and CFSR (Figs. 11d and 11e respectively), the departure in sensitivity of model parameters from the diagonal is pronounced across most of the catchments. Overall, the variability in sensitivity of model parameters is high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, thus we can conclude that the reanalysis datasets are not suitable for model calibration is these catchments of Kenya characterized by arid and semi- arid conditions. However, the catchment's water balance may highly be affected by the dams constructed in the upstream areas and the massive irrigation schemes, which results in water attenuation.

4.1.3.2. Overall sensitivity of GR4J model parameters. The parameters related to water balance, i.e., the soil moisture accounting store (X1) and ground water exchange (X2), show higher sensitivity across all the four datasets except for CFSR, in which the production accounting store is less sensitive and falls below the threshold value of 0.2 from the model TSIs (Fig. 12). The first parameter responsible for water routing (X3) is less sensitive for most datasets (except CFSR), whereas the unit hydrograph parameter (X4) is the least sensitive across all the catchments in all datasets. In comparison to observations, ERA5, ERAI and JRA55 show similar parameter sensitivities of model parameters while CFSR show distinct higher variability and a difference in the parameters' sensitivity, which points to high uncertainty in the CFSR dataset. This result shows that the sensitivity of the model parameters can change with the input datasets, having very different hydrological characteristics.

4.1.4. Comparison of reanalysis datasets using Model Suitability Index

When the sensitivity indices and performance statistics are considered, it is difficult to determine which dataset is more
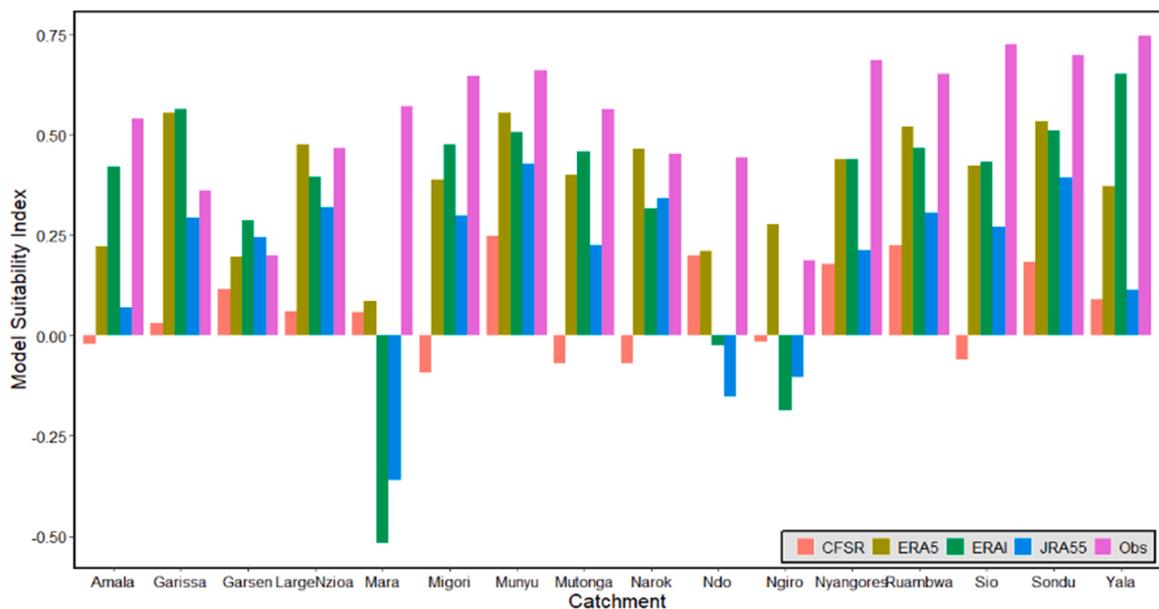
Fig. 13. Bar chart showing a comparison of model suitability in terms of performance and parameter sensitivity across different reanalysis using the Model Suitability Index (MSI).

appropriate**.** ERA5 and ERAI datasets, for example, had good and clear parameter sensitivities that captured their purposes, and the model performance score median values were higher than in CFSR and JRA55. However, the range of the performance statistics across the catchments was sometimes wider than in the other datasets, resulting in higher simulation uncertainty. When compared to the other two methods, the MSI, which considers both sensitivity indices and performance statistics, has the advantage of being easier and clearer to judge the superiority and inferiority of the datasets in terms of both model performance and parameter identifiability. We determined that a value of 0.5 for the MSI is a good MSI threshold, (Moriasi et al., 2007) We give the same weight to model performance and sensitivity, as described in the subsection 'MSI'; thus, the threshold value for good MSI is 0.5.

Combining the model performance and sensitivity indices discussed in the preceding subsections, Fig. 13 shows the MSIs for all the reanalysis datasets. The ERA5 has the highest MSI compared to observations across the nineteen catchments, followed by the ERAI reanalysis. As a result, the ERA5 and ERAI reanalysis are appropriate, at least for the selected sample of Kenyan catchments, whereas CFSR and JRA55 are least appropriate as they show lower MSI values across most of the catchments. CFSR shows negative MSI values for Amala, Migori, Mutonga, Narok, Ewaso Ngiro and Sio catchments meaning it is not appropriate for application in these Kenyan catchments. Overall, the four reanalysis datasets obtained relatively lower MSI values in Mara, Ndo, Ewasi Ngiro and Tana Garsen catchments. These catchments are mainly in arid and semi-arid areas of Kenya, like results in Section 5.2.

4.2. Discussion

4.2.1. Overall Performance of reanalysis precipitation products

In this study, we assessed four reanalysis precipitation products relative to observations for the period 1981–2016 on monthly, seasonal, and annual timescales. We also assessed how best they simulate streamflow using the GR4J model and sensitivity analysis for 19 catchments located in distinct geographical and climatic environments. Results show that the ERA5 reanalysis outperforms the other reanalysis products on monthly and seasonal scales, whereas CFSR outperforms ERA5 on annual and seasonal timescales. In general, ERA5 data were often closer to observations than other reanalysis data, which corresponds with earlier research on the datasets in different regions (e.g., Betts et al., 2019; Gleixner et al., 2020; Tarek et al., 2020), even though these studies considered different evaluation period, spatiotemporal resolution, hydrologic models, climates. However, the performance scores for the reanalysis products over the Kenyan catchments were lower which contrasts some of the studies carried out in other parts of the world with varying climates (e.g., Dhanya et al., 2017; Harada et al., 2016; Mahto et al., 2019; Wang et al., 2019), which obtained higher scores for their study areas. The low performance scores may be due to variations in the initial resolution of the datasets (Chen et al., 2018; Lemma et al., 2019) and the interpolation approach is likely to have some influence on the evaluation of various reanalysis data (Rapaić et al., 2015; Zhang et al., 2016). It is also worth noting that while the observed precipitation data are the best estimates available, they are likely to be subject to errors too (Beck at al., 2017a, Dinku et al., 2019). In addition, the seasonality of rainfall over Kenya is greatly influenced by weather phenomena such as El Niño -Southern Oscillation (ENSO) and Indian Ocean dipole (IOD) (Ayugi et al., 2020; Ojara et al., 2021; Onyutha, 2016) and play a major role in extreme rainfall events and inter-annual variability (Ongoma et al., 2015). For example, the warm phase of ENSO/El Niño results in unusually heavy rainfall, causing rare floods like as the 1997/1998 occurrence (Takaoka, 2005). ERA5, ERAI and JRA55 picked the enhanced annual precipitation totals of the strong El Niño years such as 1997/98 and 2015. However, relative to observations, ERA5 and ERAI underestimated the rainfall, and this may be attributed to incorrect configuration in the reanalysis products. For example, ERA5 precipitation is not customized to pick up the perturbations caused by the changes in the ocean – atmosphere interactions and the mountainous regions and so may miss picking up the extremes caused by events such as ENSO, thus the low performance scores.

Standardized precipitation anomalies in ERA5, ERAI and CFSR show a positive anomaly over the central highland and the western parts of Kenya (Fig. 6a, pan1) and a negative anomaly in arid and semi-arid parts in eastern and coastal lowlands in the three seasons (MAM, OND, JJA) except for JRA55 which has a stronger negative and positive anomaly in MAM and OND seasons respectively. This is consistent with study by Ongoma et al. (2018), which indicates a rise in the severity of severe precipitation events shown by a positive standardized rainfall anomaly over East Africa including the mentioned regions in Kenya. With the changing climate, temperatures in the region are projected to rise by the end of the twenty-first century, leading to an increase in rainfall extreme occurrences (Ongoma and Chen, 2017), thus, exacerbating flood risk.

Our analysis of the accuracy of precipitation reanalysis with respect to observations, over different timescales from monthly to annual, showed a positive but relatively small bias in CFSR, ERA5 and ERAI and a larger negative bias in JRA55 in MAM and OND seasons. Moreover, the first three reanalysis datasets showed a good average correlation at the monthly and seasonal scales. Therefore, the three reanalysis products have the potential to capture the rainfall seasonality and events in the study area. Recent worldwide research show that the frequency, severity, geographical range, length, and timing of climatic severe events are changing (Wainwright at al, 2021). A rise in rainfall severe events such as very wet days (R95p) and very wet days (R99p) anticipated for the future (2021–2100) (Gudoshava et al., 2020), is likely to cause the loss of life and property devastation owing to an increase in flood intensity (Finney, 2020). Therefore, further work should assess the capacity of the reanalysis datasets in capturing extreme rainfall event characteristics, such as timing and daily peaks. Performance of reanalysis as inputs into a hydrological model.

Using a hydrological model as integrator to compare simulated and observed streamflow, which can operate as an independent validation variable, is one approach to assess the quality of observation and reanalysis precipitation data. Each of the reanalysis precipitation and estimated potential evapotranspiration was supplied to the GR4J model, which were subsequently calibrated for each combination (using consistently precipitation and potential evapotranspiration from the same dataset), to independently analyze the quality of input data for each dataset relative to the observed streamflow gauge data. Streamflow gauges, of course, are subject to a variety of inaccuracies (Baldassarre and Montanari, 2009), but they represent the best available estimates for this study. Results of KGE

scores show that ERA5 is better than ERAI, JRA55 and CFSR but on overage, all the reanalyses are less skillful relative to observations across the catchments in this study and this is entirely due to the precipitation data quality. However, there is a marked improvement in the KGE scores for the catchments in the central highlands and western wet catchments which agrees with some studies on the datasets in other regions (e.g., Tarek et al., 2020; Essou et al., 2017; Lakew et al., 2020), pointing to the fact that reanalysis data can be used as a replacement for observations.

There are overall high-performance scores (KGE>0.5) in calibration mode in about half of the catchments for all datasets except CFSR, which suggests problems in using CFSR to reproduce the hydrological water balance in the region that cannot be solved or compensated by calibration (Diro et al., 2009).

4.2.2. Sensitivity analysis of model parameters

Sensitivity analysis is useful in identification of the parameters that have a strong impact on the model outputs, which in turn influences the effectiveness of the model. The greater the sensitivity of the model response to a parameter, the closer and sooner will that parameter be optimized so a high sensitivity is good. Such an in-depth analysis of a hydrological model may (i) help to identify any potential deficiencies in model structure and formulation; (ii) provide guidance for model parameterization; and (iii) provide the information content of available input data.

Based on provision of information content of the input data, different reanalyses show different sensitivities of model parameters and one that provides a higher sensitivity of model response means that it has less uncertainty and may be a lot easier to parameterize the values, but then this in practical sense does not reach the real value and the dataset with low sensitivity has got high uncertainty and model parameterization may be a lot difficult (Zeng et al., 2019). In comparison to observations, ERA5, ERAI and JRA55 show similar parameter sensitivities of model parameters while CFSR show distinct higher variability and a difference in the sensitive parameters, which points to high uncertainty in the CFSR dataset. This result shows that the sensitivity of the model parameters can change with the input datasets, having very different hydrological characteristics. Overall, the variability in sensitivity of model parameters is high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, thus we can conclude that the reanalysis datasets are not suitable for model calibration is these catchments of Kenya characterized by arid and semi- arid conditions. MSI considers both model performance and uncertainty quantitatively, therefore it can be used to compare any catchment. The ERA5 has the highest MSI compared to observations across the nineteen catchments, followed by the ERAI and JRA55, whereas CFSR has least MSI values. MSI's dependability may be increased by including more sensitivity indices and performance scores as well as assigning weights to the scores.

5. Summary and conclusion

This study addresses a notable gap that was found in the literature for evaluating the accuracy of multiple precipitation reanalysis datasets across data-scarce regions like Kenya, and for assessing their potential to supplement scarce rain gauge observations for hydrological modelling. Four different state-of-the-art reanalysis datasets were assessed. Precipitation data from ERA5 shows the highest average correlation coefficient value (0.71) on monthly timescale compared to observations and is consistently higher across all months than the other reanalysis. ERAI and CFSR have good average correlation but show larger drops in some months (especially in the drier month of August). JRA55 obtained a poor correlation coefficient of 0.46 on average. ERA5, ERAI and CFSR show higher correlations to observations in rainy months (March-April-May and October- November-December) and lower in the dry months (June-July- August), whereas JRA55 shows worst correlations during both rainy seasons. On annual timescales, CFSR, ERAI and ERA5 showed better CC indices of 0.60, 0.46, 0.52 respectively whereas JRA obtained lower CC of 0.25. ERA5 and JRA55 show a positive bias of 45% and 171% respectively, whereas ERAI and CFSR show negative bias of -26 and -85%.

Spatial rainfall patterns directly affect temporal distribution key in driving runoff and soil erosion processes, which is useful in management of hydrological risks and generation of sediments from rainwater. Monthly standardised anomaly maps in ERA5, ERAI and JRA55 showed a positive anomaly over the central highland and western parts. In the arid and semi-arid parts in the eastern and coastal lowlands parts of Kenya the three datasets showed enhanced negative anomaly. On seasonal timescales, ERA5, ERAI and CFSR show positive anomalies in Western and central highland in the three seasons except for JRA55 which has a stronger negative and positive anomaly in MAM and OND seasons respectively. Extreme precipitation showed a positive bias in CFSR, ERA5 and ERAI in MAM and OND seasons whereas JRA55 has enhanced negative bias in most parts of the country except for isolated positive bias in the central highlands' region in JJA and OND seasons.

The performance of GR4J model when forced with different reanalysis in the 19 catchments reveals a bigger role of localized catchment characteristics and process in model calibration. Wetland catchments in the western and highlands of Kenya obtained relatively better calibration scores compared to those in the semi-arid regions, with Yala, Sio, Nzioa and Gucha (wetland catchments) performing best and Perkerra, Ndo, Tsavo, Thiba and Tana (semi-arid catchments) performing worst. For each of the catchments, ERA5 showed better calibrated KGE scores compared to observations while CFSR and JRA55 obtained poorer KGE scores. The range of KGE values was relatively larger compared to observations; thus, a relatively unstable prediction ability is expected for streamflow in reanalysis for Kenyan catchments. The range of performances is more variable in ERA5 and JRA55 and less variable in ERAI. Overall, the variability in KGE values is highest in validation than in calibration across all the reanalysis compared to observations, as expected.

Sensitivity analysis allows the reduction of parameters incorporated in optimization by determining the convergence most influential model parameters. Sensitivity analysis revealed that in all the four datasets, the routing parameter (X4) related to the unit hydrograph was evidently the least sensitive, followed by the capacity of the routing store (X3), whereas the two parameters governing the water-balance, i.e. the soil moisture accounting store (X1) and the groundwater exchange coefficient (X2), are the most sensitive across the datasets in most of the catchments, except for CFSR, where X1 was less sensitive, with ERA5 showing a highest sensitivity in

the model parameters. However, the variability in sensitivity of model parameters was high in Thiba, Munyu, Pekerra and Ewaso Ngiro across all the datasets, thus we conclude that model calibration in arid and semi- arid catchments of Kenya does not yield skillful results using the reanalysis data. The MSI aggregates both sensitivity indices & performance statistics, providing a clear index to judge the superiority (or inferiority) of a reanalysis with respect to observations. On average ERA5, ERAI (& JRA55) have better MSI scores across most of the Kenyan catchments: ERAI & ERA5 perform better than JRA55 & CFSR, and lead to more robust model parameters. Using a catchment model and combined sensitivity - model performance analysis, allows an evaluation of the impact of the variability in the rainfall products throughout the catchment modelling process.

In conclusion, in this study we have demonstrated the usefulness of reanalysis rainfall products as potential alternatives for hydrological applications in Kenya. We assessed the suitability of reanalysis precipitation datasets for hydrological modelling across Kenyan catchments, but first assessed the propagation of errors when reanalysis is compared to observations. We performed the assessment on monthly, seasonal, and annual timescales. Then, using a lumped bucket-style hydrological model, we assessed the model performance via the KGE criterion and parameter uncertainty via Sobol's Sensitivity Analysis for four different reanalyses: - ERA5, ERAI, JRA55 and CFSR across 19 catchments. The parametric and model input uncertainty is investigated using the sensitivity indices and the comprehensive model performance analysis is used to examine the model's input strength, i.e., the amount to which the model captures the dynamics of rainfall–runoff processes with respect to different forcing. We also coupled the results of the performance scores and sensitivity indices to compute MSI for the 19 catchments.

We acknowledge the value and need of additional work, if reliable data at higher temporal frequency becomes available and can be used, as it contains more information. However, this is a big limitation for the current study due to high data gaps in the daily data (river discharge data used in the current study) and the lack of higher temporal resolution hydrological data. Future work should concentrate on assessing the sub-daily performance of hydrological modelling with reanalysis, testing its quality on other additional catchments in countries in the region with quality observed gauge data, but prior investments in data collection in Kenya seem to be needed. Our approach may be extended to various conceptual rainfall -runoff models as well as physically based distributed rainfall–runoff models. The MSI analysis is a practical method for weeding out the appropriate model input on a catchment scale basis, however a more robust analysis where weights are assigned would yield some improvements in the results. To fully ascertain the potential of alternative model forcing, catchments characteristics and human influence such as dams and reservoirs should be modeled.

Finally, it is essential to note that this work does not promote the use of products such as reanalysis to replace observed data from weather stations, nor can it be understood as giving reason to continue the present trend of retiring additional stations. Quality controlled ground observations still act as the best data for research. The ERA5 results demonstrate that atmospheric reanalysis has likely reached the stage where they can consistently supplement data from weather stations and offer trustworthy proxies in places with less dense station networks, at least across Kenya. Overall, reanalysis can be a viable alternative to observations in ungauged catchments, but the associated uncertainties need to be carefully communicated for informed choice of hydrological modelling tools.

CRediT authorship contribution statement

Maureen A. Wanzala: Conceptualization, Data curation- Lead, Methodology- Lead, Investigation, Software, Formal analysis - Lead, Visualization, Writing – original draft, Writing – review & editing. **Andrea Ficchì**: Conceptualization, Supervision, Data curation, Supporting, Methodology, Supporting, Investigation, Software, Analysis, Supporting, Visualization, Writing – review & editing, Supporting. **Hannah L. Cloke**: Conceptualization, Methodology, Investigation, Writing – review & editing, Project administration, Supervision. **Elisabeth M. Stephens**: Conceptualization, Methodology, Investigation, Writing – review & editing, Supporting, Project administration, Supervision, Funding acquisition. **Heou M. Badjana**: Methodology, Data curation, Supporting, Writing – review & editing, Supporting. **David A. Lavers**: Data curation, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

No. NE/R004668/1).

References

Acharya, Suwash Chandra, Nathan, Rory, Wang, Quan J., Su, Chun-Hsu, Eizenberg, Nathan, 2019. An evaluation of daily precipitation from a regional atmospheric reanalysis over Australia. Hydrol. Earth Syst. Sci. 23 (8), 3387–3403.

Adeyeri, O.E., Laux, P., Arnault, J., Lawin, A.E., Kunstmann, H., 2020. Conceptual hydrological model calibration using multi-objective optimization techniques over the transboundary Komadugu-Yobe basin, Lake Chad Area, West Africa. J. Hydrol. Reg. Stud. https://doi.org/10.1016/j.ejrh.2019.100655.

Alemayehu, Tadesse, Kilonzo, Fidelis, van Griensven, Ann, Bauwens, Willy, 2018. Evaluation and application of alternative rainfall data sources for forcing hydrologic models in the Mara Basin. Hydrol. Res. 49 (4), 1271–1282.

Arshad, Muhammad, Ma, Xieyao, Yin, Jun, Ullah, Waheed, Liu, Mengyang, Ullah, Irfan, 2021. Performance evaluation of ERA-5, JRA-55, MERRA-2, and CFS-2 reanalysis datasets, over diverse climate regions of Pakistan. Weather Clim. Extrem. 33, 100373.

Ayugi, Brian, Tan, Guirong, Niu, Rouyun, Dong, Zeyao, Ojara, Moses, Mumo, Lucia, Ongoma, Victor, 2020. Evaluation of meteorological drought and flood scenarios over Kenya. East Afr. Atmosph. 11 (3), 307.

Ayugi, Brian Odhiambo, Wen, Wang, Chepkemoi, Daisy, 2016. Analysis of spatial and temporal patterns of rainfall variations over Kenya. J. Environ. Earth Sci. 6 (11).

Baldassarre, G.Di, Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrol. Earth Syst. Sci. 13 (6), 913–921.

Beck, Hylke E., van Dijk, Albert, I.J.M., de Roo, Ad, Dutra, Emanuel, Fink, Gabriel, Orth, Rene, Schellekens, Jaap, 2017a. Global evaluation of runoff from ten state-of-the-art hydrological models. Hydrol. Earth Syst. Sci. 21 (6), 2881–2903.

Beck, Hylke E., Vergopolan, Noemi, Pan, Ming, Levizzani, Vincenzo, van Dijk, Albert I.J.M., Weedon, Graham P., Wood, Eric F., 2017b. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrol. Earth Syst. Sci. 21 (12), 6201–6217.

