Journal of Water Resource and Protection, 2021, 13, 362-393
https://www.scirp.org/journal/jwarp
ISSN Online: 1945-3108
ISSN Print: 1945-3094

Flood Change Detection and Attribution Using Simulation Approach in Data-Scarce Watersheds: A Case of Wabi Shebele River Basin, Ethiopia

Fraol Abebe Wudineh[1*], Semu Ayalew Moges[2], Belete Berhanu Kidanewold[1]

[1]School of Civil and Environmental Engineering, Addis Ababa University, Addis Ababa, Ethiopia
[2]Department of Civil and Environmental Engineering, University of Connecticut, Connecticut, USA
Email: *fraolabebe21@gmail.comm, semu.moges.2000@gmail.com, bb_academy@yahoo.com

How to cite this paper: Wudineh, F.A., Moges, S.A. and Kidanewold, B.B. (2021) Flood Change Detection and Attribution Using Simulation Approach in Data-Scarce Watersheds: A Case of Wabi Shebele River Basin, Ethiopia. *Journal of Water Resource and Protection*, **13**, 362-393.
https://doi.org/10.4236/jwarp.2021.135023

Received: January 24, 2021
Accepted: May 24, 2021
Published: May 27, 2021

Abstract

Flood events vary with sub-regions, sites and time and show complex characteristics. This study investigated temporal variabilities in flood discharges and relationships with principal driving factors in data scarce Wabi Shebele River Basin. The preliminary analysis using exploratory data analysis (EDA) on annual and seasonal maximum discharge reveals that there are cycles of extreme flows at five- and ten-year intervals respectively throughout the basin. The statistical verification using the Mann-Kendall test and Quantile perturbation method indicates a significant trend in flood magnitude and frequency entire the basin in the early 21st century. For longest period (1980-2010) annual maximum stream flow shows significant positive trend (p-value < 0.05) in middle catchments and negative trend (p-value < 0.05) in eastern catchments. The years: 1986-1995, 2006-2010 are the years in which positive significant anomalies occurred in all seasons, while the years: 1980-1985, 1996-2005 are the occurrence years of significant negative anomalies. Rainfall from climate drivers; DA, BE, VS and fraction of sand from environmental background drivers; fraction of forest and population density from external factors were identified as the powerful driving factors of flood variabilities in the Wabi Shebele River Basin.

Keywords

Flood Events, Watersheds, Wabi Shebele River, Hydrological Model, Driving Factors of Flood

1. Introduction

Flood is excessive water availability, which is caused by above normal stream flow, leading to inundation of areas that are normally not covered by water. In flood generation, antecedent conditions which refer to the saturation of natural storage in the catchment are critical factors [1] [2]. It is the consequence of earlier precipitation or snowmelt in the catchment. If the saturation is at maximum capacity, a consecutive moderate amount of rain can also generate large floods [3]. Further, low permeability of the surface due to dry and crusted soil after a prolonged period without rain can also rapidly convert heavy rainfall to a runoff which usually results in a flash flood [4]. Intense and/or long-lasting rainfall is the main cause of large flood events in tropical regions [5] [6]. In these regions, more water vapor and heat in the atmosphere brings storms and consequently floods will be become more intense [7].

Ethiopia has experienced two major types of floods: flash and river (fluvial) floods. The occurrences and extents of river floods for the last six decades are increased from decade to decade [8] [9]. Spatial distribution of major flood events over major river basins shows: Awash Basin recorded 13 flood events and followed by Wabi Shebele, Rift Valley Lakes, and Genale Dawa with 12, 11, and 11 flood events, respectively [9]. In the eastern and southeastern part of Ethiopia, several areas have been afflicted by floods for decades, killing hundreds and displacing thousands of people. Akola *et al*., [8] reported that, in Diredawa city (Ethiopia), the frequency of floods has been increasing from 1945, 1977, 1981, 1997, 2001, 2004, 2005 to 2006. Similarly, the recent history of the Wabi Shebele River Basin is marked by frequent destructive floods [10] [11] [12] [13] (e.g., floods in 1996, 1999, 2000, 2003, 2005, 2006, 2008, 2010 and 2016).