Beck, Hylke E., Pan, Ming, Miralles, Diego, G., Reichle, Rolf, H., Dorigo, Wouter, A., Hahn, Sebastian, Parinussa, Robert M., 2021. Evaluation of 18 satellite-and model-based soil moisture products using in situ measurements from 826 sensors. Hydrol. Earth Syst. Sci. 25 (1), 17–40.

Betts, Alan K., Chan, Darren Z., Desjardins, Raymond L., 2019. Near-surface biases in ERA5 over the Canadian prairies. Front. Environ. Sci. 7, 129.

Beven, Keith, 2012. Rainfall-Runoff Modelling: The Primer: Second Edition. Rainfall-Runoff Modelling: The Primer, second ed. John Wiley and Sons. https://doi.org/10.1002/9781119951001.

Bitew, Menberu M., Gebremichael, Mekonnen, Ghebremichael, Lula T., Bayissa, Yared A., 2012. Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia. J. Hydrometeorol. 13 (1), 338–350.

Black, Emily, Slingo, Julia, Sperber, Kenneth R., 2003. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. Mon. Weather Rev. 131 (1), 74–94.

Blau, M.T., Ha, K.-J., 2020. The Indian Ocean dipole and its impact on East African short rains in two CMIP5 historical scenarios with and without anthropogenic influence. J. Geophys. Res. Atmosph. 125 (16), e2020JD033121.

Blöschl, G.ünter, Sivapalan, Murugesu, 1995. Scale issues in hydrological modelling: a review. Hydrol. Process. 9 (3–4), 251–290.

Chen, Jie, Brissette, François P., Chen, Hua, 2018. Using reanalysis-driven regional climate model outputs for hydrology modelling. Hydrol. Process. 32 (19), 3019–3031.

Collier, Emily, & Mölg, Thomas. (2016). Analysis and simulation of recent climate variability in the high-mountain regions of East Africa. In EGU General Assembly Conference Abstracts (pp. EPSC2016–6136).

Coron, Laurent, Delaigue, Olivier, Thirel, Guillaume, Perrin, Charles, Michel, Claude, Andréassian, Vazken, … Mathevet, Thibaut. (2019). airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling (v. 1.2. 13.16).

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Vitart, F., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137 (656), 553–597. https://doi.org/10.1002/qj.828.

Delaigue, Olivier, Thirel, Guillaume, Coron, Laurent, & Brigode, Pierre. (2019). airGR and airGRteaching: two packages for rainfall-runoff modeling and teaching hydrology. In 15th edition of the International R User Conference (p. 1).

Dhanya, C.T., Villarini, Gabriele, 2017. An investigation of predictability dynamics of temperature and precipitation in reanalysis datasets over the continental United States. Atmos. Res. 183, 341–350.

Dile, Yihun Taddele, Srinivasan, Raghavan, 2014. Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. JAWRA J. Am. Water Resour. Assoc. 50 (5), 1226–1241.

Dinku, T., 2019. Challenges with availability and quality of climate data in Africa. Extreme Hydrology and Climate Variability. Elsevier, pp. 71–80.

Dinku, Tufa, Funk, Chris, Peterson, Pete, Maidment, Ross, Tadesse, Tsegaye, Gadain, Hussein, Ceccato, Pietro, 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Q. J. R. Meteorol. Soc. 144, 292–312.

Diro, G.T., Grimes, D.I.F., Black, E., O'Neill, A., Pardo-Iguzquiza, E., 2009. Evaluation of reanalysis rainfall estimates over Ethiopia. Int. J. Climatol. A J. R. Meteorol. Soc. 29 (1), 67–78.

Emerton, Rebecca, Cloke, H.L., Stephens, E.M., Zsoter, E., Woolnough, S.J., Pappenberger, F., 2017. Complex picture for likelihood of ENSO-driven flood hazard. Nat. Commun. 8 (1), 1–9.

Essou, Gilles R.C., Brissette, François, Lucas-Picher, Philippe, 2017. The use of reanalyses and gridded observations as weather input data for a hydrological model: comparison of performances of simulated river flows based on the density of weather stations. J. Hydrometeorol. 18 (2), 497–513.

Ficchì, Andrea, Perrin, Charles, Andréassian, Vazken, 2016. Impact of temporal resolution of inputs on hydrological model performance: an analysis based on 2400 flood events. J. Hydrol. 538, 454–470.

Ficchì, Andrea, Perrin, Charles, Andréassian, Vazken, 2019. Hydrological modelling at multiple sub-daily time steps: model improvement via flux-matching. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2019.05.084.

Finney, Declan L., Marsham, John H., Walker, Dean P., Birch, Cathryn E., Woodhams, Beth J., Jackson, Lawrence S., Hardy, Sam, 2020. The effect of westerlies on East African rainfall and the associated role of tropical cyclones and the Madden–Julian Oscillation. Q. J. R. Meteorol. Soc. 146 (727), 647–664.

Fortin, Vincent, Roy, Guy, Donaldson, Norman, Mahidjiba, Ahmed, 2015. Assimilation of radar quantitative precipitation estimations in the Canadian Precipitation Analysis (CaPA). J. Hydrol. 531, 296–307.

Funk, Chris, Peterson, Pete, Landsfeld, Martin, Pedreros, Diego, Verdin, James, Shukla, Shraddhanand, Michaelsen, Joel, 2015. The climate hazards infrared precipitation with stations - a new environmental record for monitoring extremes. Sci. Data 2, 1–21. https://doi.org/10.1038/sdata.2015.66.

Garibay, Victoria M., Joseph, H.A.Guillaume, Crokeab, Barry F.W., Jakemanbc, Anthony J., 2021. Evaluation of Reanalysis Precipitation Data and Potential Bias Correction Methods for Use in Data-Scarce Areas. Water Resources Management 35, 1587–1602. https://doi.org/10.1007/s11269-021-02804-8.

Gabellani, S., Boni, G., Ferraris, L., Von Hardenberg, J., Provenzale, A., 2007. Propagation of uncertainty from rainfall to runoff: a case study with a stochastic rainfall generator. Adv. Water Resour. 30 (10), 2061–2071.

Gleixner, Stephanie, Demissie, Teferi, Diro, Gulilat Tefera, 2020. Did ERA5 improve temperature and precipitation reanalysis over East Africa? Atmosphere 11 (9), 996.

Golian, Saeed, Murphy, Conor, 2021. Evaluation of sub-selection methods for assessing climate change impacts on low-flow and hydrological drought conditions. Water Resour. Manag. 35 (1), 113–133.

Gosset, Marielle, Viarre, Julien, Quantin, Guillaume, Alcoba, Matias, 2013. Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. Q. J. R. Meteorol. Soc. 139 (673), 923–940. https://doi.org/10.1002/qj.2130.

Gudoshava, Masilin, Misiani, Herbert O., Segele, Zewdu T., Jain, Suman, Ouma, Jully O., Otieno, George, Osima, Sarah, 2020. Projected effects of 1.5C and 2C global warming levels on the intra-seasonal rainfall characteristics over the Greater Horn of Africa. Environ. Res. Lett. 15 (3), 34037.

Guo, Binbin, Zhang, Jing, Xu, Tingbao, Croke, Barry, Jakeman, Anthony, Song, Yongyu, Liao, Weihong, 2018. Applicability assessment and uncertainty analysis of multi-precipitation datasets for the simulation of hydrologic models. Water 10 (11), 1611.

Gupta, Hoshin V., Kling, Harald, Yilmaz, Koray, K., Martinez, Guillermo F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377 (1), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Hamon, W.Russell, 1960. Estimating Potential Evapotranspiration. Massachusetts Institute of Technology.

Harada, Yayoi, Kamahori, Hirotaka, Kobayashi, Chiaki, Endo, Hirokazu, Kobayashi, Shinya, Ota, Yukinari, Takahashi, Kiyotoshi, 2016. The JRA-55 Reanalysis: Representation of atmospheric circulation and climate variability. J. Meteorol. Soc. Jpn. Ser. II 94 (3), 269–302.

Hargreaves, George H., Samani, Zohrab A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1 (2), 96–99.

Hartman, Adam T., 2018. An analysis of the effects of temperatures and circulations on the strength of the low-level jet in the Turkana Channel in East Africa. Theor. Appl. Climatol. 132 (3), 1003–1017.

Hersbach, H. (2018). Operational global reanalysis: progress, future directions and synergies with NWP. European Centre for Medium Range Weather Forecasts.

Indeje, M., Semazzi, F.H.M., 2000. Relationships between QBO in the lower equatorial stratospheric zonal winds and East African seasonal rainfall. Meteor. Atmos. Phys. vol. 73, 227–244. https://doi.org/10.1007/s007030050075.

Khan, Sadiq I., Adhikari, Pradeep, Hong, Yang, Vergara, H., F Adler, R., Policelli, F., Okello, L., 2011. Hydroclimatology of Lake Victoria region using hydrologic model and satellite remote sensing data. Hydrol. Earth Syst. Sci. 15 (1), 107–117.

Kilavi, Mary, MacLeod, Dave, Ambani, Maurine, Robbins, Joanne, Dankers, Rutger, Graham, Richard, Todd, Martin C., 2018. Extreme rainfall and flooding over central Kenya including Nairobi city during the long-rains season 2018: causes, predictability, and potential for early warning and actions. Atmosphere 9 (12), 472.

Kinuthia, Joseph Hiri, 1992. Horizontal and vertical structure of the Lake Turkana jet. J. Appl. Meteorol. Climatol. 31 (11), 1248–1274.

Klemeš, Vit, 1986. Operational testing of hydrological simulation models. Hydrol. Sci. J. 31 (1), 13–24.

Kling, Harald, Fuchs, Martin, Paulin, Maria, 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. J. Hydrol. 424, 264–277.

Knoben, Wouter J.M., Freer, Jim E., Woods, Ross A., 2019. Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23 (10), 4323–4331.

Kobayashi, Shinya, Yukinari, Ota, Yayoi, Harada, Ebita, Ayataka, Moriya, Masami, Onoda, Hirokatsu, Takahashi, Kiyotoshi, 2015. The JRA-55 reanalysis: general specifications and basic characteristics. J. Meteorol. Soc. Jpn. Ser. II 93 (1), 5–48. https://doi.org/10.2151/jmsj.2015-001.

Koukoula, Marika, Nikolopoulos, Efthymios I., Dokou, Zoi, Anagnostou, Emmanouil N., 2020. Evaluation of global water resources reanalysis products in the Upper Blue Nile River Basin. J. Hydrometeorol. 21 (5), 935–952.

Lakew, Haileyesus Belay, Moges, Semu Ayalew, Asfaw, Dereje Hailu, 2020. Hydrological performance evaluation of multiple satellite precipitation products in the upper Blue Nile basin, Ethiopia. J. Hydrol. Reg. Stud. 27, 100664.

Langat, Philip Kibet, Kumar, Lalit, Koech, Richard, 2017. Temporal variability and trends of rainfall and streamflow in Tana River Basin, Kenya. Sustainability 9 (11), 1963.

Le, Alyssa M., Pricope, Narcisa G., 2017. Increasing the accuracy of runoffand streamflow simulation in the Nzoia Basin, Western Kenya, through the incorporation of satellite-derived CHIRPS data. Water 9 (2). https://doi.org/10.3390/w9020114.

Lemma, Estifanos, Upadhyaya, Shruti, Ramsankaran, R.A.A.J., 2019. Investigating the performance of satellite and reanalysis rainfall products at monthly timescales across different rainfall regimes of Ethiopia. Int. J. Remote Sens. 40 (10), 4019–4042.

MacLeod, D.A., Dankers, R., Graham, R., Guigma, K., Jenkins, L., Todd, M.C., Mwangi, E., 2021. Drivers and subseasonal predictability of heavy rainfall in equatorial East Africa and relationship with flood risk. J. Hydrometeorol. 22 (4), 887–903.

Mahto, Shanti Shwarup, Mishra, Vimal, 2019. Does ERA-5 outperform other reanalysis products for hydrologic applications in India? J. Geophys. Res. Atmosph. 124 (16), 9423–9441.

Mathevet, T., Michel, C., Andreassian, V., Perrin, C., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. IAHS Publ. 307, 211.

Menne, Matthew J., Williams, Claude N., Gleason, Byron E., Rennie, J.Jared, Lawrimore, Jay H., 2018. The global historical climatology network monthly temperature dataset, version 4. J. Clim. 31 (24), 9835–9854.

Meresa, Hadush, Murphy, Conor, Fealy, Rowan, Golian, Saeed, 2021. Uncertainties and their interaction in flood hazard assessment with climate change. Hydrol. Earth Syst. Sci. 25 (9), 5237–5257.

Michel, Claude, 1983. Que peut-on faire en hydrologie avec modèle conceptuel à un seul paramètre? La Houille Blanc (1), 39–44.

Moriasi, Daniel N., Arnold, Jeffrey G., Van Liew, Michael W., Bingner, Ronald L., Harmel, R.Daren, Veith, Tamie L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Mostafaie, A., Forootan, E., Safari, A., Schumacher, M., 2018. Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data. Comput. Geosci. 22 (3), 789–814. https://doi.org/10.1007/s10596-018-9726-8.

Nicholson, S.E., Klotter, D., Zhou, L., Hua, W., 2019. Validation of satellite precipitation estimates over the Congo Basin. J. Hydrometeorol. 20 (4), 631–656.

Njogu, Humphrey Waita, 2021. Effects of floods on infrastructure users in Kenya. J. Flood Risk Manag. 14 (4), e12746.

Nkiaka, E., Nawaz, N.R., Lovett, J.C., 2017. Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: case study of the Logone catchment, Lake Chad Basin. Meteorol. Appl. 24 (1), 9–18.

Nossent, Jiri, Elsen, Pieter, Bauwens, Willy, 2011. Sobol'sensitivity analysis of a complex environmental model. Environ. Model. Softw. 26 (12), 1515–1525.

Nyenzi, B.S. (1988). Mechanisms of East African rainfall variability.

Ogallo, L.A., 1993. Dynamics of the East African climate. Proc. Indian Acad. Sci. -Earth Planet. Sci. 102 (1), 203–217.

Ogwang, Bob Alex, Ongoma, Victor, Xing, Li, Ogou, Katchele Faustin, 2015. Influence of Mascarene high and Indian Ocean dipole on East African extreme weather events. Geogr. Pannonica 19 (2), 64–72.

Ojara, Moses A., Yunsheng, Lou, Ongoma, Victor, Mumo, Lucia, Akodi, David, Ayugi, Brian, Ogwang, Bob Alex, 2021. Projected changes in East African climate and its impacts on climatic suitability of maize production areas by the mid-twenty-first century. Environ. Monit. Assess. 193 (12), 1–24.

Ongoma, V., Guirong, T., Ogwang, B.A., Ngarukiyimana, J.P., 2015. Diagnosis of seasonal rainfall variability over east Africa: a case study of 2010-2011 drought over Kenya. Pak. J. Meteorol. Vol. 11.

Ongoma, Victor, Chen, Haishan, 2017. Temporal and spatial variability of temperature and precipitation over East Africa from 1951 to 2010. Meteorol. Atmos. Phys. 129 (2), 131–144.

Ongoma, Victor, Chen, Haishan, Omony, George William, 2018. Variability of extreme weather events over the equatorial East Africa, a case study of rainfall in Kenya and Uganda. Theor. Appl. Climatol. 131 (1–2), 295–308.

Onyutha, Charles, 2016. Variability of seasonal and annual rainfall in the River Nile riparian countries and possible linkages to ocean–atmosphere interactions. Hydrol. Res. 47 (1), 171–184.

Opere, A., 2013. Floods in Kenya. In: Developments in Earth Surface Processes, Vol. 16. Elsevier.

Oudin, Ludovic, Andréassian, Vazken, Perrin, Charles, Anctil, François, 2004. Locating the sources of low-pass behavior within rainfall-runoff models. Water Resour. Res. 40, 11.

Owiti, Z.O., Ogallo, L.A., Mutemi, J., 2008. Linkages between the Indian Ocean Dipole and East African seasonal rainfall anomalies. J. Kenya Meteor. Soc. vol. 2, 3–17.

Perrin, Charles, Michel, Claude, Andréassian, Vazken, 2003. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 279 (1–4), 275–289.

Probst, E., Mauser, W., 2022. Evaluation of ERA5 and WFDE5 forcing data for hydrological modelling and the impact of bias correction with regional climatologies: a case study in the Danube River Basin. J. Hydrol.: Reg. Stud. 40, 101023.

Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., Andréassian, V., 2011. A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. J. Hydrol. 411 (1–2), 66–76.

Rapaić, Maja, Brown, Ross, Markovic, Marko, Chaumont, Diane, 2015. An evaluation of temperature and precipitation surface-based and reanalysis datasets for the Canadian Arctic, 1950–2010. Atmosph. Ocean 53 (3), 283–303.

Saha, Suranjana, Moorthi, Shrinivas, Pan, Hua-Lu, Wu, Xingren, Wang, Jiande, Nadiga, Sudhir, Goldberg, Mitch, 2010. The NCEP climate forecast system reanalysis. Bull. Am. Meteorol. Soc. 91 (8), 1015–1058. https://doi.org/10.1175/2010bams3001.1.

Saltelli, Andrea, Tarantola, Stefano, Campolongo, Francesca, 2000. Sensitivity analysis as an ingredient of modeling. Stat. Sci. 377–395.

Schulzweida, Uwe. (2019). CDO user guide. Clim Data Oper.

Shayeghi, Afshin, Azizian, Asghar, Brocca, Luca, 2020. Reliability of reanalysis and remotely sensed precipitation products for hydrological simulation over the Sefidrood River Basin, Iran. Hydrol. Sci. J. 65 (2), 296–310.

Shin, Mun-Ju, Kim, Chung-Soo, 2017. Assessment of the suitability of rainfall–runoff models by coupling performance statistics and sensitivity analysis. Hydrol. Res. 48 (5), 1192–1213.

Shin, Mun-Ju, Gitau, Margaret W., Kiggundu, Nicholas, Moriasi, Daniel, Mishili, Fulgence, 2013. Addressing ten questions about conceptual rainfall–runoff models with global sensitivity analyses in R. Journal of Hydrology. https://doi.org/10.1016/j.jhydrol.2013.08.047.

Shin, Mun-Ju, Guillaume, Joseph, H.A., Croke, Barry, F.W., Jakeman, Anthony J., 2015. A review of foundational methods for checking the structural identifiability of models: results for rainfall-runoff. J. Hydrol. 520, 1–16.

Sun, Qiaohong, Miao, Chiyuan, Duan, Qingyun, Ashouri, Hamed, Sorooshian, Soroosh, Hsu, Kuo-Lin, 2018. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Rev. Geophys. 56 (1), 79–107.

ReliefWeb, 2019b: UNICEF sends supplies to flood-affected children and families in 12counties in Kenya - Kenya | ReliefWeb. (n.d.). Retrieved March 3, 2020, from ⟨https://reliefweb.int/report/kenya/unicef-sends-supplies-flood-affected-children-and-families-12-counties-kenya⟩.

ReliefWeb, 2019a - Kenya: Floods and Landslides - Oct 2019 | ReliefWeb. (n.d.). Retrieved March 3, 2020, from ⟨https://reliefweb.int/disaster/fl-2019–000138-ken⟩.

Takaoka, S., 2005. Impact of the 1997–1998 El Niño rains on farms in the Mount Kenya region. Mt. Res. Dev. 25 (4), 326–331.

Tang, Guoqiang, Clark, Martyn P., Papalexiou, Simon Michael, Ma, Ziqiang, Hong, Yang, 2020. Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. Remote Sens. Environ. 240, 111697.

Tapiador, Francisco J., Turk, Francis J., Petersen, Walt, Hou, Arthur Y., García-Ortega, Eduardo, Machado, Luiz A.T., Huffman, George J., 2012. Global precipitation measurement: methods, datasets and applications. Atmos. Res. 104, 70–97.

Tarek, Mostafa, Brissette, François P., Arsenault, Richard, 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrol. Earth Syst. Sci. 24 (5), 2527–2544.

Tarek, Mostafa, Brissette, François, Arsenault, Richard, 2021. Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies. Hydrol. Earth Syst. Sci. 25 (6), 3331–3350.

Tesfaye, T.W., Dhanya, C., Gosain, A., 2017. Evaluation of ERA-interim, MERRA, NCEP-DOE R2 and CFSR reanalysis precipitation data using gauge observation over ethiopia for a period of 33 years. AIMS Environ. Sci. 4, 596–620.

van Werkhoven, Kathryn, Wagener, Thorsten, Reed, Patrick, Tang, Yong, 2009. Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. Advances in Water Resources 32, 1154–1169. https://doi.org/10.1016/j.advwatres.2009.03.002.

Van Esse, W.R., Perrin, C., Booij, Martijn J., Augustijn, Dionysius C.M., Fenicia, F., Kavetski, D., Lobligeois, F., 2013. The influence of conceptual model structure on model performance: a comparative study for 237 French catchments. Hydrol. Earth Syst. Sci. 17 (10), 4227–4239.

van Griensven, A. van, Meixner, Thomas, Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. J. Hydrol. 324 (1–4), 10–23.

Viglione, Alberto, Chirico, Giovanni Battista, Komma, J.ürgen, Woods, Ross, Borga, Marco, Blöschl, G.ünter, 2010. Quantifying space-time dynamics of flood event types. J. Hydrol. 394 (1–2), 213–229.

Vischel, T., Quantin, G., Lebel, Thierry, Viarre, J., Gosset, Marielle, Cazenave, Frédéric, Panthou, G., 2011. Generation of high-resolution rain fields in West Africa: evaluation of dynamic interpolation methods. J. Hydrometeorol. 12 (6), 1465–1482.

Wainwright, Caroline M., Finney, Declan L., Kilavi, Mary, Black, Emily, Marsham, John H., 2021. Extreme rainfall in East Africa, October 2019–January 2020 and context under future climate change. Weather 76 (1), 26–31.

Wang, Gefei, Zhang, Xiaowen, Zhang, Shiqiang, 2019. Performance of three reanalysis precipitation datasets over the qinling-daba mountains, eastern fringe of Tibetan Plateau, China. Adv. Meteorol. 2019.

Worqlul, Abeyou W., Yen, Haw, Collick, Amy S., Tilahun, Seifu A., Langan, Simon, Steenhuis, Tammo S., 2017. Evaluation of CFSR, TMPA 3B42 and ground-based rainfall data as input for hydrological models, in data-scarce regions: the upper Blue Nile Basin, Ethiopia. Catena 152, 242–251.

Xu, Lei, Chen, Nengcheng, Moradkhani, Hamid, Zhang, Xiang, Hu, Chuli, 2020. Improving global monthly and daily precipitation estimation by fusing gauge observations, remote sensing, and reanalysis data sets. Water Resour. Res. 56 (3), e2019WR026444.

Yang, Wenchang, Seager, Richard, Cane, Mark A., Lyon, Bradfield, 2015. The annual cycle of East African precipitation. J. Clim. 28 (6), 2385–2404.

Zaitchik, Benjamin F., Montanari, Alberto, Beven, Keith J., Kirkby, Michael J., Mike, J., Pappenberger, Floria N., Smith, Tyler Jon, Obled, Charles, 2011. Rainfall estimation in the Sahel. Part II: evaluation of rain gauge networks in the CILSS countries and objective intercomparison of rainfall products. J. Hydrol. 9 (3), 2. https://doi.org/10.1016/j.jhydrol.2019.05.084.

Zeng, Ling, Xiong, Lihua, Liu, Dedi, Chen, Jie, Kim, Jong Suk, 2019. Improving parameter transferability of GR4J model under changing environments considering nonstationarity. Water 11 (10). https://doi.org/10.3390/w11102029.

Zhang, Chi, Academy, Chinese, 2016. Hu et al-2016-JGR. J. Geophys. Res. Atmosph. 121 (June), 5648–5663. https://doi.org/10.1002/2016JD024781.Received.

Hydrological Sciences JOURNAL

Detecting trends in flood series and shifts in flood timing across Kenya

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

SCHOLARONE™
Manuscripts

Detecting trends in flood series and shifts in flood timing across Kenya

Maureen A. Wanzala[*1], Hannah L. Cloke[1,2,3,4,7], Elisabeth M. Stephens[2,5], Andrea Ficchi [1,6], Shaun Harrigan[7]

[1]*Department of Geography & Environmental Science, University of Reading, Reading, UK*

[2]*Department of Meteorology, University of Reading, Reading, UK*

[3] *Centre of Natural Hazards and Disaster Science, CNDS, Uppsala, Sweden*

[4]*Department of Earth Sciences, Uppsala University, Uppsala, Sweden*

[5]*Red Cross Red Crescent Climate Centre, The Hague, Netherlands*

[6]*Politecnico di Milano, Milan, Italy*

[7] *European Centre for Medium-Range Weather Forecasts, Reading, UK*

* **Corresponding Author:** Maureen A. Wanzala; m.a.wanzala@pgr.reading.ac.uk

Current Affiliation: - Department of Geography & Environmental Science, University of Reading, Reading, UK

Authors' contributions

Maureen A. Wanzala (65%): - Conceptualization, Data curation- Lead, Methodology- Lead, Investigation, Formal analysis - Lead, Visualization, Writing – original draft, Writing - Review & Editing. Hannah L. Cloke (15%): Supervision, Writing - Review & Editing, Project administration, Elisabeth M. Stephens (10%): Supervision, Methodology, Writing - Review & Editing- Support, Project administration, Andrea Ficchì (10%): Conceptualization, Supervision, Methodology- Support, Analysis - Support, Visualization, Shaun Harrigan (5%): Methodology - Support, Writing – Review & Editing-Support.

1

Detecting trends in flood series and shifts in flood timing across Kenya

Abstract

Floods constitutes one of the main causes of detrimental consequences arising from natural disasters, not only in Kenya, but across the globe. The frequency and magnitude of flood events in Kenya have increased over the past decade. Observations show a shift in timing and variability in flood occurrences in most parts of the country. Trend analysis is useful in detecting and supporting the evidence of change in flow series, as well as variability in flood timing. In this study, the frequency and magnitude of floods observed in peak over threshold (POT) and annual maximum (AMAX) flood series from 1981 to 2016 are compared in 19 Kenyan catchments. Flood peaks are identified using a threshold technique from Kenyan daily discharge data, and notable patterns in the AMAX series are compared to those in the POT series, which is created for three distinct exceedance criteria. The timing and variability of the annual floods is determined from the AMAX flow. Our findings show that, the AMAX series detects more trends in flood magnitude than the POT series, while the POT series detects more significant trends in flood frequency than flood magnitude. The sensitivity of the trend to the choice of exceedance threshold reveals differing trend patterns across the stations. Flood timing is in peak rainfall months of April, May and November and are less variable in most of the coastal and western stations, and highly variable in stations whose annual floods occur in dry months of June, July, and August. This information on flood characteristics can help to inform policy for disaster risk management, infrastructure design and agriculture and ultimately support improved livelihoods in Kenya.

Key words: peak over threshold, annual maximum flood, Kenya, flood magnitude, trend, flood frequency, flood timing.

1. Introduction

Flooding is among the most detrimental natural hazards worldwide (Berghuijs *et al.*, 2019) and with a changing climate there is an expected increase in flood risk globally (Arnell and Gosling, 2016; Liu *et al.*, 2018). For Kenya, floods are the most common climatic extreme and the leading hydro-meteorological disaster (Huho and Kosonei, 2014). There is a growing concern that major flooding events in many parts of Kenya in the past decade (Wainwright *at al* 2021;

Wanzala and Ogallo, 2020, Kilavi *et al.*, 2018; McLeod *et al.,* 2021) are indicative of the effects of a changing climate.

Understanding flood characteristics such as frequency, magnitude and timing is important for informing policy for disaster risk management, infrastructure design and agriculture, amongst other hydrological applications (Rosner, Vogel and Kirshen, 2014; Bezak, Brilly and Šraj, 2016). Such assessments require information on the probable year to year variations in flood characteristics (Parry *et al.*, 2007; Kundzewicz *et al.*, 2014). In addition, consideration of the trends in flood data series may result in more accurate flood timing, magnitude and frequency estimations (Berghuijs *et al.*, 2017, 2019; Mangini *et al.*, 2018; Sa'adi *et al.*, 2019). Trend analysis can be used to investigate whether there is any evidence of an increase in river floods in the observational river discharge data. Such analysis requires long records (e.g., more than 30 years) not only to explicitly distinguish climate variability from climate change induced trends (Svensson, Kundzewicz and Maurer, 2005; Vogel, Yaindl and Walter, 2011), but also incorporate the impacts of human induced activities such as deforestation and water management practices (e.g., reservoirs and irrigation).

Trend analysis on river flow series have been undertaken at global and regional scales in many parts of the world. For example, Svensson, Kundzewicz and Maurer, 2005; Cunderlik and Ouarda, 2009; Burn, Whitfield and Sharif, 2016; Vormoor *et al.*, 2016; Berghuijs *et al.*, 2017; Mangini *et al.*, 2018; Paprotny *et al.*, 2018; Ávila *et al.*, 2019; Ishak and Rahman, 2019; Zadeh, Burn and O'Brien, 2020. However, there is overgeneralization of trend patterns when considering a larger spatial extent, thus the need for trend analysis at smaller scales e.g., country scale (see Wilcox *et al.*, 2018; Giuntoli, Renard and Lang, 2019; Tramblay *et al.*, 2019). Relatively few studies have undertaken river flow trend analysis in Africa (e.g., Nka *et al.*, 2015; Diop *et al.*, 2018; Degefu *et al.*, 2019), mainly due to data quality issues which may affect trend detection (Slater and Villarini, 2017).

For Kenya, a few studies have attempted to quantify trends in streamflow (e.g., Mwangi *et al.*, 2016; Langat, Kumar and Koech, 2017; Cheruiyot, Gathuru and Koske, 2018). The highlighted studies looked at the trends and variability in rainfall in Kenya and did not look at the changes in the frequency and how this varies across different catchments in Kenya. Despite its practical significance, little is known about the temporal characteristics of streamflow and the highlighted studies considered only trends in the annual maximum flow and only for one catchment and little has been done to quantify trends in both annual floods and peak over

3

threshold (frequency and magnitude). Those studies which have incorporated the frequency and magnitude of floods were focused on single river basins, incorporating only a single station e.g., Tana River (Langat, Kumar and Koech, 2017), Malewa river (Nyokabi, Wambua and Okwany, 2021), Naivasha (Kyambia and Mutua, 2015). These studies showed that flow series for these stations had a statistically significant upward monotonic trend and seasonal variability, indicating that the streamflow regime had changed significantly.

Flood trend analysis looks at trends in the annual maximum river discharge (AMAX), i.e., a one value per year flood series (Kundzewicz *et al*. 2004; 2005). The advantage of this strategy is that the events chosen in two consecutive years are independent. However, the AMAX technique ignores flood occurrences that are less than the annual maximum in each year but are nevertheless significant for society, particularly in terms of losses, and perhaps inappropriate for climates with two distinct rainy seasons. The Peak Over Threshold (POT) technique (Burn, Whitfield and Sharif, 2016; Mangini *et al.*, 2018) selects all floods over a specific threshold that occur throughout a flow record. This allows for a trend in the frequency (counts) of floods rather than merely their magnitude to be estimated (Svensson, Kundzewicz and Maurer, 2005). Studies looking into trends in POT series do not exist in Kenya even though there has been an increase in the frequency of reported flood events across the country.

Understanding variability in flood timing and seasonality is important for water resources planning and management (Stephens *et al.*, 2015). Changes in the timing of the yearly flood have far-reaching consequences for flood-based farming systems especially to the livelihoods of people who adjust their floodplain management and agricultural activities to the rise and fall of the flood wave (Ficchì and Stephens, 2019). Significant changes in flood timing have been found around the world as shown in studies by (Cunderlik and Ouarda, 2009; Burn, Whitfield and Sharif, 2016; Blöschl *et al.*, 2017; Ficchì and Stephens, 2019). There has been an overall shift in flood timing in East Africa in recent decades (Ficchì and Stephens, 2019) but there has been no detailed study for catchments across Kenya.

Hence, there is a notable gap in the literature associated with temporal characteristics of streamflow data in terms of frequency and magnitude in both AMAX and POT as well as the country scale flood seasonality and timing in Kenya, and this gap was an important motivation for the present study. The objective of this paper is to detect the evidence of statistically significant flood trends across Kenya, where flood-based farming systems and livelihoods are extensive. High flow indices are derived from river flow discharge series: AMAX, and POT

4

indices using different magnitude and frequency thresholds. First high flow indices were derived and subjected to a trend test. Then a sensitivity analysis of the detected trends was performed for different flood peak selection criteria. Next the country-scale seasonality in flooding was characterised. Last, the changes in the timing of the annual floods. The following three research questions are addressed:

(1) What are the trends in flood magnitude and frequency across Kenya for the period 1981–2016?

(2) What is the sensitivity of the detected trends to the selection criterion used to define different flood peak series?

(3) What are the characteristics of country scale seasonality in flooding?

(4) What are the changes in seasonality and flood timing of the annual floods across Kenya?

2. Study area

The study is undertaken across Kenyan catchments at 19 river gauging stations (**Error! Reference source not found.**) with varying characteristics (Table 1). These were selected due to the frequency and magnitude of the impacts of floods, as well as the availability of river flow observations (Table 1).

Kenya exhibits high variability in physiographic and hydroclimatic conditions. The highest point is at about 5000m a.s.l, (mostly areas around central highlands) while the lowest point is about 20m a.s.l. (mainly around coastal areas). The vegetation cover is mainly a mixed tree cover, grass, and sparse vegetation in most of parts of the country and shrubs and bare land in the arid and semi-arid areas of northern Kenya. As a result, Kenya experiences different climate related extremes in terms of intensity, magnitude, and timing.

The rainfall pattern follows a bimodal rainfall seasonality (Ongoma and Chen, 2017) with high spatiotemporal variability (Hession & Moore, 2011). Three seasons are experienced: the 'long rains' of March–April - May (MAM), non-rainy months of June – July - August (JJA) and, the 'short rains' of October - November - December (OND) (Ogallo, 1988; Ongoma *et al.*, 2015). About 42% of the total annual rainfall is observed during MAM rainfall season (Ongoma and Chen, 2017), with the highest intensity observed near the water bodies of the Indian Ocean, Lake Victoria, and the Kenyan highlands.

3. Data and methods

3.1. Observed discharge data

For the study, daily river flow data from 19 sites distributed across the country were employed (Table 1, Figure 1). A common data period of 1981 – 2016 was adopted for the analysis. Two of the stations (Ndo and Perkerra) did not have the most recent data records and only 13 and 21 years of data were available respectively, which may affect the shifts in the trend patterns in the two stations.

Six distinct high flow indices were derived from the daily mean river discharge data for the 19 stations, and then subjected to the Mann-Kendall test to check for trends. The magnitude of the trend slope was also determined for each index for each station. The discharge data used in this study does not incorporate in-depth metadata, thus the trends may include the effects of changes in flow attenuation (e.g, reservoirs) and land use as well as any impacts of climate change.

3.2. Flow indices

The features of high flow regimes were described using six distinct indices. First, the yearly maximum daily mean river flow (AMAX) was derived. The second index is the Peak-over threshold (POT). Three peak magnitudes were tested as thresholds with the size of the criterion established so that, on average, one, three and five POT occurrences were chosen every year (Mangini *et al*., 2019): POT1*mag;* POT3*mag* and POT5*mag*. If the peaks in a POT series were separated by a specific time, they were deemed independent of each other. After an inspection of the flow series, a declustering method, which uses a 7 day separation time interval between the peak was applied as most of the catchments are less than $40,000 \text{km}^2$ (Collet, Beevers and Prudhomme, 2017) and thus a concentration time of less than 7 days. This separation time generally allow for the flow to recede appreciably between peaks thus selecting a lower separation window period for smaller catchments is highly recommended. Counting the number of POTs that occur each year can be used to quantify flood frequency (annual counts), and thus three flood frequency indices were derived corresponding to the POT magnitudes: POT1*freq*, POT3*freq* and POT5*freq*.The three POT1, POT3 and POT5 series represent the magnitude and frequency of the most severe, moderate, and minor floods respectively. In this work, the sensitivity of trend results to this threshold selection is evaluated and trends in the POT1 flood magnitude data are compared to those of AMAX.

6

A multiple index (MI) was adopted from Mangini *et al.,* (2019). Multiple Index (MI) shows the ratio of the magnitude of the mean discharge to the magnitude of maximum discharge of a flood series as well as the average yearly outflow for specific stations, which can be used to show differences in hydrological flood characteristics. Thus, useful in describing the varied flood regimes across Kenya's different hydroclimatic regions (e.g., wet, arid and semi-arid), as well as the effects of human influence such as dam construction and irrigation activities on the catchments' water balance. The MI is the ratio of the mean discharge of a flood series to the mean annual flow at each individual station.

$$MI = Q_F/Q_A \tag{1}$$

Q_F represents the mean discharge value of the flow series and Q_A is the mean discharge recorded for an individual station. A significant divergence from the mean flow is indicated by higher MI values and for smaller exceedance thresholds, low MI values would be obtained.

3.3.   *Trend detection in AMAX and POT series*

The non-parametric Mann Kendall's (MK) test (Mann, 1945; Kendall, 1975) was used to detect trends in AMAX and POT*mag* flood series. The modified MK test was applied to test for trend which incorporates Yue and Wang's (2004) variance correction approach. The effective sample size was calculated using serial correlation coefficients for all lags and the slope magnitude was estimated using the Thiel-Sen slope algorithm (Sen, 1968).

Trends in POT*freq* were estimated using the Chi-Squared test with parametric Poisson regression because unlike the non-parametric Mann Kendall test, it accounts for the hierarchical count series that may contain several paired values (Vormoor *et al.*, 2016; Mangini *et al.*, 2018). A two- tailed trend test was applied at 10% significance level to test for the statistically significant trends. This was required because the direction of the trends to be tested was unknown.

3.4 Sensitivity analysis of trends to threshold selection

The number of peaks considered in each flood series are affected by the exceedance threshold (λ), where POT1 ($\lambda =1$), yields the highest floods recorded at each of the stations and POTn ($\lambda =n$) yields n flood events in the series. This means that selecting a higher λ would result in a lower threshold thus yielding higher number of flood peaks in the series. Different thresholds ($\lambda =1$, 3 and 5) corresponding to POT1, POT3 POT5 flood series were derived and sensitivity analysis of trends to different POT series to the selection of different thresholds (λ) analysed.

7

Threshold selection was aided by the creation of different plots such as mean residual life plot that can aid in the determination of a suitable threshold level (Burn, Whitfield and Sharif, 2016). Mean residual life plot is a plot of the mean flood excess over a given threshold versus a range of threshold values.

3.5. *Flood seasonality*

Seasonality measures (Parajka *et al.*, 2010) are used to characterize the timing and variability of the extreme flood events. These are defined by directional statistics (Mardia, 1975). The date of occurrence of a flood event is defined as directional statistic through conversion of the Julian date of the occurrence of an event to an angular value (Berghuijs *et al.*, 2019; Ficchì and Stephens, 2019), where January 1 is Day 1 and December 31 is Day 365 of the flood occurrence for the event *i* following Eq. (2).

$$\theta_i = \left(JulianDate \right)_i \left(\frac{2\pi}{365} \right)$$

(2)

Where θ_i is the angular value (in radians) for the flood date for flood event *i* , following an interpretation that a flood date is a vector with a unit magnitude and a direction given by θ_i . For a given sample of n flood events, the mean flood date can be given by the x- and y - coordinates using Equations (2) and (4).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{i} \cos \left(\theta_i \right)$$

(3)

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{i} \sin \left(\theta_i \right)$$

(4)

Where $\bar{x}$ and $\bar{y}$ represent the x- and y-coordinates of the mean flood date and lie within, or on, the unit circle. To obtain the mean direction ($\bar{\theta}$) of the flood dates, Eq. (5) was used,

$$\bar{\theta} = \tan^{-1} \left(\frac{\bar{y}}{\bar{x}} \right)$$

(5)

Where $\bar{\theta}$ can be converted back to a day of the year using Eq. (6):

$$MD = \bar{y} \frac{365}{2\pi}$$

(6)

8

The measure of the average time of occurrence of flood events for a given catchment is represented by the variable MD, which is expected to be similar for catchments with similar hydrologic characteristics (such as size and location).

To determine a measure of variability of the n flood occurrence about the mean date, a mean resultant is defined using Equation (7).

$$\bar{r} = \sqrt{\bar{x}^2 + \bar{y}^2}$$

(7)

Where the dimensionless measure of the data spread is defined by $\bar{r}$ and may assume values from 0 to 1. A value close to 1 is an indication that all floods in a given sample occurred on the same day of the year, while values close to zero point to a higher variability in the date of occurrence of flood events for a catchment. Which follows that the higher values of $\bar{r}$ are associated with higher regularity in the timing of flood events (higher predictability). More details of directional statistics can be found in Burn (1997), and this method was adopted for this analysis because it has been used in several flood seasonality studies across different parts of the world e.g., (Berghuijs *et al.*, 2019; Ficchì and Stephens, 2019).

4. Results

4.1 Frequency of peak events in the flood series

The mean number of discharge peaks per year varies spatially across Kenyan catchments from about 4 to 13 events (Figure 2). The highest number of flood events per year are recorded at Perkerra, Mara and Garsen stations. The notably high number of events at Perkerra may be due to the shorter data series (large percentage of missing data as outlined in Table 1). However, the Garsen and Mara stations have relatively good data records and are mainly arid and semi-arid, therefore we can be more confident that the high numbers of events are likely to be due to the sporadic torrential rains that occur during the rainy months. The mean number of flood events per year for the other stations ranges between 6-8.

4.2 Trends in AMAX and POT1 flood series

There is considerable heterogeneity in the trend results calculated for the 36-year research period (Figure 3). However, positive trends are dominant across 12 stations in the AMAX flood series across Kenyan and only 3 stations (Nyangores, Ewaso-Ngiro and Gucha-Migori) show negative trends.

9

Statistically significant positive and negative trends at 10% confidence interval in the AMAX series are detected in 12 out of 19 stations (Figure 4). The significant positive trends are dominant at Garissa, Garsen, Mutonga, Athi-Munyu, Athi-Tsavo, Sondu, Perkerra, Sio and Ruambwa stations. The significant negative trends are dominant at Gucha-Migori, Nyangores and Ewaso-Ngiro stations. Only three stations (Perkerra, Ewaso-Narok and Ndo) do not show trends in the AMAX series. Interestingly, Pekerra has the highest mean number of independent discharge peaks but no visible trends in the peaks. However, in general trends in the same direction can be seen for stations in proximity to each other which is evidence of a spatially coherent pattern.

Positive trends in POT1 flood magnitude are found for 11 stations and negative trends for 8 stations respectively (Figure 5 right). For flood frequency only three stations (LargeNzioa, Munyu and Migori) show negative trends for POT1 (Figure 5 left), whereas the rest of the stations show positive trends.Garsen, Yala, Narok, Perkerra, Miriu, Sio, Nyangores and Garissa stations show increasing positive trends, and Thiba, Munyu, and LargeNzioa show decreasing trends both in flood frequency and magnitude.