Studying the changes in river discharge and precipitation patterns has critical importance as a climatic indicator for environmental risk problems such as droughts and floods. Among these analyzing a river discharge gives the entire picture of a catchment response for water resource planning and management [13] [14]. The seasonality of streamflow and cost of extreme weather events has been found a rapid upward trend in recent decades throughout the world including Ethiopia [11] [15] [16] [17] [18] [19]. This makes the search for trends in flood data series area of scientific interest. In addition, the study of trends in flood time series requires decision-makers to better understand the ongoing changes in hydrologic extremes to make preparations for the possibility of changing conditions. Flooding is the result of hydraulic basin condition, environmental susceptibility of areas to flooding, anthropic impacts and precipitation. A possible correlation with climate change needs to be considered a given scenario, which may be changed in rainfall regimes (probably the most common).

This paper aims to study possible flood changes through a flood simulation approach in watersheds, in conditions of data scarcity and identify correlations with potential driving forces. In Wabi Shebele River Basin obtaining historical river, records are hardly difficult. Some gauging stations that exist in the upper part of the basin have inconsistent data which needs extra data extension mechanisms. Thus, an alternative approach (hydrological modeling) to generate flow from weather and catchment variables is required in such areas.

2. Study Area and Data

2.1. Study Area

The Wabi Shebele Basin is a transboundary basin located at Horn of Africa, situated in between Ethiopia and Republic of Somalia. It originates from Bale highlands ranges of the Galama to Ahmar of Ethiopia, about 4000 m altitude and drains portion of Somalia before draining to Indi-an Ocean. More than 70% of the catchment (202,220 km^2) is lying in Ethiopia. The Wabi Shebele Basin in this study is used to represent the catchment that is lying in Ethiopia within 4˚45'N to 9˚45'N latitude and 38˚45'E to 45˚45'E longitude [20]. The air temperature of the Wabi Shebele Basin varies with altitude [10]. The mean monthly temperature of the basin is 19.9˚C. The climate of the basin is dependent on the altitude and strong latitudinal movement of the intertropical convergence zone (ITCZ) [20]. The highlands are cool and densely populated while the lowlands are arid and sparsely populated with recorded annual rainfall of 1213 mm and 268 mm respectively [10] [20].

The basin is divided into four geographical areas based on morphometric characteristics and rainfall regimes [10] [21]: upper catchments which characterized by a mountainous area with abrupt valleys; middle catchments which is wider highland and rainy area; eastern catchments which characterized semi-arid areas and lower catchments which covers arid lowland area of the basin. The watershed characteristics analysis using Arc GIS indicate that flood characteristics of Wabi Shebele Basin is related to basin and relief morphometric characteristics. The mean peak flow (QMPF) in Wabi Shebele Basin has large positive association with linear morphometric parameters (like valley length, mean stream length) and with basin morphometric parameter (like drainage size, shape factor) and negative associations with relief morphometric characteristics (like basin elevation and valley slope). The basin has three climatological rainy seasons: spring (February-May), summer (June-September) and winter (October-January) [10] [18] [22]. While having the largest area coverage, the basin's annual runoff is estimated to 3.4 BCM (Billion Meter Cube) (**Figure 1**).

2.2. Data

We utilize both ground-based station observations and gridded analysis data in this paper. In Hydrologic model, digital elevation model (DEM), physiographic data of pedology, land use and occupation and classes of slopes and meteorological data were used to generate flow. Digital elevation model (DEM) with a spatial resolution of 90 meters obtained from SRTM GDEM official website.

The soil information was acquired from FAO Digital Soil Map of the World (DSMW) at a scale of 1:500,000 downloaded from FAO Soil Map and Database