The total number of stations exhibiting statistically significant trends is higher in the POT1mag than in AMAX (Figure 6). There is a consistency pattern in the flood change in some catchments depicting significantly increasing trends in both the MAX and POT1mag series: Garissa, Athi- Munyu, Athi- Tsavo, Ruambwa, Sio, Mutonga etc. Only one station (Ewaso – Ngiro) has a significantly decreasing trend both in the AMAX and POT1mag flood series, whereas Gucha-Migori has significantly increasing and decreasing trend in the AMAX and POT1mag respectively.

POT1*freq* flood series show statistically significant trends in 8 out of the 19 stations, predominantly in stations in the western part of the country. This suggests a clear pattern in the trends in flood frequency being more significant in the western parts of Kenya than in the rest of the country, with the exception of the Garissa and Garsen stations.

4.3 Sensitivity of flood trends to the selection of different flood exceedances thresholds

There are notable differences in trends in some catchments for the AMAX and POT1mag flood series as seen in Figures 4 and 6, which points to the effects of selecting exceedance thresholds in deriving POT series. Higher exceedance threshold ($\lambda = 1$) gives the most extreme flood events, with an average of one event pet year, whereas lower exceedance threshold ($\lambda = 5$) gives a lower threshold, thus an increase in the number of (smaller) peak events (Figure 7).

10

MIs from the AMAX are the most significant, followed by the MI of the POT1mag series as well as the remaining POT series, whereas smallest MIs are derived for the POT6mag flood series (Figure 8a).MI values are relatively higher in the small sized arid and the semi-arid catchments such as Ewaso-Ngiro, Ewaso-Narok, Perkerra and Mara (Figure 8b). The arid and semi-arid climate of these catchments causes generally low mean discharge across the year which is offset by intense precipitation events leading to high discharge values. Additionally, small catchments, are likely to show an intrinsic high variance in the daily hydrograph, thus the high MI values. Mutonga and Thiba show highest MI values. This could be attributed to the large dam releases into the rivers resulting in higher mean discharge, especially during rainy months. In general, catchments in the western part of the country have medium MI values due to rain falling all year round. This is because rainfall falling out of typical seasonal rainfall contributes substantially to the mean annual rainfall.

There is a high sensitivity of the results of the trend analysis to different exceedance thresholds in terms of both magnitude and frequency (Figure 9). Half of the stations show significant positive and negative trends for high λ. More than half of the stations show significant positive trends for $\lambda = 3$ and $\lambda = 4$. However, there no defined pattern in sensitivities (increasing with increase in λ or decreasing with a decrease in across most of the stations for flood magnitude (Figure 9a)

Trend results show high sensitivity to thresholds for flood frequency series as seen in Figure 9b. There are notable significant trends detected in the POTfreq series across most of the stations for increasing value of λ. Half of the stations depict significant increasing positive and negative trends with clear negative trends in Ewaso-Narok, Perkerra and Gucha-Migori. for threshold values of $\lambda = 3$ and $\lambda = 5$. There is clear pattern of significantly increasing trends in flood frequency across all the stations.

4.4 Flood timing and variability

The occurrence of the annual flood in most of the stations is around the months of March, April, May, November, and December (Figure 11 panel a). These coincides with the occurrence of the 'long rains' (March – May) and the 'short rains' (October – December). Some stations such as Ndo, Perkerra, Mara, Ewaso- Ngiro and Ewaso-Narok, have a mean date of occurrence around the months of June, July, and August.

Predictability is high (indicated by values of r > 0.4) in most of the stations whose annual floods occur during the short and long rains, whereas values of r < 0.4 can be seen in the stations in

11

which the flood timing is around the dry months of June, July, and August (Figure 11 panel b). There is a consistent spatial pattern in the predictability; stations in the western (Nzio, Sio, Yala, LargeNzioa), central (Thiba, Mutonga, Munyu), and coastal regions (Tsavo, Garsen, Garissa) show a high tendency in the predictability, whereas stations in the Rift valley region (Mara, Amala, Nyongores, Ndo and Perkerra), have low tendency in predictability.

5. Discussion

5.1. *Trends in AMAX and POT1 series*

In this study, statistically significant trends are detected in the AMAX and POT flow series in most of the 19 stations across the country, with only three showing significantly decreasing trends in the AMAX. The decreasing trends at Nyangores, Ewaso Ngiro and Migori stations are mainly due shorter data records of the flow series used in this study. The available data does not include the most recent years as compared to other stations and this may fail to capture the most recent floods in those stations. The trends are consistent spatially for stations on Tana River (Thiba, Garissa and Garsen), Athi (Tsavo, Munyu) and Nzioa (Ruambwa). This may be attributed to the increasing frequency of rainfall events in these parts of the country (Huho and Kosonei, 2014; Wainwright *et al.*, 2021).

The significant trend in both AMAX and POT1*mag* varies from one station to another. For instance, there is a general tendency of increasing trend in flood frequency and magnitude at Garissa, Athi- Munyu, Athi- Tsavo, Ruambwa, Sio, Mutonga, while Gucha-Migori, Nyangores and Ewaso-Ngiro show decreasing trends. Trend patterns at Garissa, Ruambwa and Sio agree to the findings of Nasambu (no date; Langat *et al.*, 2020) et al., (2018); Langat et al (2020), who found significant increasing trends due to increased frequency of rainfall events at these stations. However, the above results also point to a mixed pattern in trends in stations that may be from the same region. These trends patterns are important in indicating the existence of different flood drivers in those regions. For example, Ruambwa and Gucha are both located in western part of Kenya, which receive rainfall all year round, but the two have contrasting trend patterns. Thus, there is a lot of variability within the trends and their significance, and there is often a difference between the magnitude and the frequency. This finding is like studies from around the world underlining the complexity of making regional generalisations about trends such as Mangini *et al*., (2019) and Arheimer and Lindström (2015). Garsen station at the downstream point on Tana River has a significantly increasing trend in flood frequency but not in magnitude, whereas Mara and Amala have decreasing trends in two indices. This may be

12

mainly because rivers in semi-arid environments, such as the Tana and Mara, are particularly vulnerable to fluctuations in water availability caused by decreases in rainfall, increases in water withdrawals, and changes in seasonal flows (Langat et al 2020). Tana for example, is highly affected by human activities such as irrigation schemes, domestic consumption, and hydropower generation (Langat, Kumar and Koech, 2017; Langat *et al.*, 2019).

In addition, both Tana and Mara rivers are classified as water limited catchments with a dryness index of 1.1 annually. This means that there is imbalance in water demand and any increase in the rainfall is accompanied by higher temperatures resulting in high evapotranspiration, thus increasing the dryness index further. Therefore, most of the rainfall in these catchments is mostly lost to evapotranspiration (Mwangi *et al.*, 2016).

5.2.   *Sensitivity of trends to different POT threshold selection*

Flood frequency trends can be seen across all thresholds whereas for POT-M series, with more than one event per year on average, the percentage of stations exhibiting significant trends in flood magnitude is high, and only a few significant positive trends are detected for $\lambda = 1$. Sensitivity of trends to different POT series when different thresholds are selected show different trend patterns and varies from one station to another. Generally, there is no clear pattern in trends in POT*mag* series in response to different thresholds. This is because the flood series is highly sensitive to the threshold selection and varies considerably from one station to another. This may also be alluded to non-homogeneity in flood characteristics at different stations across Kenya.

When we considered smaller floods (POT5), we observed a general increasing trend in flood magnitude for stations in western parts, while a clear pattern toward decreasing flood magnitude in the southern and coastal stations. This is because western stations are characterised by less intense rainfall spread across all year round, unlike ethe coastal and southern regions which receive rains mainly in rainy months (March, April, May, October, November) and mostly dry for the rest of the months. These decreasing trends are pronounced at Tsavo and Mara station. The overall spatial pattern of decreasing flood trends in the POT3 series is same to studies in other regions such as to Mediero *et al.,* (2015).

5.3.   *Flood timing and variability*

The timing of the annual flood is assessed using the AMAX flow for the 19 stations across the country. Results show flood timing occurs in the peak rainfall months of April and May for the 'long rains' (March – April – May) and October and November for the 'short rains mostly for

13

western, central highlands and the coastal stations. Our findings are like those in Stephens *et al.*, (2015), which showed some extent of correlation of precipitation and floodiness in East Africa by comparing the anomalies in precipitation and floodiness on monthly basis. However, some stations lying on the same river, show varied timings of floods in different months, but the reasons for this are not clear. For example, Tana River and Athi river in the coastal area and Nzioa river in the western areas. The upstream gauging stations on Tana River (Mutonga and Garissa) show the occurrence of annual floods in the months in November and December. The downstream gauging stations (Garsen and Hola) show this occurrence in April and May. The reverse is observed for stations located along the Athi river. One possible reason for this observation may be alluded to the high predictability of the short rains (October-November-December) (Kilavi *et al.*, 2018), which is captured in the daily hydrograph of these downstream stations. However, more research is needed to determine the extent of this nonlinearity at different temporal or spatial scales, considering the role of different precipitation periods in flood generation in different regions (Froidevaux *et al.*, 2015). Also, there are stations (e.g., Mara, Amala, Ewaso-Ngiro, Perkerra) showing the occurrence of floods in the dry months of July and August. These stations are in arid and semi-arid areas of the country and are mainly characterised by low flows and any sporadic rainfall falling in those offseason months may lead to a rise in flow, thus depicted in the AMAX index. This implies increased likelihood of flooding events both in rainy and non-rainy seasons, which may result in upward trends in flood frequency in these stations.

There is a high tendency of flood predictability in rivers that flood during peak rainfall months in Kenya and the reverse is true. For example, flood predictability is high for the stations (Ruambwa, Garissa, Garsen, Munyu) whose flood timing occurs in peak rainfall months and low for those whose floods occur in dry months (Perkerra, Ndo, Ewaso-Ngiro, Mara, Amala). The results also show defined regional spatial pattern in flood predictability across the country. For example, there is correspondence in predictability for stations located in the western and costal stations, which are characterised by the annual floods occurring in the for peak rainfall months. Whereas for the stations located in the Rift valley region, there is a clustered pattern for annual floods occurring in dry months.

6. Conclusion

Our efforts to identify regional and basin scale trend signals yielded no spatial flood trend pattern especially for the POT*mag* flood series. In every catchment in Kenya, there emerged mixed trends for most of the flood indices used. Mixed and inverse trend signals were also

14

observed between adjacent gauging stations in the same and neighbouring river catchments. The observed complex trend signals between adjust streamflow stations may be attributed to the presence of very complex climate, topography, land cover, and land use systems in different parts of the country, which showed greater variation with short horizontal distances.

With a few exceptions, such as the trends in flood frequency, this study shows that the number of catchments exhibiting significant trends differ across the country and are not consistent across all flood series. Also, the trend significance in both AMAX and POT1*mag* vary from one station to another. For this reason, it would be more informative to consider trend analysis within larger scale hydro-climatic regions, because trend signal within a region can be considered less sensitive. Additionally, There is need therefore to understand such trends, but this requires careful identification of triggers and hydrological processes (Slater and Wilby, 2017; Berghuijs *et al.*, 2019). This is useful in unveiling the degree of this nonlinearity over different temporal and spatial scales, whilst considering the influence of the role of different precipitation periods for flood generation in different regions across Kenya. However, such studies require more reliable data which is currently lacking in Kenya. For example, the cause of flood trends in Kenya is still unclear due to a limited understanding of regional variations in flood-generating mechanisms e.g. in (Tramblay *et al.*, 2019), land- use changes, reservoir construction and other local effects (Svensson, Kundzewicz and Maurer, 2005) and the uncertainty in projections of future flooding under climate change (Kundzewicz *et al.*, 2014; Berghuijs *et al.*, 2019).

In conclusion, the presence of statistically significant trends in flood series across Kenyan catchments investigated in this study. In comparison to previous trend studies, three novel aspects are explored. First, significant trends detected across the country in two flood series: annual maximum (AMAX) and the peak over threshold (POT) are compared. Then a sensitivity analysis of trends in floods to the selection of different exceedance thresholds in the POT flood series is analysed. Finally, the timing and variability in the annual floods across the stations using the annual maximum flow is explored. This study supports the importance of analysis of trends at country level as it highlights key characteristics that may not be captured in regional or global analysis. The findings are key in providing information on flood characteristics to help inform policy for disaster risk management, infrastructure design and agriculture and ultimately support improved livelihoods in Kenya.

15

Acknowledgements

References

Arnell, N.W. and Gosling, S.N. (2016) 'The impacts of climate change on river flood risk at the global scale', *Climatic Change*, 134(3), pp. 387–401.

Ávila, Á., Guerrero, F.C., Escobar, Y.C. and Justino, F. (2019) 'Recent precipitation trends and floods in the Colombian Andes', *Water*, 11(2), p. 379.

Berghuijs, W.R., Aalbers, E.E., Larsen, J.R., Trancoso, R. and Woods, R.A. (2017) 'Recent changes in extreme floods across multiple continents', *Environmental Research Letters*, 12(11), p. 114035.

Berghuijs, W.R., Harrigan, S., Molnar, P., Slater, L.J. and Kirchner, J.W. (2019) 'The relative importance of different flood-generating mechanisms across Europe', *Water Resources Research*, 55(6), pp. 4582–4593.

Bezak, N., Brilly, M. and Šraj, M. (2016) 'Flood frequency analyses, statistical trends and seasonality analyses of discharge data: a case study of the L itija station on the S ava R iver', *Journal of Flood Risk Management*, 9(2), pp. 154–168.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R.A.P., Merz, B., Arheimer, B., Aronica, G.T., Bilibashi, A., Bonacci, O. and Borga, M. (2017) 'Changing climate shifts timing of European floods', *Science*, 357(6351), pp. 588–590.

Burn, D.H. (1997) 'Catchment similarity for regional flood frequency analysis using

16

seasonality measures', *Journal of hydrology*, 202(1–4), pp. 212–230.

Burn, D.H., Whitfield, P.H. and Sharif, M. (2016) 'Identification of changes in floods and flood regimes in Canada using a peaks over threshold approach', *Hydrological Processes*, 30(18), pp. 3303–3314.

Cheruiyot, M., Gathuru, G. and Koske, J. (2018) 'Quantity and trends in streamflows of the Malewa River Basin, Kenya', *Journal of Environmental Science and Engineering B*, 7(1), pp. 1–11.

Collet, L., Beevers, L. and Prudhomme, C. (2017) 'Assessing the impact of climate change and extreme value uncertainty to extreme flows across Great Britain', *Water*, 9(2), p. 103.

Cunderlik, J.M. and Ouarda, T.B.M.J. (2009) 'Trends in the timing and magnitude of floods in Canada', *Journal of hydrology*, 375(3–4), pp. 471–480.

Degefu, M.A., Alamirew, T., Zeleke, G. and Bewket, W. (2019) 'Detection of trends in hydrological extremes for Ethiopian watersheds, 1975–2010', *Regional Environmental Change*, 19(7), pp. 1923–1933. Available at: https://doi.org/10.1007/s10113-019-01510-x.

Diop, L., Yaseen, Z.M., Bodian, A., Djaman, K. and Brown, L. (2018) 'Trend analysis of streamflow with different time scales: a case study of the upper Senegal River', *ISH Journal of Hydraulic Engineering*, 24(1), pp. 105–114.

Ficchì, A. and Stephens, L. (2019) 'Climate variability alters flood timing across Africa', *Geophysical Research Letters*, 46(15), pp. 8809–8819.

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C. and Martius, O. (2015) 'Flood triggering in Switzerland: the role of daily to monthly preceding precipitation', *Hydrology and Earth System Sciences*, 19(9), pp. 3903–3924.

Giuntoli, I., Renard, B. And Lang, M. (2019) '10 Floods in France', *Changes in Flood Risk in Europe* [Preprint].

Hession, S.L. and Moore, N. (2011) 'A spatial regression analysis of the influence of topography on monthly rainfall in East Africa', *International journal of climatology*, 31(10), pp. 1440–1456.

Huho, J.M. and Kosonei, R.C. (2014) 'Understanding extreme climatic events for economic development in Kenya', *IOSR Journal of Environmental Science, Toxicology and Food Technology*, 8(2), pp. 14–24. Available at: https://doi.org/10.9790/2402-08211424.

Ishak, E. and Rahman, A. (2019) 'Examination of changes in flood data in Australia', *Water*, 11(8), p. 1734.

Kendall, M. (1975) 'Rank correlation methods, book series, Charles Griffin'. Oxford University Press, USA, London.

Kilavi, M., MacLeod, D., Ambani, M., Robbins, J., Dankers, R., Graham, R., Titley, H., Salih, A.A.M. and Todd, M.C. (2018). Extreme rainfall and flooding over central Kenya including Nairobi city during the long-rains season 2018: causes, predictability, and potential for early warning and actions. *Atmosphere*, **9** (12): 472. doi:10.3390/atmos9120472.

Kundzewicz, Z.W., Kanae, S., Seneviratne, S.I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L.M., Arnell, N. and Mach, K. (2014) 'Flood risk and climate change: global and regional perspectives', *Hydrological Sciences Journal*, 59(1), pp. 1–28.

17

Kyambia, M.M. and Mutua, B.M. (2015) 'Detection of trends in extreme streamflow due to climate variability in the Lake Naivasha basin, Kenya', *International Journal of River Basin Management*, 13(1), pp. 97–103.

Langat, P.K., Kumar, L. and Koech, R. (2017) 'Temporal variability and trends of rainfall and streamflow in Tana River Basin, Kenya', *Sustainability*, 9(11), p. 1963.

Langat, P.K., Kumar, L., Koech, R. and Ghosh, M.K. (2019) 'Hydro-morphological characteristics using flow duration curve, historical data and remote sensing: effects of land use and climate', *Water*, 11(2), p. 309.

Langat, P.K., Kumar, L., Koech, R. and Ghosh, M.K. (2020) 'Characterisation of channel morphological pattern changes and flood corridor dynamics of the tropical Tana River fluvial systems, Kenya', *Journal of African Earth Sciences*, 163, p. 103748.

Liu, H., Wang, Y., Zhang, C., Chen, A.S. and Fu, G. (2018) 'Assessing real options in urban surface water flood risk management under climate change', *Natural Hazards*, 94(1), pp. 1–18.

MacLeod, D.A., Dankers, R., Graham, R., Guigma, K., Jenkins, L., Todd, M.C., Kiptum, A., Kilavi, M., Njogu, A. and Mwangi, E. (2021). Drivers and subseasonal predictability of heavy rainfall in equatorial East Africa and relationship with flood risk. *Journal of Hydrometeorology*, **22** (4): 887-903.

Mangini, W., Viglione, A., Hall, J., Hundecha, Y., Ceola, S., Montanari, A., Rogger, M., Salinas, J.L., Borzì, I. and Parajka, J. (2018) 'Detection of trends in magnitude and frequency of flood peaks across Europe', *Hydrological Sciences Journal*, 63(4), pp. 493–512. Available at: https://doi.org/10.1080/02626667.2018.1444766.

Mann, H.B. (1945) 'Nonparametric tests against trend', *Econometrica: Journal of the econometric society*, pp. 245–259.

Mardia, K.V. (1975) 'Statistics of directional data', *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(3), pp. 349–371.

Mediero, L., Kjeldsen, T.R., Macdonald, N., Kohnova, S., Merz, B., Vorogushyn, S., Wilson, D., Alburquerque, T., Blöschl, G. and Bogdanowicz, E. (2015) 'Identification of coherent flood regions across Europe by using the longest streamflow records', *Journal of Hydrology*, 528, pp. 341–360.

Mwangi, H.M., Julich, S., Patil, S.D., McDonald, M.A. and Feger, K.-H. (2016) 'Relative contribution of land use change and climate variability on discharge of upper Mara River, Kenya', *Journal of Hydrology: Regional Studies*, 5, pp. 244–260.

Nasambu*, O.B., Masinde, N.E. and Mwikali, K. V (no date) 'Examining the Geographical Coverage of Floods Using Satellite Images and Discharge In LNRB'.

Nka, B.N., Oudin, L., Karambiri, H., Paturel, J.-E. and Ribstein, P. (2015) 'Trends in floods in West Africa: Analysis based on 11 catchments in the region', *Hydrology and Earth System Sciences*, 19(11), pp. 4707–4719.

Nyokabi, E.W., Wambua, R.M. and Okwany, R.O. (2021) 'Assessment of rainfall, streamflow and reservoir level trends for Malewa River Catchment, Naivasha, Kenya', *Journal of Civil, Construction and Environmental Engineering*, 6(1), pp. 1–8.

Ogallo, L.J. (1988) 'Relationships between seasonal rainfall in East Africa and the Southern

18

Oscillation', *Journal of Climatology*, 8(1), pp. 31–43.

Ongoma, V., Guirong, T., Ogwang, B.A. and Ngarukiyimana, J.P. (2015) *Diagnosis of Seasonal Rainfall Variability over East Africa: A Case Study of 2010-2011 Drought over Kenya*, *Pakistan Journal of Meteorology*.

Ongoma, V. and Chen, H. (2017). Temporal and spatial variability of temperature and precipitation over East Africa from 1951 to 2010. *Meteorology and Atmospheric Physics*, **129** (2): 131–144.

Paprotny, D., Sebastian, A., Morales-Nápoles, O. and Jonkman, S.N. (2018) 'Trends in flood losses in Europe over the past 150 years', *Nature communications*, 9(1), p. 1985.

Parajka, J., Kohnová, S., Bálint, G., Barbuc, M., Borga, M., Claps, P., Cheval, S., Dumitrescu, A., Gaume, E. and Hlavčová, K. (2010) 'Seasonal characteristics of flood regimes across the Alpine–Carpathian range', *Journal of hydrology*, 394(1–2), pp. 78–89.

Parry, M.L., Canziani, O., Palutikof, J., Van der Linden, P. and Hanson, C. (2007) *Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC*. Cambridge University Press.

Rosner, A., Vogel, R.M. and Kirshen, P.H. (2014) 'A risk-based approach to flood management decisions in a nonstationary world', *Water Resources Research*, 50(3), pp. 1928–1942.

Sa'adi, Z., Shahid, S., Ismail, T., Chung, E.S. and Wang, X.J. (2019) 'Trends analysis of rainfall and rainfall extremes in Sarawak, Malaysia using modified Mann–Kendall test', *Meteorology and Atmospheric Physics*, 131(3), pp. 263–277. Available at: https://doi.org/10.1007/s00703-017-0564-3.

Sen, P.K. (1968) 'Estimates of the regression coefficient based on Kendall's tau', *Journal of the American statistical association*, 63(324), pp. 1379–1389.

Slater, L. and Villarini, G. (2017) 'On the impact of gaps on trend detection in extreme streamflow time series', *International Journal of Climatology*, 37(10), pp. 3976–3983.

Slater, L.J. and Wilby, R.L. (2017) 'Measuring the changing pulse of rivers', *Science*, 357(6351), p. 552.

Stephens, E., Day, J.J., Pappenberger, F. and Cloke, H. (2015) 'Precipitation and floodiness', *Geophysical Research Letters*, 42(23), pp. 10–316.

Svensson, C., Kundzewicz, Z.W. and Maurer, T. (2005) 'Trend detection in river flow series: 2. Flood and low-flow index series', *Hydrological Sciences Journal*, 50(5), pp. 811–824. Available at: https://doi.org/10.1623/hysj.2005.50.5.811.

Tramblay, Y., Mimeau, L., Neppel, L., Vinet, F. and Sauquet, E. (2019) 'Detection and attribution of flood trends in Mediterranean basins', *Hydrology and Earth System Sciences*, 23(11), pp. 4419–4431.

Vogel, R.M., Yaindl, C. and Walter, M. (2011) 'Nonstationarity: Flood magnification and recurrence reduction factors in the united states', *Journal of the American Water Resources Association*, 47(3), pp. 464–474. Available at: https://doi.org/10.1111/j.1752-1688.2011.00541.x.

Vormoor, K., Lawrence, D., Schlichting, L., Wilson, D. and Wong, W.K. (2016) 'Evidence for

19

changes in the magnitude and frequency of observed rainfall vs. snowmelt driven floods in Norway', *Journal of Hydrology*, 538, pp. 33–48.

Wainwright, C.M., Finney, D.L., Kilavi, M., Black, E. and Marsham, J.H. (2021) 'Extreme rainfall in East Africa, October 2019–January 2020 and context under future climate change', *Weather*, 76(1), pp. 26–31.

Wilcox, C., Vischel, T., Panthou, G., Bodian, A., Blanchet, J., Descroix, L., Quantin, G., Cassé, C., Tanimoun, B. and Kone, S. (2018) 'Trends in hydrological extremes in the Senegal and Niger Rivers', *Journal of Hydrology*, 566, pp. 531–545.

Yue, S. and Wang, C.Y. (2004) 'The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series', *Water resources management*, 18(3), pp. 201–218. Available at: https://doi.org/10.1023/B:WARM.0000043140.61082.60.

Zadeh, S.M., Burn, D.H. and O'Brien, N. (2020) 'Detection of trends in flood magnitude and frequency in Canada', *Journal of Hydrology: Regional Studies*, 28, p. 100673.

List of figures

(POT1), three (POT3) and five (POT5). (b) MI for POT1mag flood magnitude at the each of the 19 gauging stations across Kenya.

Figure 9: The sensitivity of trends in the POT series to the selection of different exceedances thresholds (λ) for different mean number of floods per year for POT1, POT3 and POT5 (a)magnitude and (b)frequency. Significant trends (expressed as percentages) in flood magnitude at 10% significance level with a threshold level λ range of 1, 3 and 5 mean events per year Different colors and symbols represent the different 19 stations considered in the study.

Figure 10:Annual flood seasonality measure for Kenya stream-gauge data. (a) Average flood timing (θ) and (b) Interannual variability of flood timing (r [−]) shown by different color saturation (the higher the variability is, the lighter the saturation is). The filled rectangular dots are reported at each of the 19 river gauging stations.

List of tables

Table 1: Characteristics of the Kenyan catchments studied, human influence and daily mean river flow gauge data availability information .Missing data is expressed as a percentage of the available period.

21

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
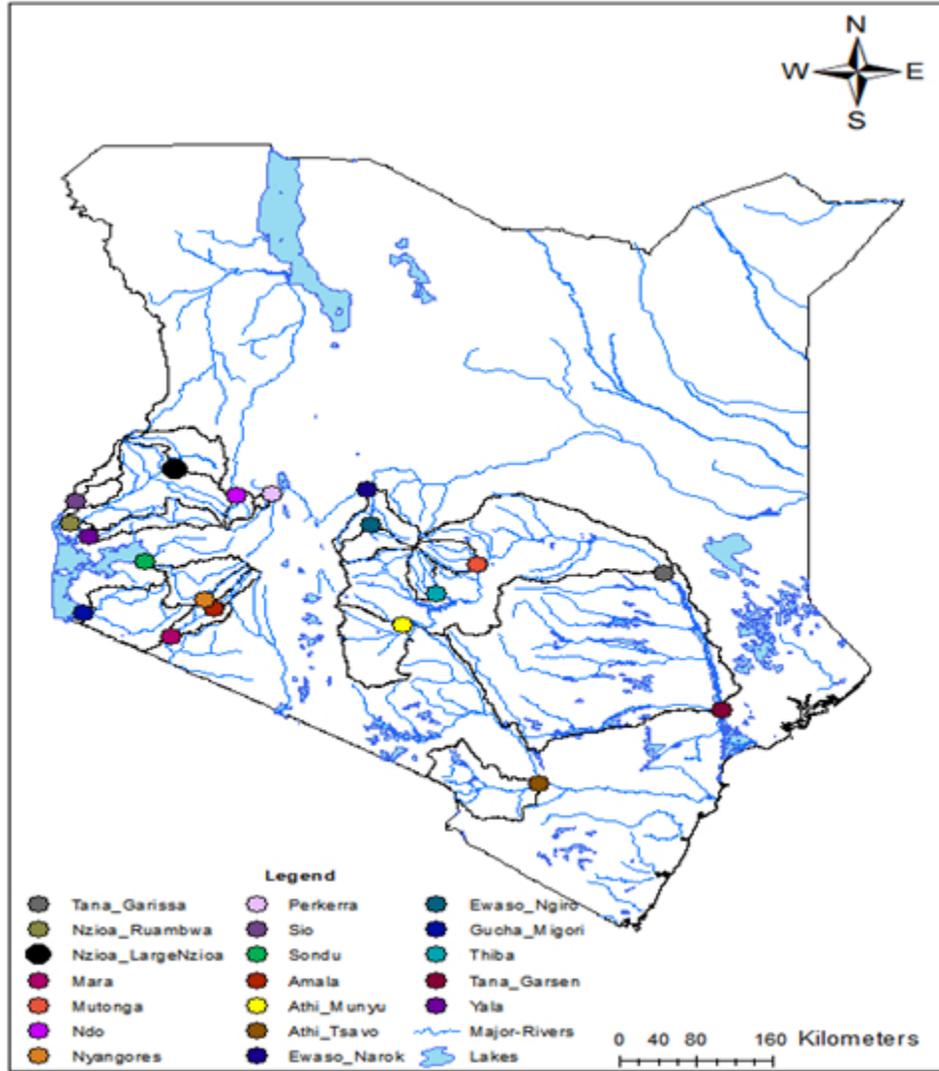
Figure 1: Location of the 19 river gauging stations located in the Kenyan catchments studied.

324x364mm (38 x 38 DPI)

41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
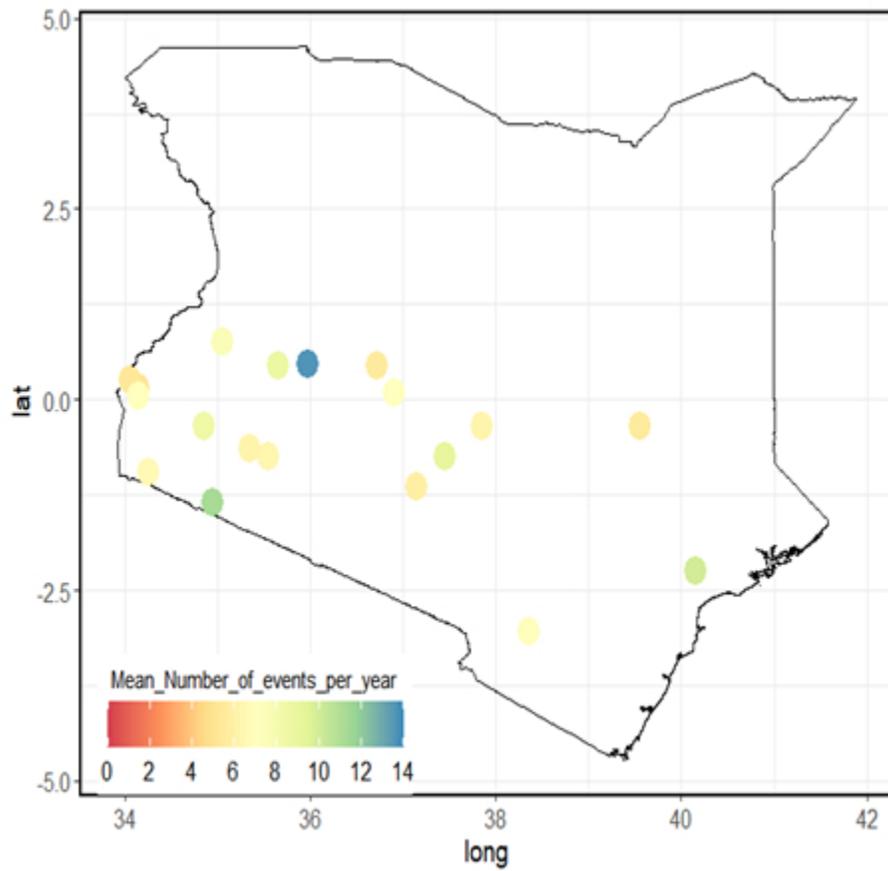
Figure 2: Mean number of independent flood peaks per year at each of the 19 gauging stations in Kenya.

302x295mm (38 x 38 DPI)
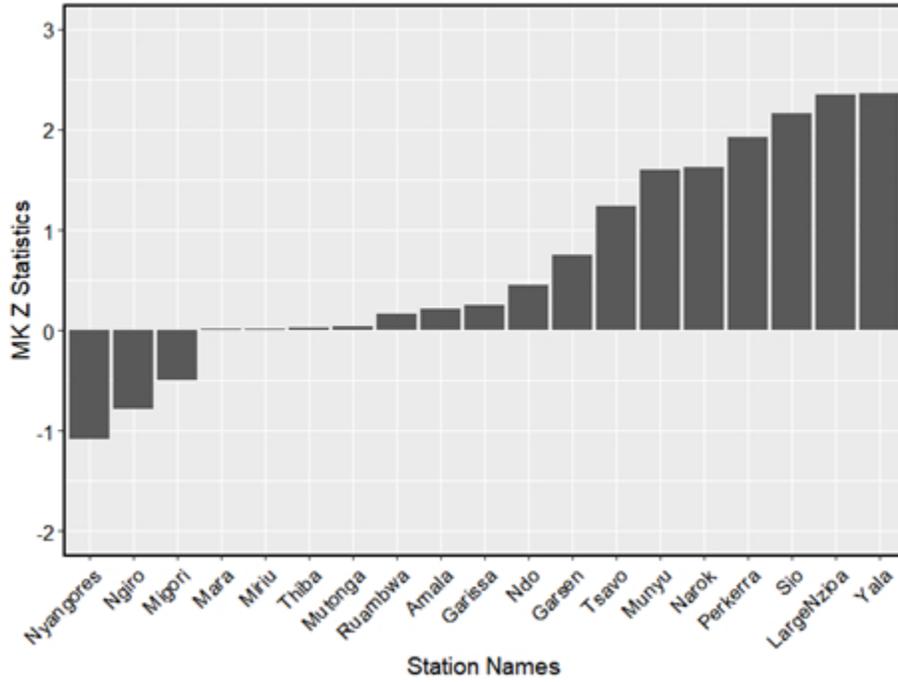
Figure 3: Mann–Kendall (MK) test Zs statistics in the AMAX discharge series across 19 Kenyan stations studied for the period 1981- 2016.

307x232mm (38 x 38 DPI)
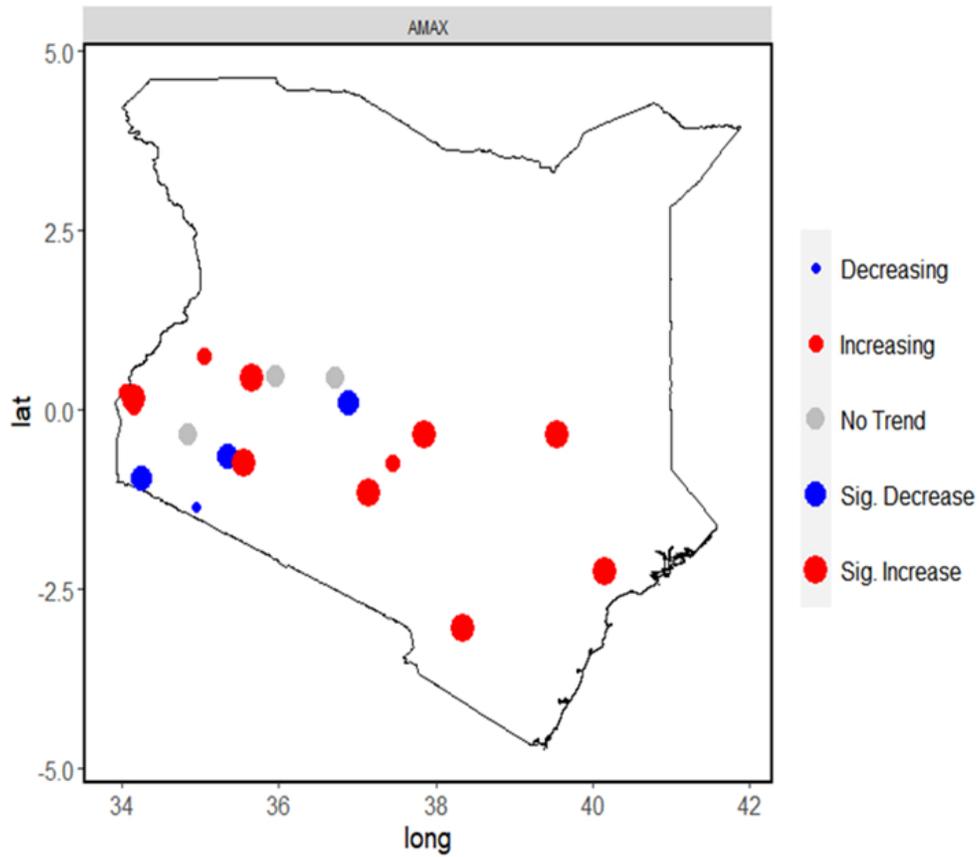
Figure 4: Trends in the annual maximum flood series (AMAX) for the period 1981 -2016. Filled circular symbols indicate the direction of the trend slope positive (red), negative (blue) and grey (no trend) trends at 10% significance level. The size of the circles indicates the statistically significant trends.

313x274mm (59 x 59 DPI)
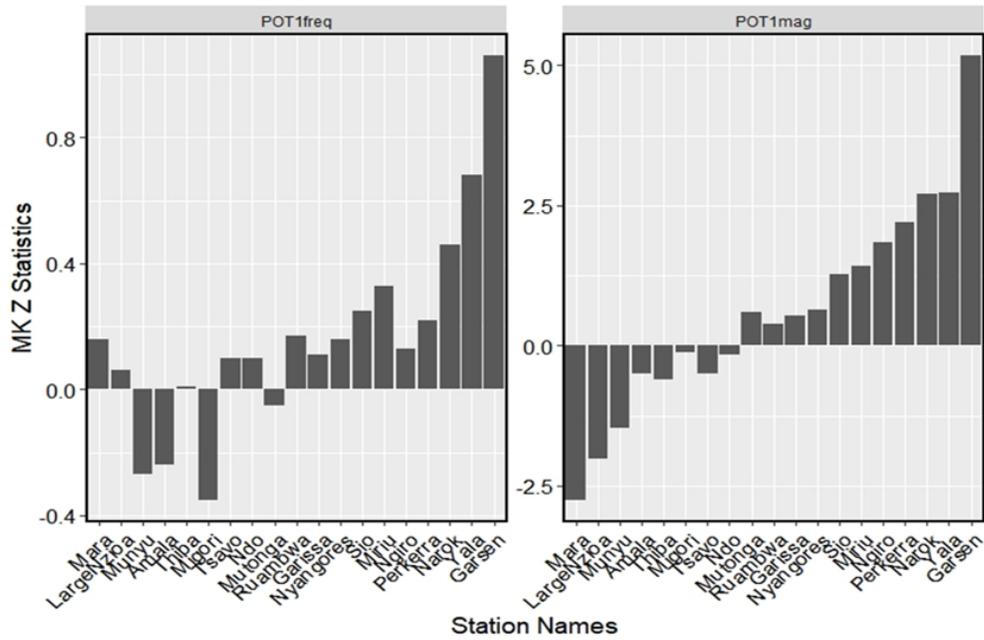
Figure 5: Mann–Kendall (MK) test Zs statistics in the POT1freq (left) and POT1mag (right) flood series across the 19 stations studies for the for the period 1981- 2016.

362x236mm (59 x 59 DPI)
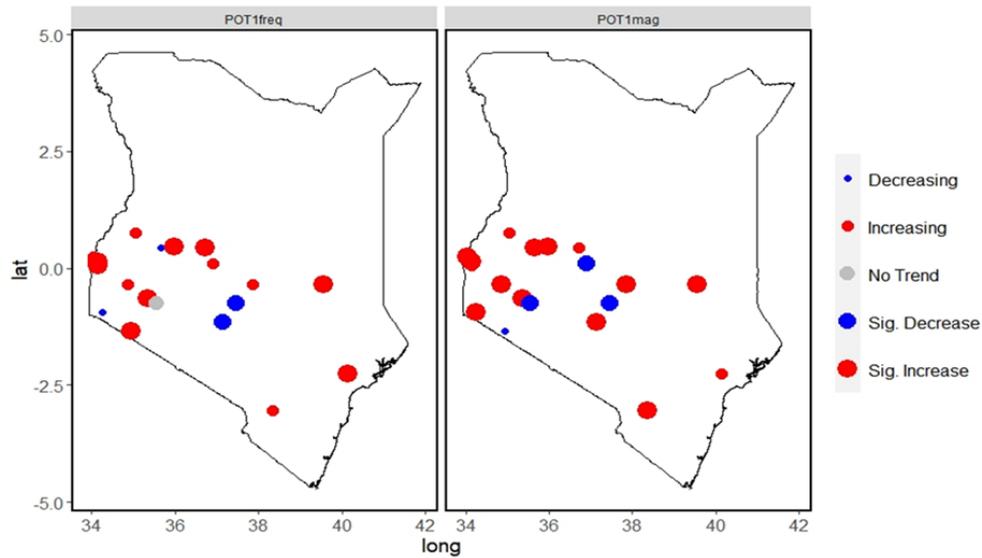
Figure 6:Trends in the POT1freq (left) and POT1mag (right) for the period 1981 -2016. Filled circular symbols indicate direction of the trend slope, significant positive (red), negative (blue) and grey (no trend) trends at 10% significance level. The size of the symbol indicates the statistically significant trends.

400x226mm (59 x 59 DPI)
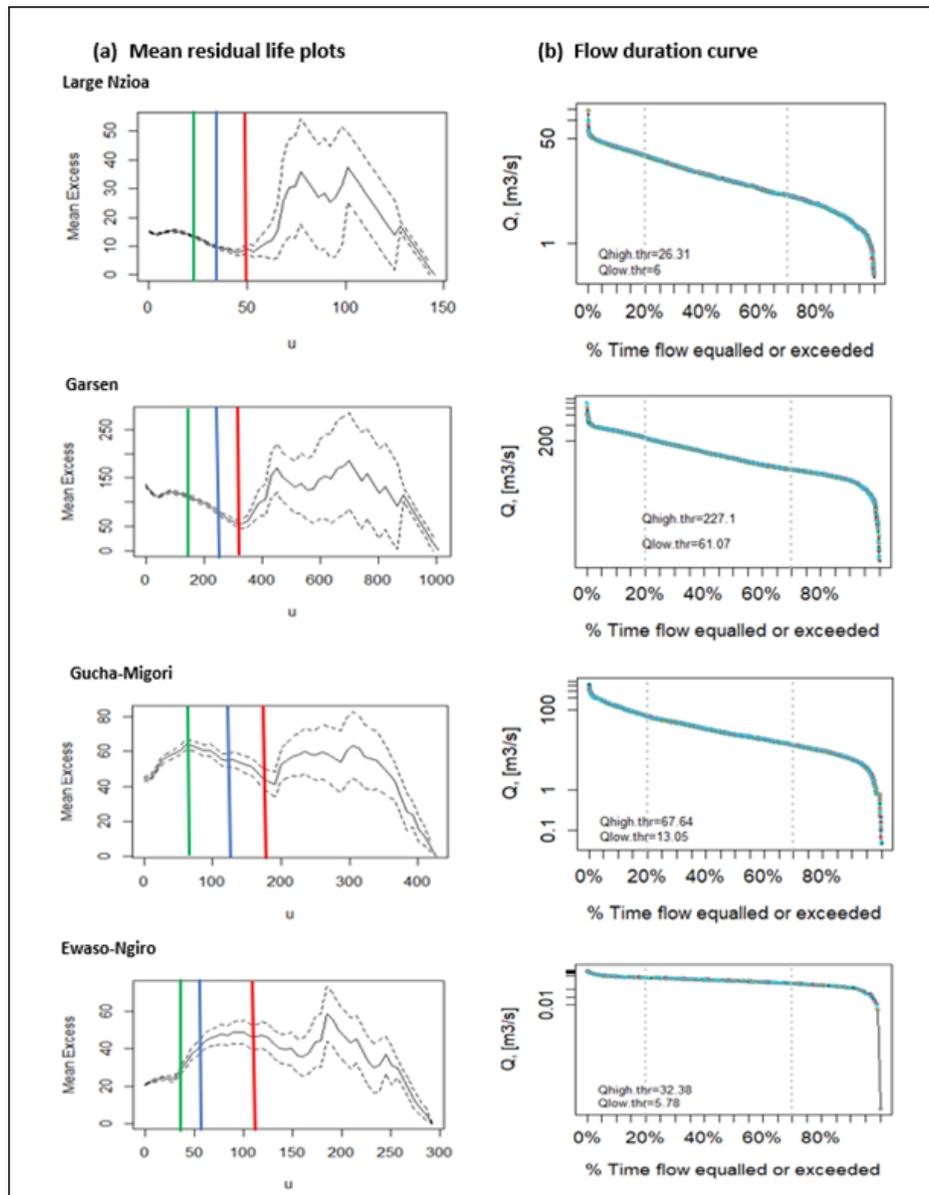
Figure 7: Visual plots for exceedance threshold selection using (a) mean residual life plots (left panel) and (b) flow duration curves for selected study stations. The vertical lines show different thresholds (λ =1 for red, λ =3 for blue and λ =5 for green).