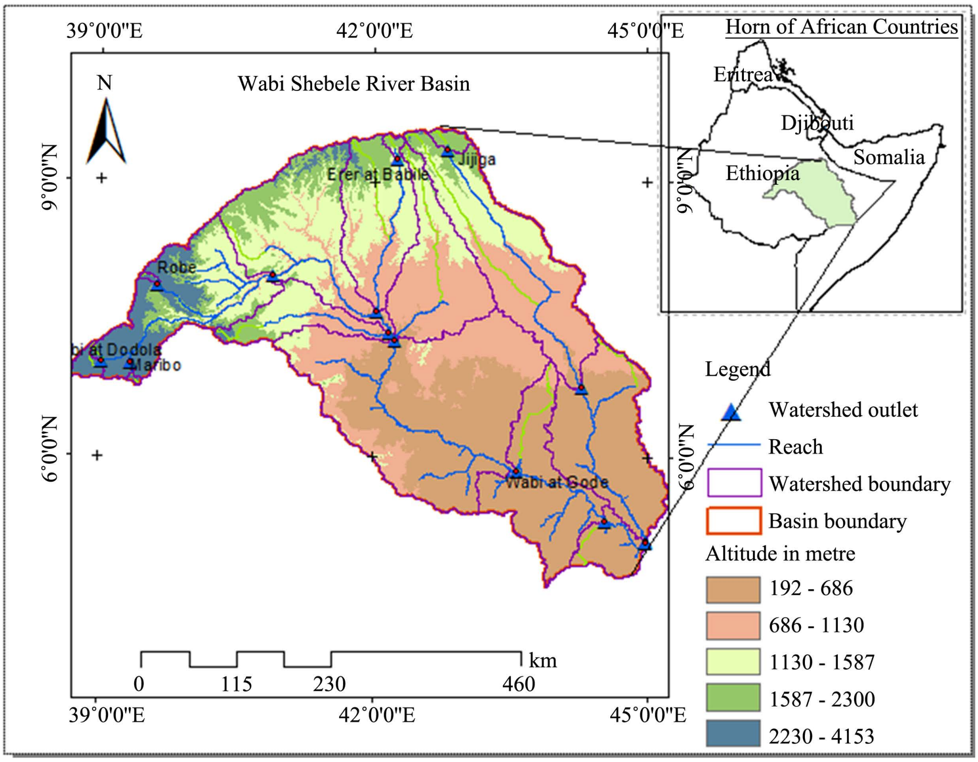


Figure 1. Study area map.

website [23]. The soil data with the resolution of 1 km mainly include soil texture, soil depth, and soil drainage attributes needed for the SWAT model will be derived from Harmonized World Soil Databasev1.2, a database that combines existing regional and national soil information in combination with information provided by FAO-UNESCO soil map. Land-use and land cover data were obtained from Ministry of Water, Irrigation and Energy.

From the data described above, the hydrologic response units (HRU) were established. After HRU definition, the data from climatic stations located in the study basin were inserted on the SWAT model. These data refer to rainfall (mm), maximum and minimum air temperatures (˚C), relative humidity (%), solar radiation (KJ·m^2) and wind speed (m/s). These data were obtained from the National Meteorology Agency (NMA). The data sets provide daily observations for stations exist in the basin. Table 1 shows fourteen weather stations selected in the study, which has good quality and spatial distribution in the basin with a minimum record length of 10 years in between 1980 to 2010. Accordingly, five station from upper catchments; six stations from middle catchments; two stations from Eastern catchments and one station from lower catchments.

Table 2 shows measured discharge data used for model calibration and validation, collected from hydrology department of the Ministry of Water and Energy, Ethiopia (MoWRIE). According to the Ministry of Water and Energy, the measurements of river levels follow the guidelines of the World Metrological Organization (WMO) [10].

Table 1. Meteorological data stations used in SWAT model.

S. No	Station Name	Controller	Coordinates*		Altitude (m)	Station aspect
			Latitude	Longitude		
1	Adaba	NMA	543,691	773,113	2420	NW
2	Dodola	NMA	519,746	772,054	2580	NW
3	Kofele	NMA	479,397	781,976	2620	NW
4	Merero	NMA	538,334	822,503	2940	NW
5	Hawassa	NMA	443,083	781,757	1750	NW
6	Robe (Arsi)	NMA	568,694	868,627	2400	N
7	Sinana	NMA	624,339	777,736	2400	M
8	Deder	NMA	767,132	1,032,669	2350	M
9	Jara	NMA	661,100	804,664	1960	M
10	Haromaya	NMA	832,842	1,040,832	2125	N
11	Gursum	NMA	873,588	1,034,746	1900	M
12	Jijiga	NMA	915,127	1,033,159	1775	E
13	Gode	NMA	1,006,927	654,440	295	S
14	Degehabour	NMA	1,001,635	911,616	1070	E

*UTM Adindan 1984, Zone 37N, NW = North western, N = North, M = Middle, E = East, S = South.

Table 2. Discharge data used for calibration and validation of SWAT model.