410x524mm (38 x 38 DPI)
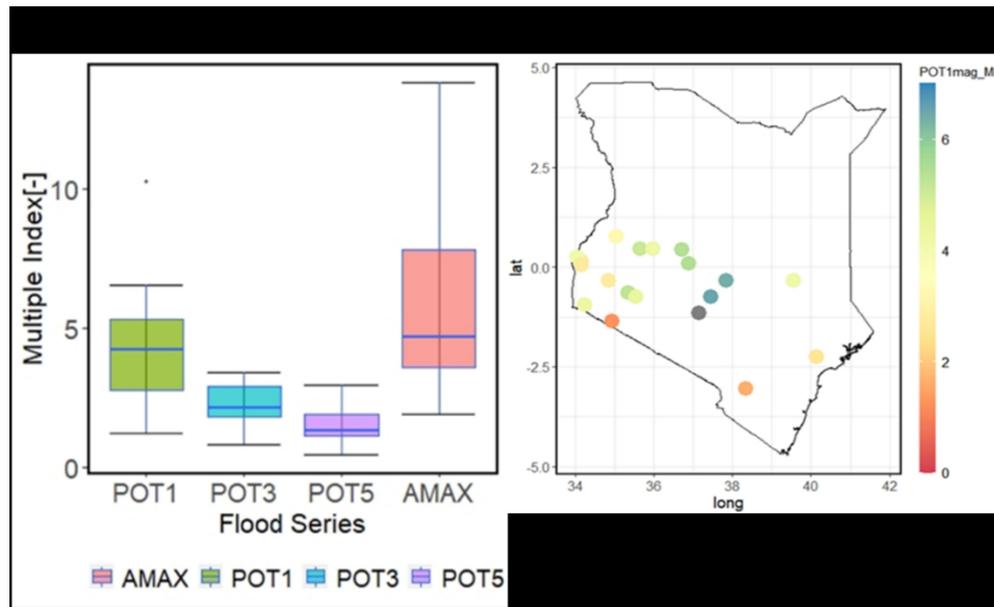
Figure 8:(a) Box plots for the Multiple Index (MI) for the different peak over thresholds (POT) flood series and the annual maximum (AMAX) for all the 19 stations. The bold line represents the 50th percentile; boxes and whiskers show the 25th and 75th percentiles, and the 10th and 90th percentiles.The mean number of events per year for POT thresholds (λ) considered are one (POT1), three (POT3) and five (POT5). (b) MI for POT1mag flood magnitude at the each of the 19 gauging stations across Kenya.

360x218mm (87 x 87 DPI)
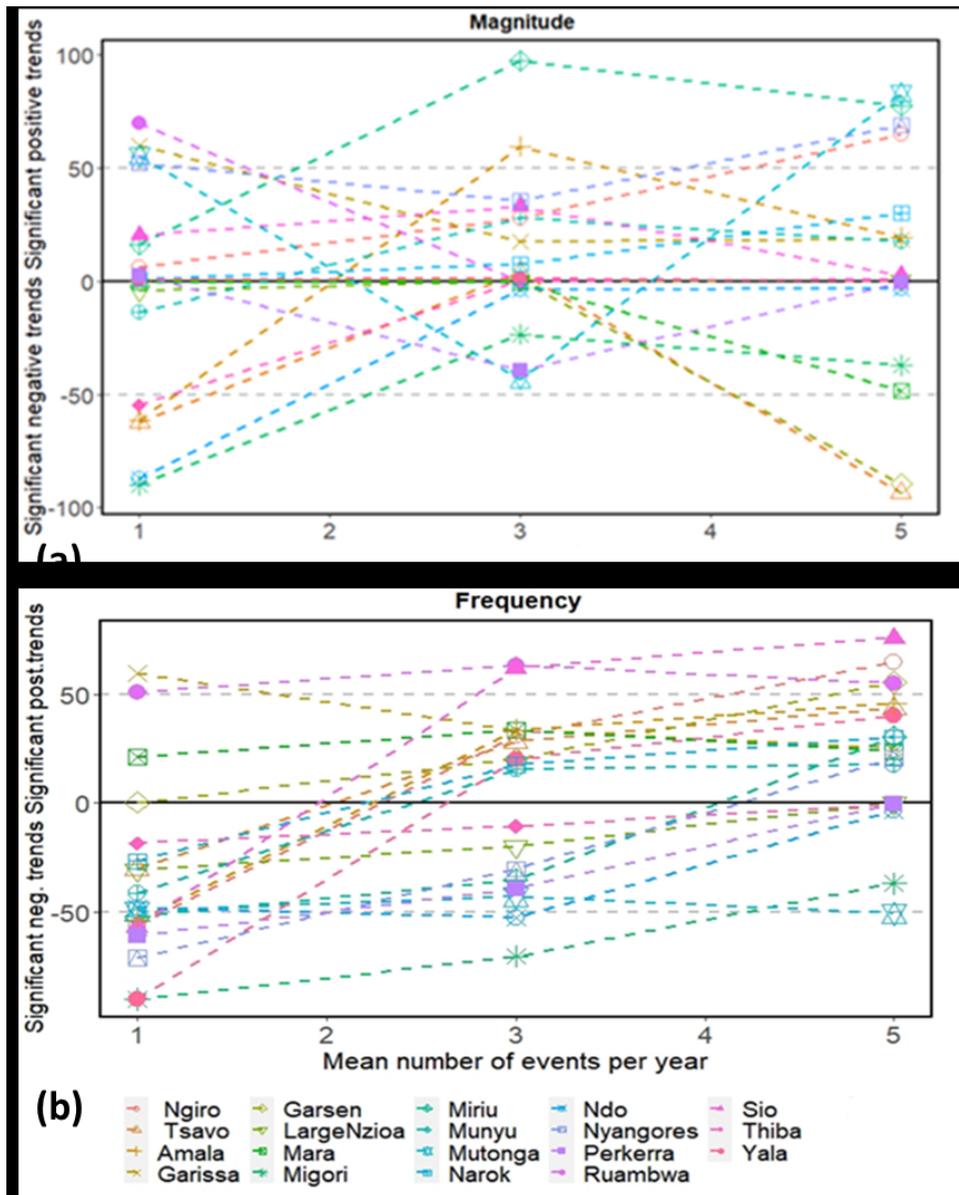
Figure 9: The sensitivity of trends in the POT series to the selection of different exceedances thresholds (λ) for different mean number of floods per year for POT1, POT3 and POT5 (a)magnitude and (b)frequency. Significant trends (expressed as percentages) in flood magnitude at 10% significance level with a threshold level λ range of 1, 3 and 5 mean events per year Different colors and symbols represent the different 19 stations considered in the study.

375x463mm (59 x 59 DPI)
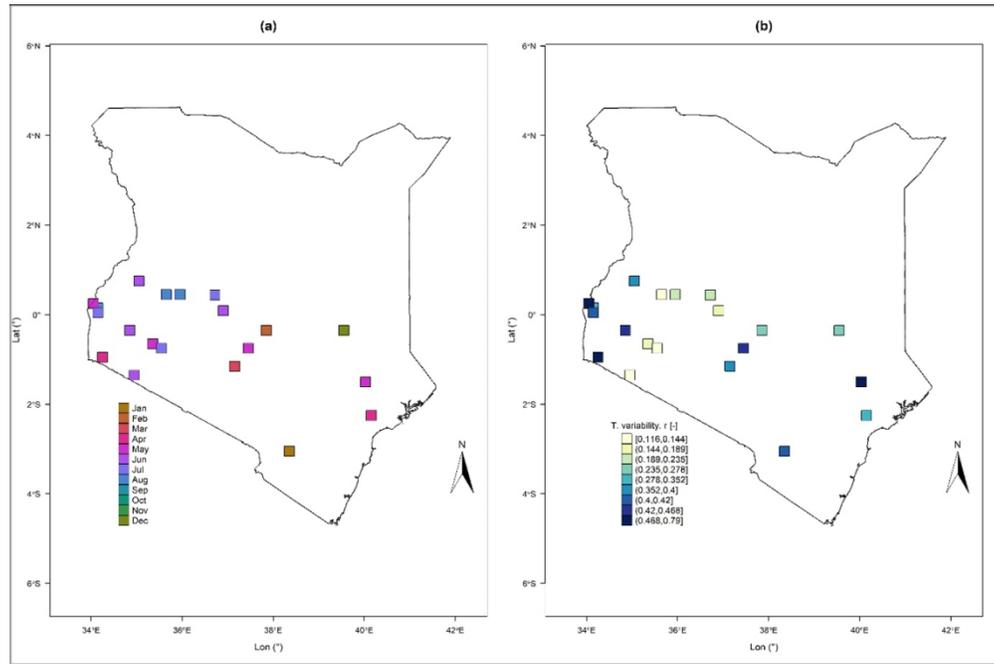
Figure 10: Annual flood seasonality measure for Kenya stream-gauge data. (a) Average flood timing (θ) and (b) Interannual variability of flood timing (r [−]) shown by different color saturation (the higher the variability is, the lighter the saturation is). The filled rectangular dots are reported at each of the 19 river gauging stations.

409x271mm (87 x 87 DPI)

River Name	Catchment Outlet point	Station ID	Lon	Lat	Drainage Area (km²)	Mean Elevation (m.a.s.l)	Mean Annual Rainfall (mm)	Annual Discharge (m³s⁻¹)	Catchment Characteristics	Human Influence		First & Last year of record	Record length (years)	Amount missing (%)
										Dams	Irrigation schemes			
Tana	Tana Garsen	4G02	40.11	-2.28	80 760	720	672	135.8	Semi-arid plains	9	11	1981-2016	36	58.2
Tana	Tana Garissa	4G01	39.7	-0.45	32 695	870	868	169.3	Highlands on the upstream & semi-arid plains in the lowland	8	7	1981-2018	38	14.2
Nzioa	RuambwaBridge	1EF01	34.09	0.12	12 643	1740	1488	151.2	Dense forest cover (highlands) & low trees & bushes (lower reaches)	2	4	1981-2018	38	13.6
Galana	Galana Tsavo	3G02	38.47	-2.99	6560	930	628	3.3	Semi-arid savannah plains	3	1	1981-2015	35	59.6
Gucha	Gucha Migori	1KB05	34.21	-0.95	6 310	1650	1435	45.0	Eastern lowlands with dense vegetation cover	0	2	1981-2015	35	47.8
Athi	Athi Munyu	3DA02	37.19	-1.09	5689	1730	822	18.8	Highlands and forest cover	3	1	1981-2017	37	21.6
Nzioa	Large Nzioa	1BD02	35.06	0.76	3878	1720	1267	15.3	Dense forest cover	1	1	1981-2011	31	28.8
Sondu	Sondu Miriu	1JG04	34.80	-0.33	3444	2017	1614	53.9	Low lying plains (western) & highland (Eastern)	2	2	1981-2018	38	64.4

Mara	Mara	1LA04	35.04	-1.23	2977	2100	1262	11.8	Low lying shrubs, semi-arid	0	1	1981-2015	35	77.7
Yala	Yala	1FG02	34.27	0.04	2700		1696	40.8	Swampy	0	0	1981-2019	39	59.6
Ewaso	Ewaso Narok	5AC10	36.73	0.43	2597	1600	880	5.3	Low lying shrubs & mainly semi-arid	0	2	1981-2018	38	26.5
Tana	Mutonga	4EA07	37.89	-0.38	1867	1830	1427	35.5	Highlands and forest cover	0	1	1981-2016	36	44.2
Ewaso	Ewaso Ngiro	5BC04	36.91	0.09	1837	1700	972	20.6	Low lying shrubs & mainly semi-arid	0	0	1981-2019	39	35.0
Sio	Sio	2EE07A	34.14	0.39	1011	1650	1822	15.5	Low trees & bushes & swampy in lower reaches	0	1	1981-2018	38	18.1
Turkwel	Ndo	2C07	35.65	0.45	897	1133	1371	9.1	Extensive palaeo-floodplain & arid conditions	0	1	1981-1993	13	47.2
Mara	Amala	1LB02	35.44	-0.89	695	2100	1377	6.8	Low lying shrubs, semi-arid	0	0	1981-2017	37	25.6
Mara	Nyangores	1LA03	35.35	-0.79	692	2008	1262	11.8	Semi-arid savannah plains, low lying shrubs, semi-arid	0	0	1981-2017	37	15.5
Turkwel	Perkerra	2EE07A	35.97	0.46	371	1023	832	5.7	Extensive palaeo-floodplain and arid conditions	1	1	1985-2005	21	50.1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

A4: GR4J model simulations and observations across some of the Kenyan catchment
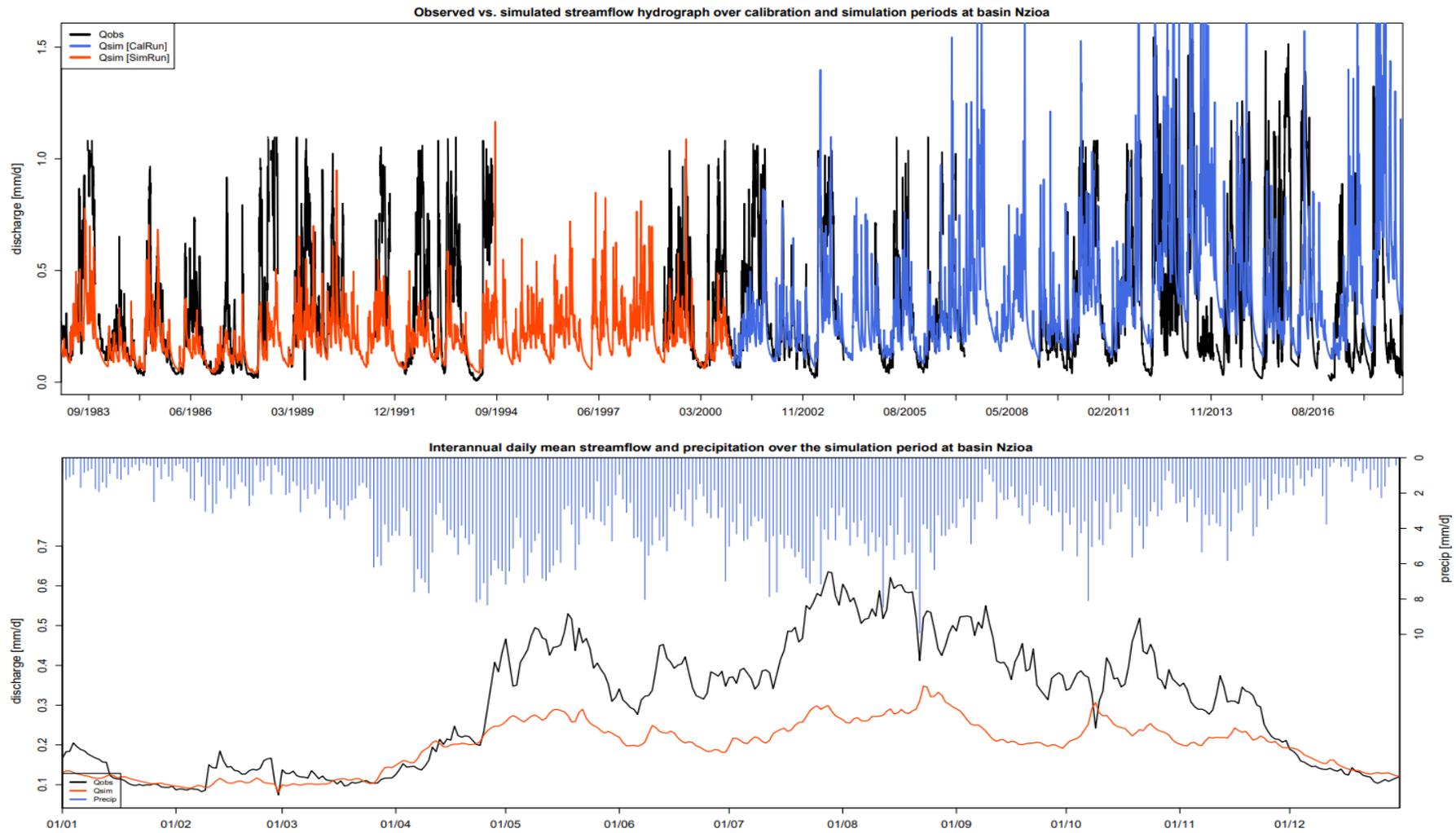
Figure A4.1: - Observed vs. simulated hydrograph over calibration (red) and simulation (blue; top panel) and interannual daily mean streamflow and Precipitation (over simulation period) at Nzioa -Ruambwa
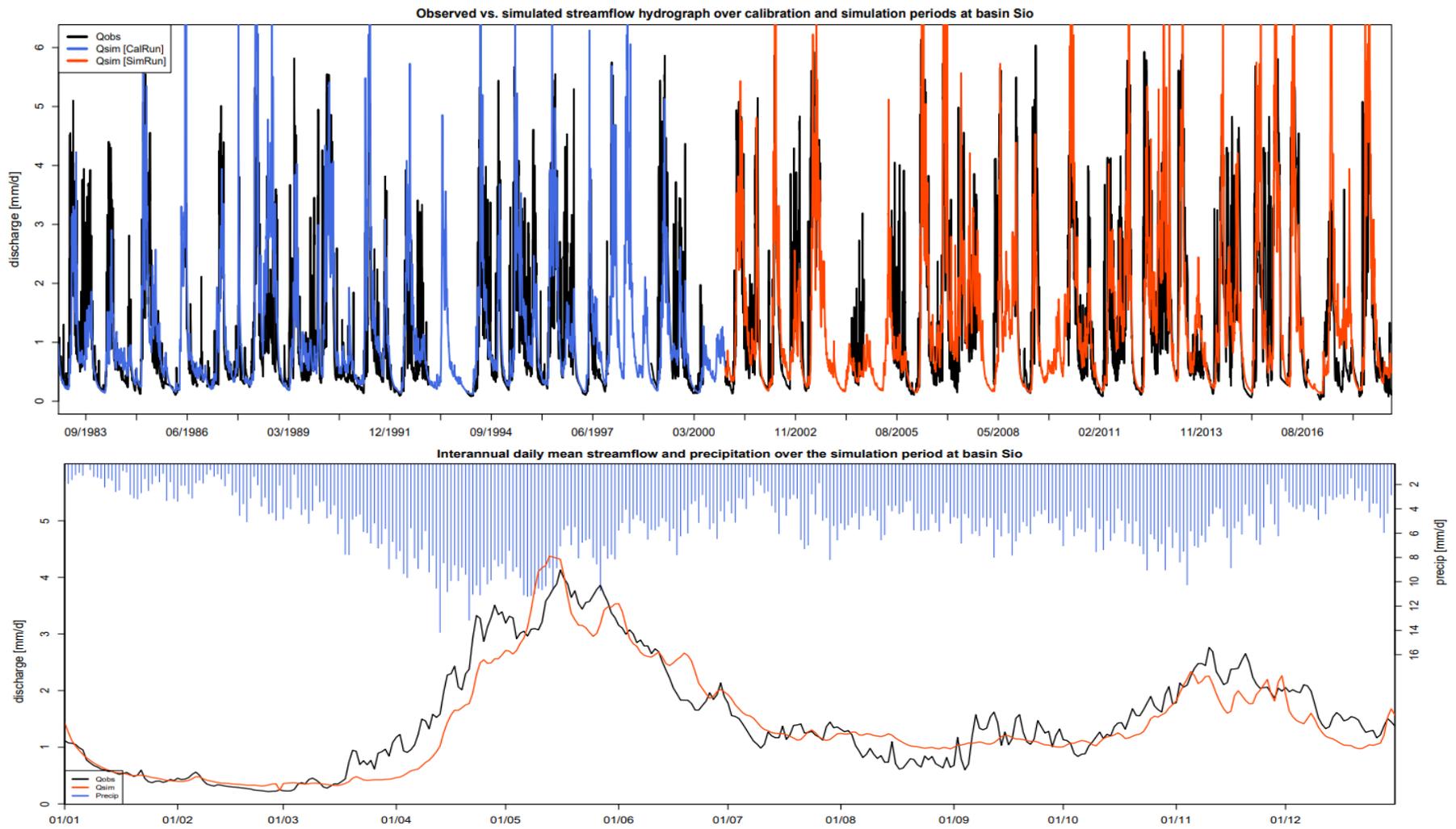
Figure A4.2: - Observed vs. simulated hydrograph over calibration (red) and simulation (blue; top panel) and interannual daily mean streamflow and Precipitation (over simulation period) at Sio
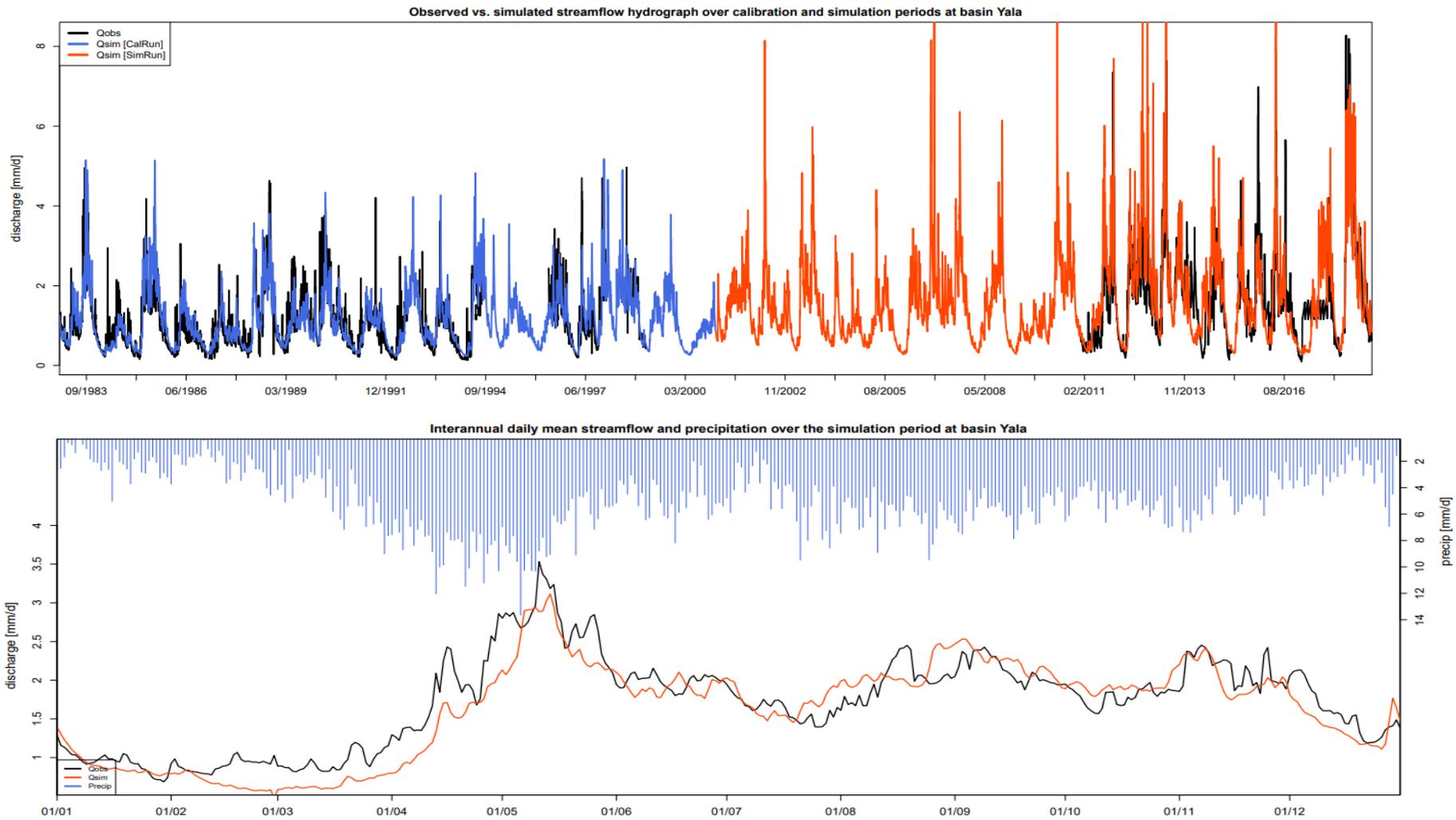
Figure A4.3: - Observed vs. simulated hydrograph over calibration (red) and simulation (blue; top panel) and interannual daily mean streamflow and Precipitation (over simulation period) at Yala
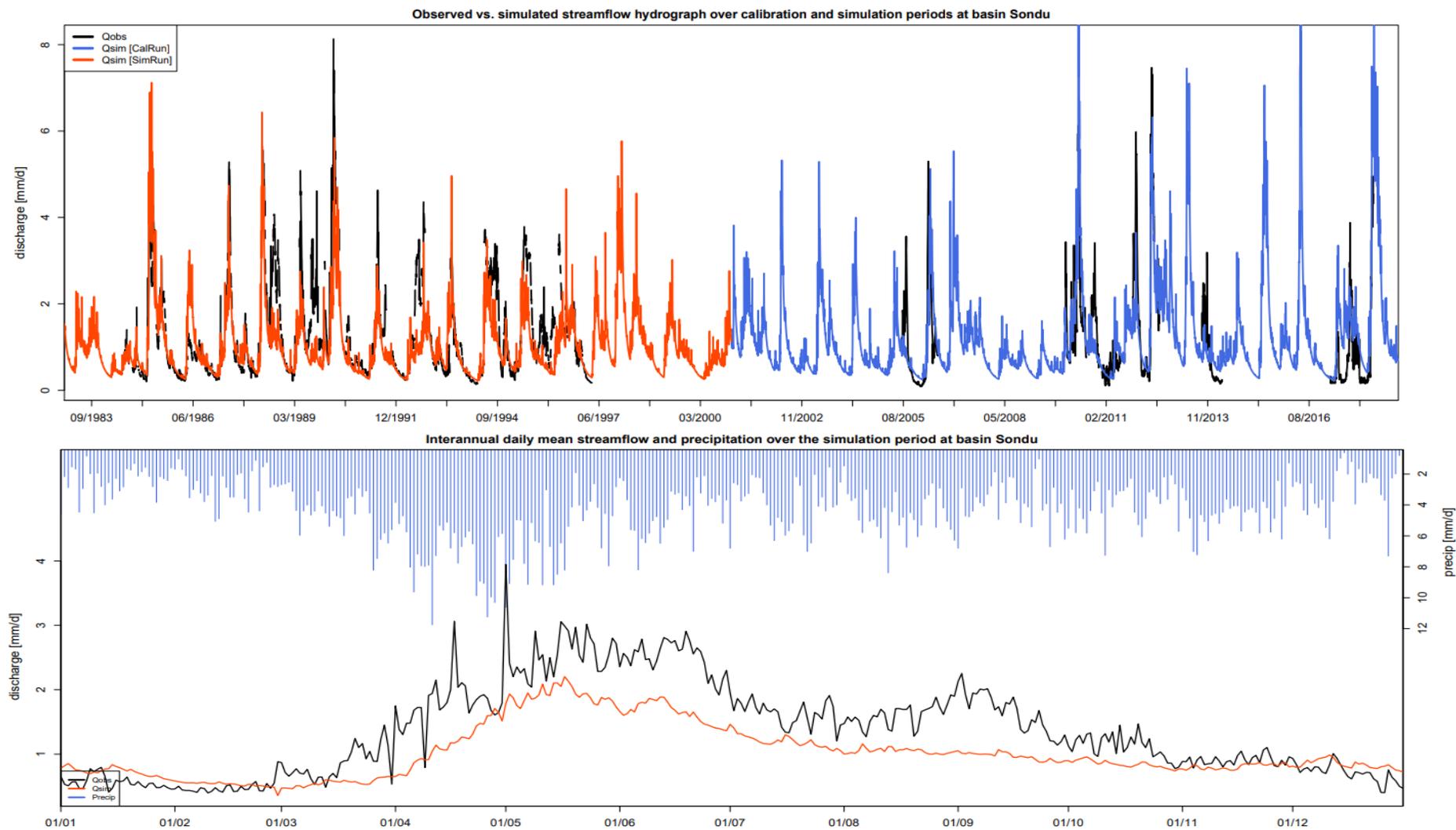
Figure A4.4: - Observed vs. simulated hydrograph over calibration (red) and simulation (blue; top panel) and interannual daily mean streamflow and Precipitation (over simulation period) at Sondu-Miriu
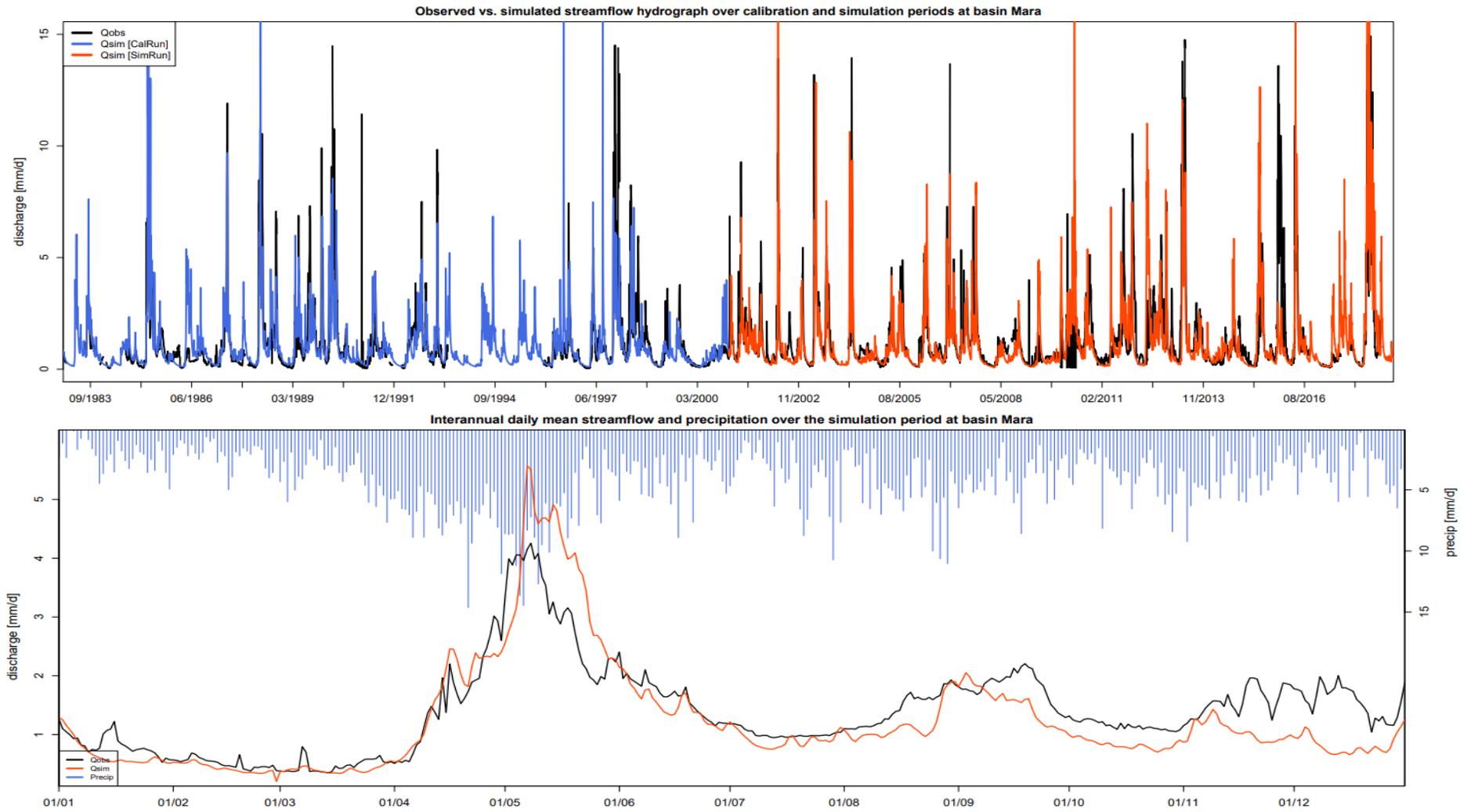
Figure A4.5: - Observed vs. simulated hydrograph over calibration (red) and simulation (blue; top panel) and interannual daily mean streamflow and Precipitation (over simulation period) at Mara.
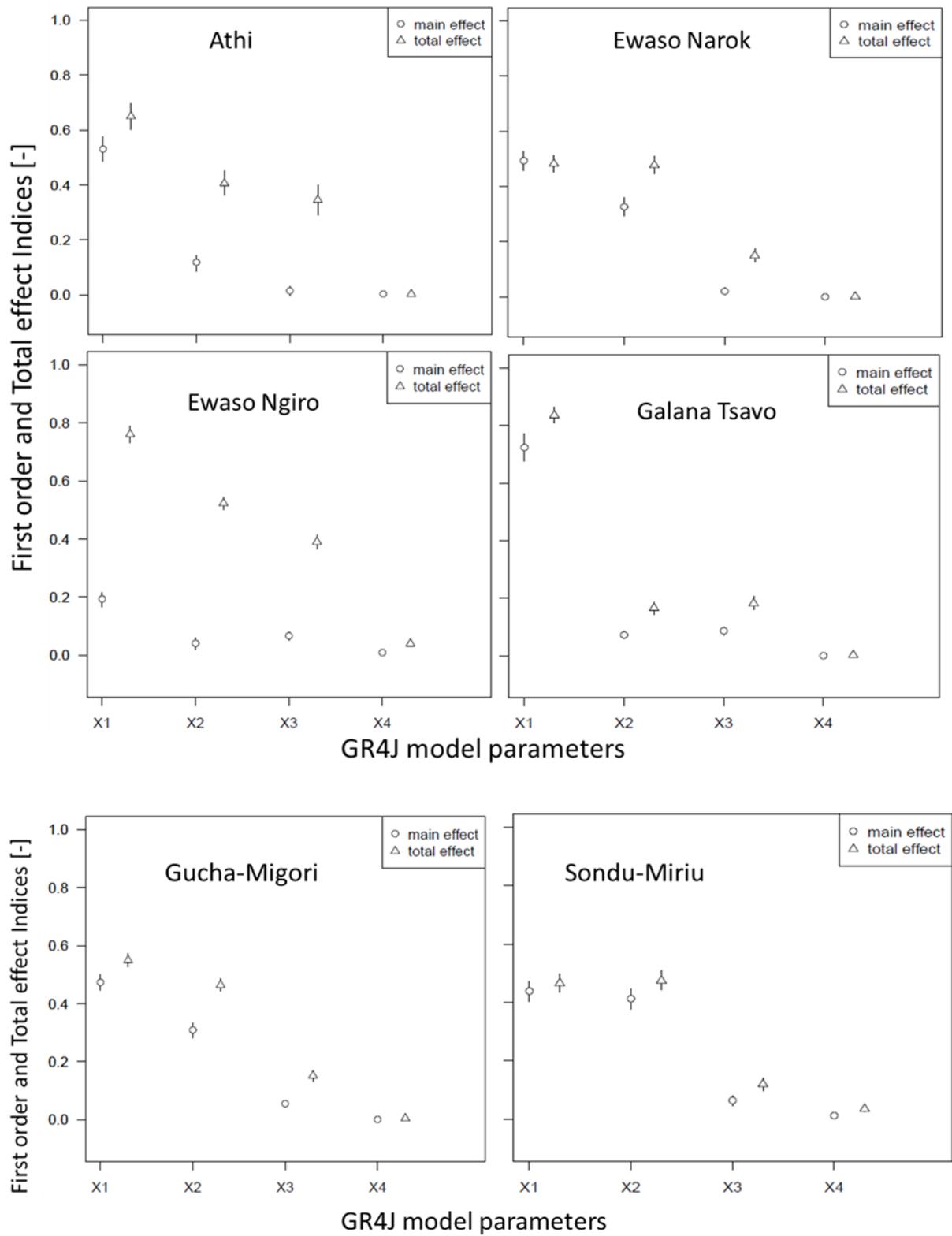
Figure A5.1: - First order and total effect Sobol' sensitivity indices for GRJ4 model parameters across Kenyan catchments
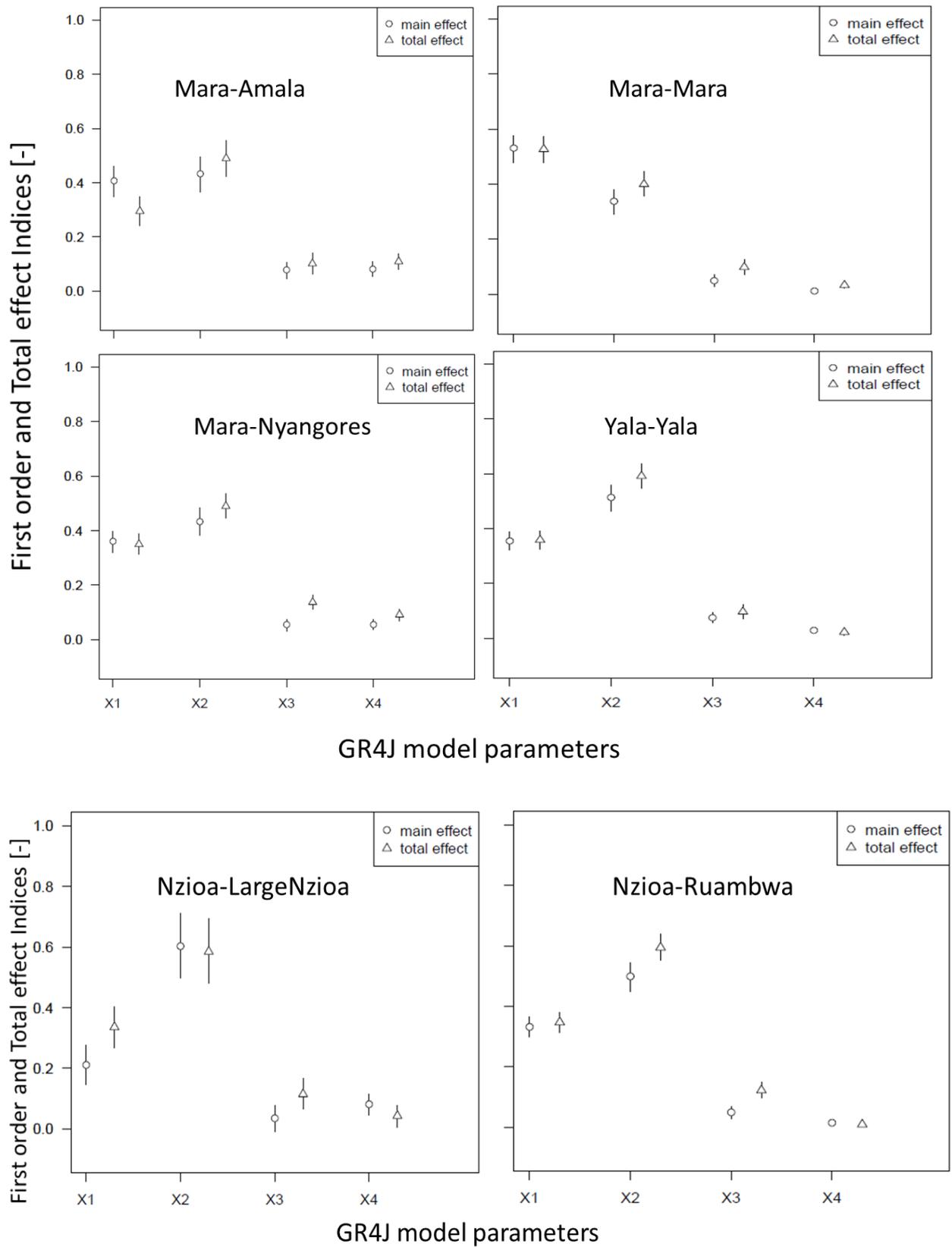
Figure A5.2: - First order and total effect Sobol' sensitivity indices for GRJ4 model parameters across Kenyan catchments
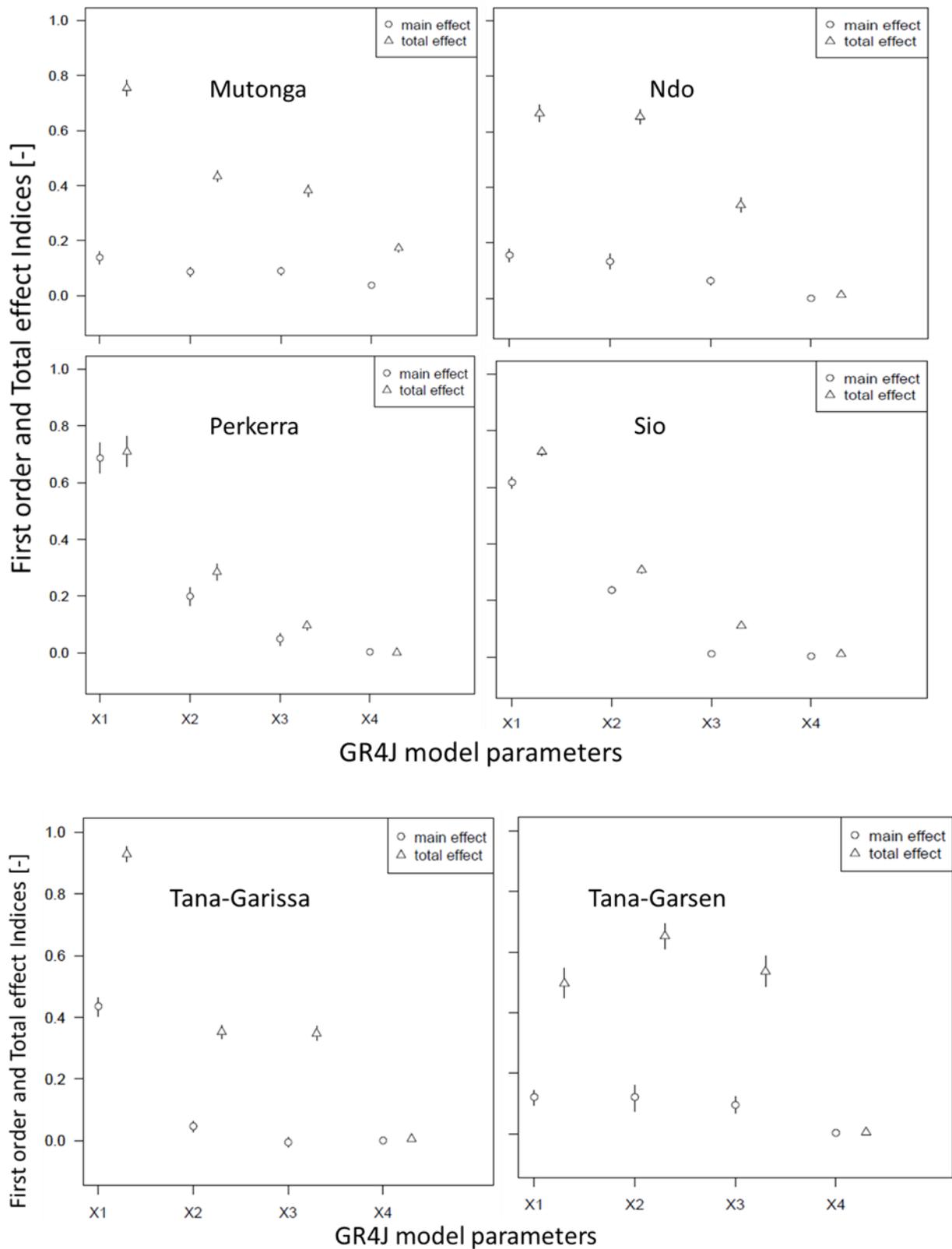
Figure A5.3: - First order and total effect Sobol' sensitivity indices for GRJ4 model parameters across Kenyan catchments