S. No	Station Name	Controller	Coordinates*		Altitude (m)	Catchment area (km^2)
			Latitude	Longitude		
1	Wabi @ Dodola	MoWIE	775,521	503,682	-	1040
2	Maribo @ Adaba	MoWIE	773,692	536,818	-	192
3	Wabi @ Legahida	MoWIE	881,015	709,436	-	19793
4	Erer Nr. Babile	MoWIE	1,021,721	197,816	-	494
5	Wabi @ Gode	MoWIE	654,138	341,331	-	124,108
6	Fafen @ Jijiga	MoWIE	1,034,208	258,365	-	731

*UTM Adindan 1984, Zone 37N.

3. Methodology

The methods implemented in this paper comprises a semi distributed macroscale hydrological modeling for the simulation of daily runoff, implementation of Exploratory data analysis (EDA) to explore simulated data, data distributions and examine clusters in the data or relationships between variables and/or sample locations and a nonparametric trend test to detect trends in annual and seasonal maximum runoff.

3.1. Hydrological Model

Soil and Water Assessment Tool (SWAT) is a continuous time, spatially distributed model designed to simulate water, sediment, nutrient and pesticide transport at a catchment scale on a daily time step. In this study, the model was used to generate flows. The model is driven by metrological data like precipitation, temperature, relative humidity, solar radiation and wind speed and physiographic data of pedology, land use and occupation, and classes of slopes.

It uses hydrologic response units (HRUs) that consist of specific land use, soil and slope characteristics. The HRUs are used to describe the spatial heterogeneity in terms of land cover, soil type and slope class within a watershed [24]. The model estimates relevant hydrologic components such as evapotranspiration, surface runoff and peak rate of runoff, groundwater flow and sediment yield for each HRUs unit.

The water in each HRU in SWAT is stored in four storage volumes: snow, soil profile (0 - 2 m), shallow aquifer (typically 2 - 20 m), and deep aquifer. Surface runoff from daily rainfall is estimated using a modified SCS curve number method, which estimates the amount of runoff based on local land use, soil type, and antecedent moisture condition. Peak runoff predictions are based on a modification of the Rational Formula [25]. The watershed concentration time is estimated using Manning's formula, considering both overland and channel flow. The SCS curve number is described by Equation (2).

$$Q_{Surf} = \frac{\left(P_i - I_a\right)^2}{P_i - I_a + S} = \frac{\left(P_i - 0.2S\right)^2}{P_i + 0.8S} \tag{2}$$

In which, Q_{surf} is the accumulated runoff or rainfall excess (mm/day), P_i is the rainfall depth for the day (mm), I_a is the initial abstraction lost from canopy interception, surface storage, and infiltration prior to runoff (mm H_2O; commonly approximated as 0.2S), S is the retention parameter (mm). The retention parameter is defined by Equation (3).

$$S = 25.4\left(\frac{1000}{\text{CN}} - 1\right) \tag{3}$$

The SCS curve number is a function of the soil's permeability, land used and antecedent soil water conditions. Specifically, the CN values are based on the hydrologic soil group of the area, land use, management, and initial hydrologic condition; with the hydrologic soil group and land use being the most important variables.

For climate, SWAT uses the data from the station nearest to the centroid of each sub basin. Calculated flow, sediment yield, and nutrient loading obtained for each sub basin are then routed through the river system. Channel routing is simulated using the variable storage or Muskingum method. The soil percolation component of SWAT uses a water storage capacity technique to predict flow through each soil layer in the root zone. Downward flow occurs when field capacity of a soil layer is exceeded and the layer below is not saturated. Percolation from the bottom of the soil profile recharges the shallow aquifer. If the soil temperature in a particular layer reaches less than or equal 0˚C, no percolation is allowed from that layer. Groundwater flow contribution to total stream flow is simulated by routing a shallow aquifer storage component to the stream [26] [27]. The model computes evaporation from soils and plants separately. Potential evapotranspiration can be modelled with the Penman-Monteith [28], Priestley-Taylor [29], or Hargreaves methods [30], depending on data availability. In this study, the Penman-Monteith method was used to determine potential evapotranspiration.

3.2. Defining an Extreme Event

In this paper six extreme hydrologic indices: Annual maximum discharge (AMAX), Peak over threshold (3rd quartile) frequency (POTF), Peak over threshold (3rd quartile) magnitude, seasonal maximum discharge for winter (SMW), Seasonal maximum discharge for spring (SMSp) and Seasonal maximum discharge for summer (SMSu) are used to define extreme high discharges.

In extreme value analysis ensuring independence of samples is initial task. In this study the time interval approach is used to ensure the independence of flow discharges. Time intervals between 5 to 14 days between successive peaks; 5 days for catchments < 10,000 km^2 and 14days for catchments ≥ 10,000, is used in this study. This approach is reported, a strong flood-frequency estimations approach e.g., [31] [32].

3.3. Flood Change Detection and Attribution

3.3.1. Flood Change Detection

Two distribution-free (nonparametric, e.g., rank-based) tests and exploratory data analysis (EDA) are used to detect changes in flood discharge. Exploratory data analysis (EDA) is used as preliminary analysis to explore initial hypotheses on changes in data time series to be confirmed by statistical analysis, nonparametric Mann-Kendall trend test to detect trends in flood discharges, and Quantile perturbation method (QPM) approach to see temporal variabilities in extreme discharges. In practice there is a continuum between "trend" and "change".

Exploratory data analysis (EDA): involves mainly plotting graphs, and allowed to explore some features in data and assess the first hypotheses to be confirmed by the statistical analysis. In addition, the linear regression gradient plot in the EDA allowed testing of potential trends data time series.

Mann-Kendall (MK) test: The MK test statistic, S is defined as [33] [34] and the test is conducted by applying Equation (4) to Equation (9):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}\left(X_j - X_i\right) \tag{4}$$

where X_j are the sequential data values, n is the length of the data set, and:

$$\operatorname{sgn}(\theta) = \begin{cases} 1 & \text{for } \theta > 0 \\ 0 & \text{for } \theta = 0 \\ -1 & \text{for } \theta < 0 \end{cases} \tag{5}$$

When $n \geq 8$, the statistic S is approximately normally distributed with mean and variance given by Mann [34] and Kendall [35]:

$$E[S]=0 \tag{6}$$

$$V(S)=\frac{n(n-1)(2n+5)-\sum_{i=1}^{n}t_1(l-1)(2l+5)}{18} \tag{7}$$

where t_1 is the number of ties of extent l. the standardized test statistic Z is computed by:

$$Z=\begin{cases}\frac{S-1}{\sqrt{Var(S)}} & \text{for } S>0\\ 0 & \text{for } S=0\\ \frac{S+1}{\sqrt{Var(S)}} & \text{for } S<0\end{cases} \tag{8}$$

The standardized MK statistic Z follows the standard normal distribution with a mean of zero and variance of one under the null hypothesis of no trend. A positive Z value indicates an upward trend, while a negative one indicates a downward trend.

The Kendall rank coefficient is often used as a test statistic to establish whether two variables may be regarded as statistically dependent. Under the null-hypothesis of independence of X_i and X_j, the sampling distribution of "tau" has an expected value of zero. The value of "tau" ranges from −1 (100% negative association, or perfect inversion) to +1 (100% positive association or perfect agreement). A value of zero indicates the absence of association.

The p-value of the MK statistic S of sample data can be estimated using the normal cumulative distribution function:

$$p=0.5-\varnothing(|Z|) \text{ where } \varnothing(|Z|)=\frac{1}{\sqrt{2\pi}}\int_0^{|Z|}\mathrm{e}^{\frac{r^2}{2}}\mathrm{d}t \tag{9}$$

If the p-value is small enough, the trend is quite unlikely to be caused by random sampling. At the significance level of 0.05, if $p \leq 0.050$, then the existing trend is assessed to be statistically significant.

Quantile Perturbation method (QPM): is a statistical analysis specially designed for analyzing extreme conditions, to study trends and multi time period oscillation patterns in hydro-climatic extremes [14] [35] [36] [37]. The method has two concepts: 1) the frequency aspect which focuses on how often an extreme event (quantile) may occur and, 2) the perturbation aspect which determines the changes in the extremes for a particular return period. The method compares the long-term baseline period extreme value quantiles with that of a selected sub-period quantile. The selected sub-period (block period) is a subseries taken from long-term baseline time series representing the period of interest. To select appropriate value of block length in between 5- and 15-year intervals, Tabari *et al.*, [37] recommends applying QPM to extreme time series to different block of years and select the one which shows a high variability at a given time interval. To check the significance of perturbation factor in extreme quantiles 95% CI (Confidence Interval) is computed using non-parametric bootstrapping method is performed. The perturbation values fall outside of confidence intervals are identified as statistically significant perturbation value.