A5b: Sensitivity analysis sample results: - GR4J model parameters interactions
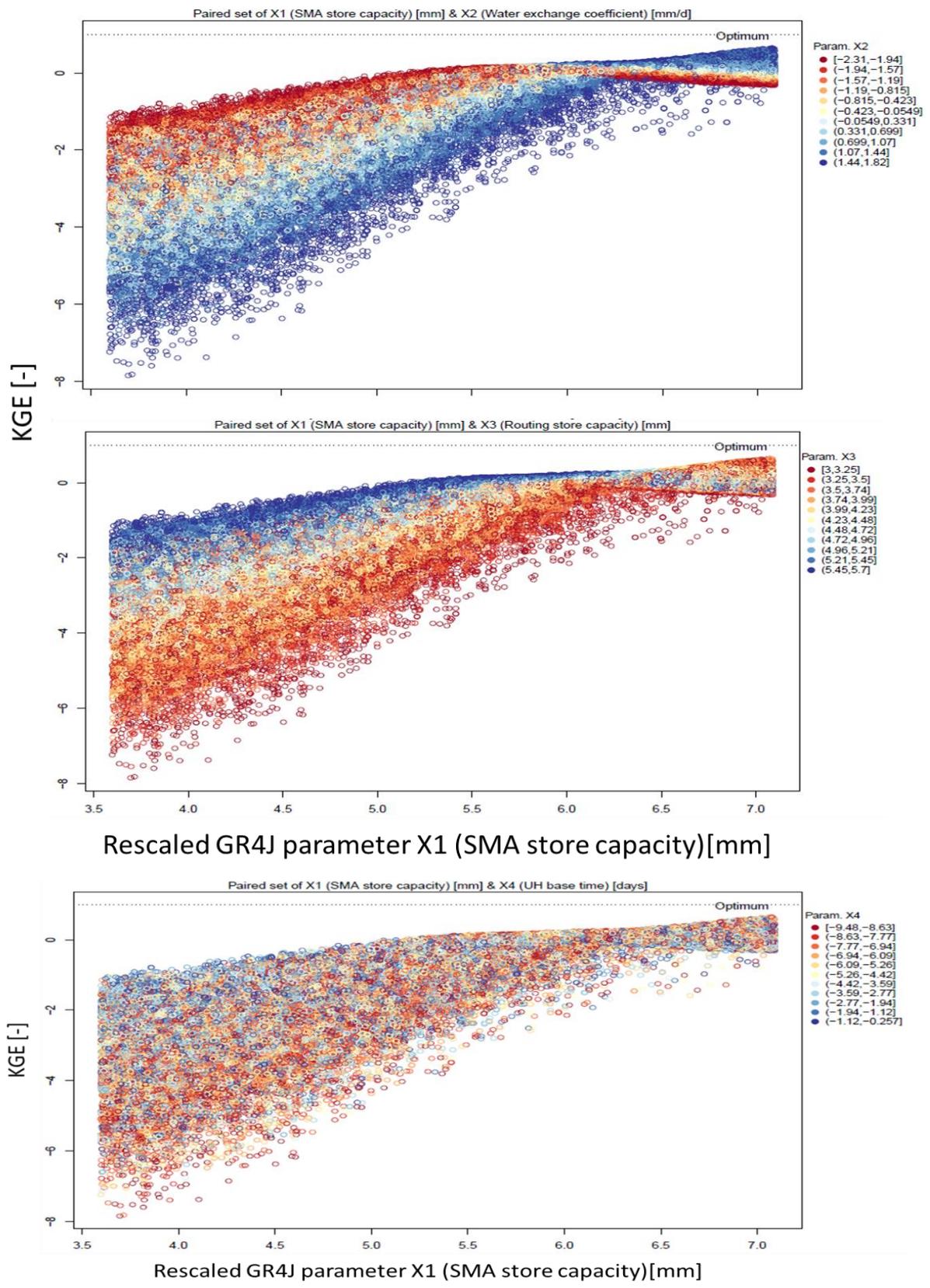
Figure A5.4: - KGE vs, GR4J model parameter interaction from Sobol' sensitivity analysis for Tana Garissa
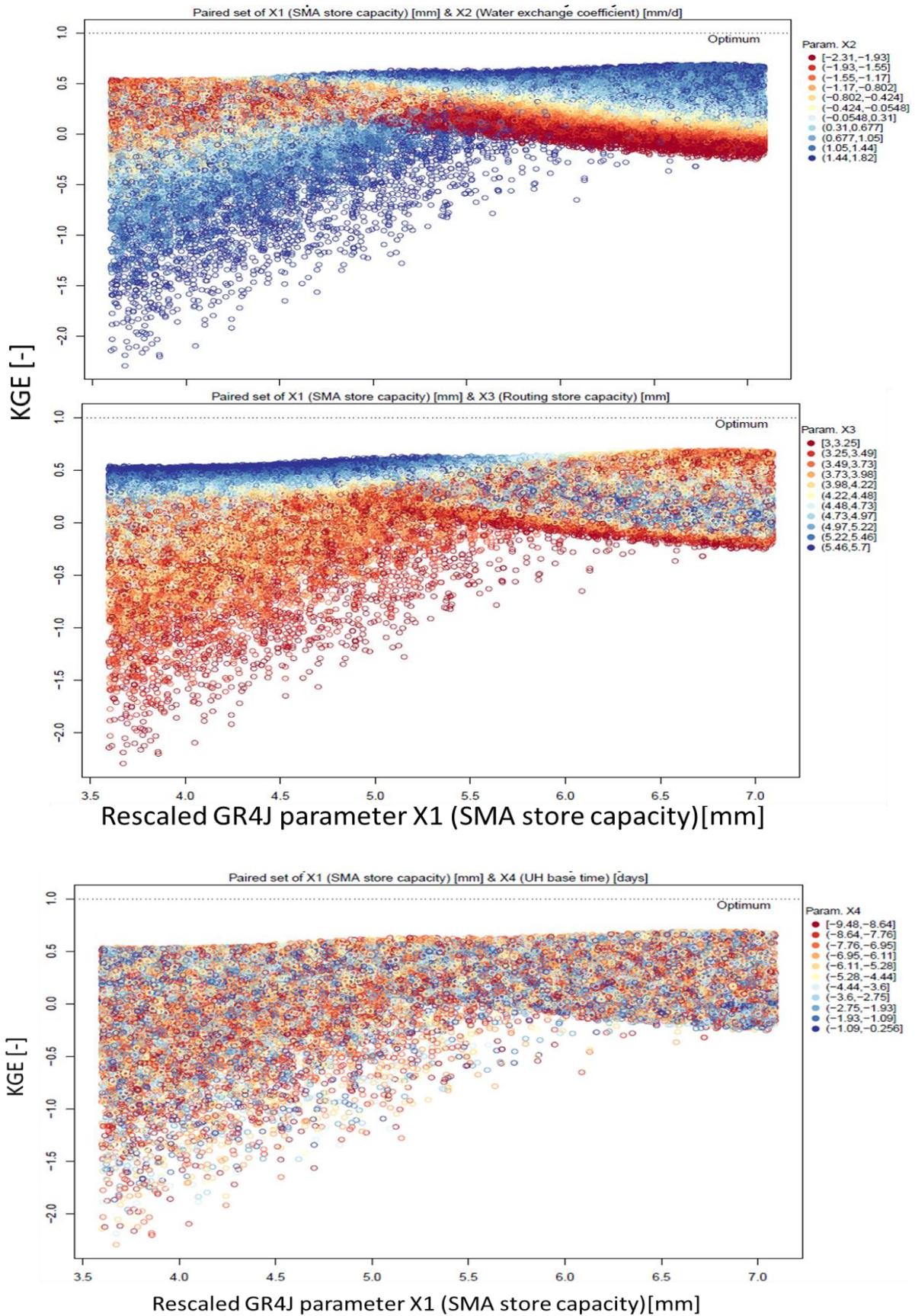
Figure A5.5: - KGE vs, GR4J model parameter interaction from Sobol' sensitivity analysis for Nzioa Ruambwa

1. ***Paper Title:*** Unveiling flood generating mechanisms across Kenya

2. ***Message of the paper:*** Acknowledging that there has been a shift in the frequency and magnitude of floods in recent decades across Kenya, we assess different **causes/flood generating drivers** to understand their **contribution to floodiness**, with particular interest to quantitatively establish their **relative importance and change** over time in different catchments.

3. ***Introduction***

   Flooding is classified among the most detrimental natural hazards worldwide (Berghuijs et al., 2019; Kundzewicz et al., 2014; Paprotny et al., 2018) , and with the changing climate there is expected increase in flood risk globally (Arnell et al., 2016; Liu et al., 2018). For Kenya, floods is the most common climatic extreme and the leading hydro-meteorological disaster (Huho & Kosonei, 2014).

   Flood timing, magnitude and frequency are useful in informing the trends in river flooding (Berghuijs et al., 2019;2017). For instance, the timing of floods in East Africa has shifted in recent decades (StephensElisabeth et al., 2015), but trends in flood magnitudes are less clear (Coughlan de Perez et al., 2017). There is need therefore to understand such trends, but this requires carefully identification of triggers and hydrological processes (Berghuijs et al., 2019; Slater et al., 2017). The causes of historical flood trends is still unclear due to a limited understanding of regional variations in flood-generating mechanisms (Tramblay et al., 2019)and the uncertainty in projections of future flooding under climate change (Berghuijs et al., 2019; Kundzewicz et al., 2014).

   Factors other than rainfall play an important role in controlling floods (Slater et al., 2015). These factors include storage components such as the land surface and subsurface memory (groundwater, soil moisture) and transfer components such as the interaction between the spatial and temporal rainfall patterns and the river network configuration and the catchment concentration time (the time it takes precipitation to reach the river mouth), as well as man-made interventions such as reservoirs (Coughlan de Perez et al., 2017). Accordingly, it follows that the most extreme amount of monthly precipitation ever recorded (for example) may not correlate with the most extreme flood (Stephens et al., 2016). Therefore, a number of studies have documented causes of flooding but only for small river basins or number of flood events (Blöschl et al., 2013), identified only a single dominant mechanism driving river flooding at each site (Berghuijs et al., 2016), thus overlooking the fact that even at a single site, individual floods can arise through different mechanisms (Blöschl et al., 2013).

   Classification of flood generating mechanism across numerous catchments are available for many parts of the world (Berghuijs et al., 2017), such as continental United States (Berghuijs et al., 2016), Europe (Berghuijs et al., 2019) and Sub-Saharan Africa (Coughlan de Perez et al., 2017). However, for Kenya, there is no reproducible quantitative mapping of the importance of different flood- generating mechanisms. Also, under the changing climate, the significance of different flood-generating mechanisms may shift over time, hence, the relative importance of flood drivers

may shift over time, but to date there has been no systematic effort to detect whether such changes are taking place in Kenya.

4. *Research Paper Objectives*

The main objective of this paper is to assess the associated flood generating mechanism, identify which is most significant in simulating flood peaks and how they have changed overtime in Kenyan catchments.

Specific research questions are:

i. What are the **dominant flood generating mechanisms** in Kenyan catchments?

ii. What is the average **timing** of river flooding in terms of **mean date of occurrence** and **floodiness** over Kenya?

iii. What is the **relative importance** of the identified **flood generating** mechanisms?

iv. Has the relative importance of flood generating mechanisms **changed over time**?

5. *Data and Methods used*

5.1.1. *Maximum annual flood dates/Water level*

The Kenya Water Resource Authority will provide the daily river discharge and water level data for the selected catchments. The maximum flood dates will be computed from the river discharge data available for the 19 stations with consistent long data records. Analyzing the dates of the floods will provide deeper insights into the processes driving change than analyzing flood magnitudes alone. In addition, the date can be identified equally well from discharge and water level data which increases the temporal and spatial.

5.1.2. *Daily Precipitation data*

This study will incorporate datasets based exclusively on reanalysis data i.e ERA-5 and CHIRPS V2.0, which is combined gauge, satellite and reanalysis. If we cannot get daily observed gauge station data from KMD, CHIRPS will be used as a base observation data.

5.1.3. *Soil moisture data*

5.2. *Methods*

5.2.1. *Definition of flood generating mechanism*

We will focus on processes that drive flooding (e.g., rainfall), rather than catchment properties (e.g., land cover) and the meteorological drivers (e.g. ENSO and IOD). We follow the three major mechanisms that may lead to flooding have been classified by Berghuijs et al. ([2016](#)) (.

5.2.2. *Seasonality characteristics*

The seasonality of flooding and flood-generating mechanisms will be characterized by circular statistic (Berghuis 2016;2019, Boschl….2017), then quantify the **mean date of occurrence** (day of year) of all the processes.

5.2.3. Relative importance of the flood generating mechanisms

We will quantify the relative importance of the established flood-generating mechanisms for each catchment following the linear algometric applied in (Berghuis et al 2019)

5.2.4. Changes in relative importance over time

To assess whether the relative importance of the established flood drivers changes over the period of our study analysis, we use the linear algometric applied in (Berghuis et al 2019), estimate the relative importance of the mechanisms separately for the periods for two split periods (e.g. 1960-1984 and 1985-2010 and compare their differences in each catchment.

6. Results / Graphs expected

1) spatial maps showing spatial pattern of the timing of maximum annual flows and the flood-driving mechanisms; 3) Spatial map showing seasonality characteristics of floods and flood generating mechanisms.

7. Discussion

Compressive discussion covering trends, seasonality, and attribution of flooding across Kenya.

8. Conclusion

Draw a conclusion from the results and how best the research question will be answered by the research and possible recommendations. Does my research agree with the earlier studies over the region applying different hydrological models, etc?

Potential journal: HESS, Hydrology and Earth System Sciences, https://www.hydrology-and-earth-system-sciences.net/

Potential partners / co-authors (TBC): *Maureen Anyango Wanzala, Hannah Cloke, Liz Stephens Andrea Ficchi, (University of Reading), Shaun Harrigan (European Centre for Medium Range Forecasts))*

References

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Dutra, E., Fink, G., Orth, R., & Schellekens, J. (2017). Global evaluation of runoff from ten state-of-the-art hydrological models. *Hydrology and Earth System Sciences*, *21*(6), 2881–2903.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., … Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, *21*(12), 6201–6217.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., … Vitart, F. (2011). The

ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Driouech, F., Déqué, M., & Mokssit, A. (2009). Numerical simulation of the probability distribution function of precipitation over Morocco. *Climate Dynamics*, *32*(7–8), 1055–1063. https://doi.org/10.1007/s00382-008-0430-6

Ficchì, A., Perrin, C., & Andréassian, V. (2019). Hydrological modelling at multiple sub-daily time steps: model improvement via flux-matching. *Journal of Hydrology*. https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.05.084

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., … Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, *2*, 1–21. https://doi.org/10.1038/sdata.2015.66

Gosset, M., Viarre, J., Quantin, G., & Alcoba, M. (2013a). Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Quarterly Journal of the Royal Meteorological Society*, *139*(673), 923–940. https://doi.org/10.1002/qj.2130

Gosset, M., Viarre, J., Quantin, G., & Alcoba, M. (2013b). Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Quarterly Journal of the Royal Meteorological Society*, *139*(673), 923–940.

Gupta, H. V, Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1), 80–91. https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.08.003

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research Atmospheres*, *113*(20). https://doi.org/10.1029/2008JD010201

Hersbach, H. (2018). *Operational global reanalysis: progress, future directions and synergies with NWP*. European Centre for Medium Range Weather Forecasts.

Hossain, F., & Huffman, G. J. (2008). Investigating error metrics for satellite rainfall data at hydrologically relevant scales. *Journal of Hydrometeorology*, *9*(3), 563–575.

KOBAYASHI, S., OTA, Y., HARADA, Y., EBITA, A., MORIYA, M., ONODA, H., … TAKAHASHI, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Le, A. M., & Pricope, N. G. (2017). Increasing the accuracy of runoffand streamflow simulation in the Nzoia Basin, Western Kenya, through the incorporation of satellite-derived CHIRPS data. *Water (Switzerland)*, *9*(2). https://doi.org/10.3390/w9020114

Li, L., Hong, Y., Wang, J., Adler, R. F., Policelli, F. S., Habib, S., … Okello, L. (2009). Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Natural Hazards*, *50*(1), 109–123. https://doi.org/10.1007/s11069-008-9324-5

Maggioni, V., Meyers, P. C., & Robinson, M. D. (2016). A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. *Journal of Hydrometeorology*, *17*(4), 1101–1117.

Pan, M., Li, H., & Wood, E. (2010). Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resources Research*, *46*(9), 1–10. https://doi.org/10.1029/2009WR008290

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, *279*(1–4), 275–289.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., … Goldberg, M. (2010). The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1058. https://doi.org/10.1175/2010bams3001.1

Seiler, R. A., Hayes, M., & Bressan, L. (2002). Using the standardized precipitation index for flood risk monitoring. *International Journal of Climatology*, *22*(11), 1365–1376.

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A. T., … Huffman, G. J. (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*, *104*, 70–97.

van der Knijff, J. M., Younis, J., & de Roo, A. P. J. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, *24*(2), 189–212. https://doi.org/10.1080/13658810802549154

Vischel, T., Quantin, G., Lebel, T., Viarre, J., Gosset, M., Cazenave, F., & Panthou, G. (2011). Generation of high-resolution rain fields in West Africa: Evaluation of dynamic interpolation methods. *Journal of Hydrometeorology*, *12*(6), 1465–1482.

Willmott, C. J., Robeson, S. M., & Matsuura, K. (2017). Climate and other models may be more accurate than reported. *Eos*, *98*(PUBART).

A7: Collaborative publications, articles, and media interviews

Papers

1. Omukuti, J., **Wanzala, M.A**., Ngaina, J. and Ganolae, P. Local decision makers in Africa need Medium- to Long-term climate information services to comprehensively manage climate risk. *Journal of Climate Resilience and Sustainability*. Manuscript ID: CLI-21-030-R2, Under Review.

2. Gudoshava, M., **Wanzala, M**., Thompson, E., Mwesigwa, J., Endris, S.H., Segele, Z., Hirons, L., Kipkogei, O., Mumbua S.C., Njoka, W., Baraibar, M., De Andrade, M.F., Woolnough, S., Atheru, Z and Artan, G. Application of Real Time S2S forecasts over Eastern Africa in the Co-production of Climate Services, Manuscript. *Journal of Climate Services*. Manuscript ID: CLISER-D-21-00142R1, Preprints.

Articles

3. Brimicombe, C., Di Napoli, C., Cloke, H. and Wanzala, M., 2021. Reviewing the summer of extreme weather in 2021. *CarbonBrief*. Available at: - *https://www.carbonbrief.org/guest-post-reviewing-the-summer-of-extreme-weather-in-2021/*.

4. Wanzala, M. and Cloke, H, 2021. How can Africa learn from other countries about combatting severe floods? *Anticipation Hub*. Available at: *https://www.preventionweb.net/news/what-can-africa-learn-other-countries-about-combatting-severe-floods*.

5. Wanzala, M.A. and Ogallo, L, 2020. Recurring Floods in Eastern Africa amidst Projections of Frequent and Extreme Climatic Events for the Region, *Icpac Medium.com. Available at https://icpac.medium.com/recurring-floods-in-eastern-africa-amidst-projections-of-frequent-and-extreme-climatic-events-for-30d20d0d6f76*

Media interviews

6. "Climate change; A risk to food security", *February 2022*, *BBC world.* Available at:- *Wanzala, M. contribution.*

7. "State of flooding in Berkshire", *May 2021, BBC Berkshire.* Link Available on request

8. "How are climate models being used for planning and adaptation?" *September 2019*: Contribution to REACH: Improving water security for the poor. Available at: - *https://www.youtube.com/watch?v=8Fu5bIzWhrg&t=16s*

Science & art communications

9. "A paint pot in the sky: meeting points between science and arts", Doctoral Research Highlights, 2020. Available at *http://www.reading.ac.uk/web/files/graduateschool/Research_Highlights_2020_web.pdf*.

10. "Beware: Floods Ahead!" A poem performed by flooding scientists, October 2020. Available at: - https://www.youtube.com/watch?v=bZ4KnJdaAjY