3.3.2. Flood Change Attribution

Studying trend and variability in hydroclimatic time series alone has no value unless the cause or driving factors of changes are followed. Merz *et al*., [38] reported in their research that the main goal of trend and variability studies in hydroclimatic time series is to test hypothesis about the influence of driving forces of floods. The detected trends were explained by the correlation with changes in another variables; climate variables, environmental background condition variables and external factors. In addition, changes in catchment and environmental background condition factors, deforestation and reforestation and alteration in agricultural management practices are analyzed.

Six watersheds representing mountainous and plain regions of Wabi Shebele River Basin were selected to identify most powerful variables in estimating the mean peak flood (QMPF) of the basin from climate, watershed's physical characteristics and external factors. The mean peak flow (QMPF) was expressed in this study as the arithmetic mean value of peak over threshold (3rd quartile) flows for the period of record. The environmental background factors are extracted from delineated watershed using Arc Hydro interface in Arc GIS. Population density and land use and land cover data were collected from MoWIE. Daily precipitation data from the meteorological stations in Wabi Shebele Basin and the surrounding areas were obtained from the National Meteorology Service Agency (NMA). Detail of data used and sources are described in **Table 3**. To highlight the most powerful parameters in flood discharge characteristics PCA is used.

Principial Component Analysis (PCA): is used in this study to see multivariate relationships between potential driving factors and mean peak flow discharge (QMPF). If the number of predictor variables increases and they are highly correlated, Multiple linear regression (MLR) models gets more unstable. PCA is one of the multivariate statistical techniques that can be used to deal with highly correlated variables in regression [39] [40]. In PCA, different types of variables: hydrologic and watershed variables are treated together. The original dataset of n variables, which are correlated to various degrees are transformed to n numbers of uncorrelated PCs. The PCs are linear transformation of the original variables in such a way that the original and the new variables have equal sums of the variances. Although the number of PCs and original variables are equal, the first few PCs explain the majority of the variance in the data set, reducing the dimensionality of the original data set [39].

The PCs are sequenced from the highest to the lowest variance, *i.e.*, the first PC describes the data's highest variance proportion. The next highest variance is explained by the second PC and so on. The values of PCs can be obtained from Equations (10) and (11):

$$\text{PC1} = a_{11}x_1 + a_{12}x_1 + \cdots + a_{1n}x_n = \sum_{J=1}^{n} a_{1j}x_j \quad (10)$$

Table 3. Potential driving factors of flood changes analyzed in this study.

Attributes	Parameters	Abbreviations	Scale or resolution	Time	Data Source
Floods	River floods	QMPF	1:50,000	1980-2010	MoWIE
Climate factors	Maximum Daily precipitation	MDR	-	1980-2010	NMA
	Mean annual rainfall	MAR	-	1980-2010	NMA
Environmental background conditions	Drainage area	DA	1:50,000		Dem 90 m
	Bain slope	BS	1:50,000		Dem 90 m
	Basin elevation	BE	1:100,000		Dem 90 m
	Shape factor	SF	1:100,000		Dem 90 m
	Drainage density	DD	1:100,000		Dem 90 m
	Valley length	VL	1:100,000		Dem 90 m
	Valley slope	VS	1:100,000		Dem 90 m
	Elongation ratio	ER	1:100,000		Dem 90 m
	Soil textures (clay, sand and loam)	clay, sand, loam	1:100,000		FAO website
Climate factors	Maximum Daily precipitation	MDR	-	1980-2013	NMA
	Mean annual rainfall	MAR		1980-2013	NMA
External factors	Population density	PD	-	2004	MoWIE
	Farmland	AGR	-	2004	MoWIE
	Forest	forest	-	2004	MoWIE

$$\text{PC2} = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = \sum_{J=1}^{n} a_{2j}x_j \tag{11}$$

where $x_1, x_2, \cdots, x_n$ are the original variables and a_{jj} are the eigenvectors. The eigenvalues are the variances of the PCs. The covariance or correlation matrix of the data set is used to derive the coefficients a_{jj}, which are the eigenvectors. The eigenvalues of the data matrix can be calculated by Equation (12):

$$|C - \lambda I| = 0 \tag{12}$$

where C is the correlation/covariance matrix, λ is the eigenvalue, and I is the identity matrix. The PC coefficients or the weights of the variables in the PCs are then calculated by Equation (13):

$$|C - \lambda I| a_{jj} = 0 \tag{13}$$

3.4. Flood Frequency Analysis

Parallel to changes in magnitudes, changes in the frequency of extreme events, which can be described by the occurrence rate has significant role in flood events. Changes in the occurrence rate reflect the clustering properties of flood events, caused by natural variability, regime shifts in the atmosphere [40] [41] or land use changes. For each sample watersheds, peak over threshold (3rd quartile) time series was extracted to see trends in magnitude and number of occurrences in annual and seasonal time level.

4. Results and Discussion

4.1. Calibration and Validation of the Hydrological Model

For model calibration and validation, the observed daily and monthly streamflow data were used from 1988 to 2000 with three years warming period. To evaluate the model performance, three parameters have been used, namely R^2 and NSE and P-bias. NSE is a normalized statistic, ranges from $-\infty$ to 1, used to indicate the relative value of residual variance compared to the variance of the observed data and values close to one shows a perfect match of the modeled with the observed data [42], Equation (10). R^2 is the proportion of the total variance in the observed data that can be explained by the model, Equation (11).

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(X_i - Y_i\right)^2}{\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2} \tag{14}$$

$$\text{R}^2 = \left[\frac{\sum_{i=1}^{N}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2}}\right]^2 \tag{15}$$

$$\text{Pbias} = \frac{\sum_{i=1}^{N}\left(Y_i - X_i\right)}{\sum_{i=1}^{N} X_i} * 100 \tag{16}$$

where x and y are observed and modelled streamflow, respectively, N is the number of data pairs. Table 4 shows statistical evaluation of model performance. Based on the model the physiographic characteristics of watersheds are presented in Table 5.

4.2. Annual Maximum Discharge

The presence of seasonal cycles and clusters in annual maximum discharges (AMAX) are examined using exploratory data analysis (EDA) (Figure 2). The magnitude of annual maximum discharge oscillates at 5 - 10-year intervals in

Table 4. Evaluation of model performance.

Stations	Area (km^2)	Location		Average Annual Flow (Mm^3)	Calibration			Validation		
		Lat	Long		R^2	NSE	Pbias (%)	R^2	NSE	Pbias (%)
Wabi at D/Bridge	1040	7.01	39.02	230.9	0.7	0.7	−3.0	0.7	0.7	−2
Maribo	192	7.00	39.20	100.2	0.5	0.6	−19	0.4	0.5	−29
Robe	169	7.51	39.38	48.5	0.5	0.4	−30	0.4	0.4	−21
Wabi at L/Hida	19,793	7.58	40.54	1848.5	0.6	0.6	−0.9	0.6	0.7	1
Erer	494	9.14	42.15	87.5	0.4	0.4	−0.2	0.1	−0.4	−54
Jijiga	731	9.21	42.48	35.4	0.2	0.5	5.3	0.1	0.2	−59
Wabi at Gode	124,108	5.56	43.33	4523.2	0.4	0.2	−29	0.2	0.1	−38

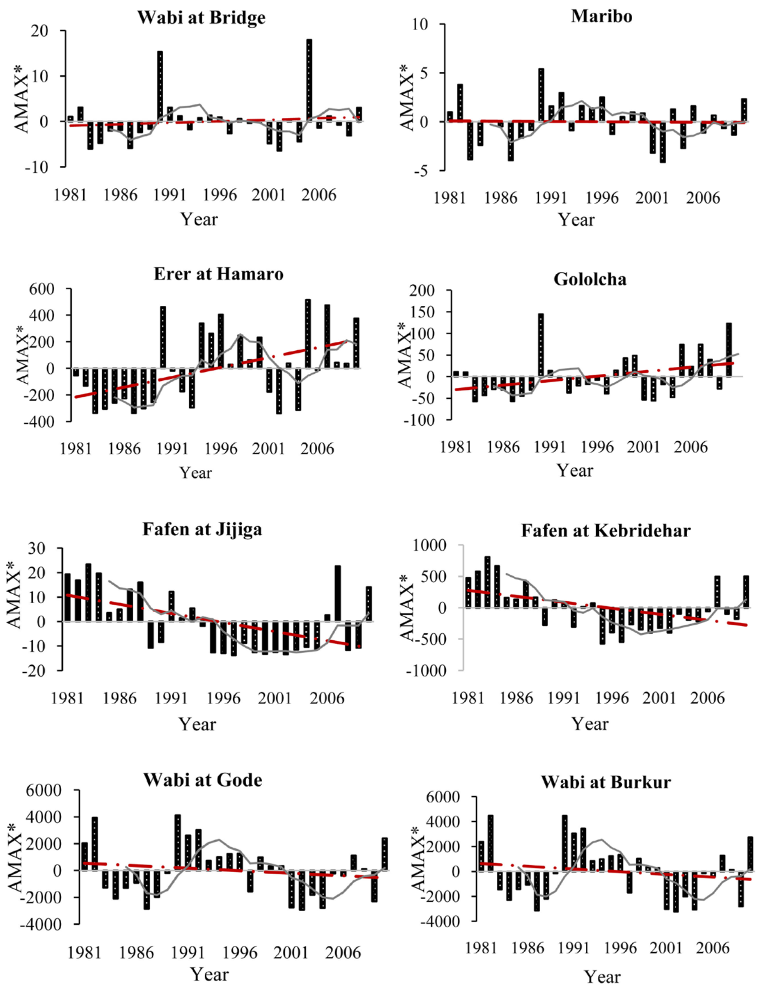


Figure 2. Standardized annual maximum discharge (AMAX*) averaged over the studied stations for the periods 1981-2010. Broken line is the linear trend, grey curve is the 5-year moving average in each sample sub basins.

Table 5. Physiographic characteristics of the 11 studied sub basins.

S. No	Station (River)	Highest elevation (m.a.s.l)	Lowest elevation (m.a.s.l)	Relief (m)	Mean elevation (m.a.s.l)	Catchment area (Km^2)	Catchment Aspect	Mean catchment slope	Time Period analyzed
1	Wabi @ Dodola	3119	2473	646	2618	1040	NW	0.11	1981-2010
2	Maribo @ Adaba	3743	2350	1393	3073	192	NW	0.14	1981-2010
3	Robe @ Robe	4055	2404	1651	2836	169	N	0.10	1981-2010
4	Wabi @ Legehidha	4153	765	3388	2066	19,793	M	0.12	1981-2010
5	Erer @ Babile	3004	1308	1696	1960	469	NE	0.18	1981-2010
6	Erer @ Hamaro	3388	503	2885	1357	14,760	M	0.14	1981-2010
7	Gololcha at Junction	2686	460	2226	1369	7139	M	0.03	1981-2010
8	Fafem @ Jijiga	2482	1634	848	1813	910	NE	0.02	1981-2010
9	Fafem @ Kebridehar	3013	513	2500	1153	24,956	SE	0.03	1981-2010
10	Wabi @ Gode	3374	258	3116	916	124,108	S	0.105	1981-2010
11	Wabi @ Burkur	977	215	762	428	146,804	S	0.10	1981-2010

Relief is calculated as the difference between the highest and the lowest elevation, Aspect is the averaged aspect of the basin (N North; NE North-East; S South; SW South-West).

most of sample stations. Annual maximum discharges indicate less than mean annual maximum values in 1990s for the middle and eastern catchments of Wabi Shebele Basin. However, annual maximum discharge in 2000s is observed above mean maximum annual discharge in all gauging stations entire the basin.

The MK test applied at each site for the period 1981-2010 showed less decreasing trend in upper and lower catchments. For the longest period, 55% of the stations indicate weak to significant decreasing trends in annual maximum discharge. However, the watersheds in middle basin indicate a majority of significant increasing trend annual maximum discharge, $p < 0.05$. Too see multi-temporal changes in annual maximum discharge Mann-Kendall trend tests are analyzed at 5, 10, 15, 25 and 30 years intervals.

From both Figure 2 & Figure 3, some conclusions are extracted on multi-temporal trend analysis. First, blue colors are more frequent than red colors,

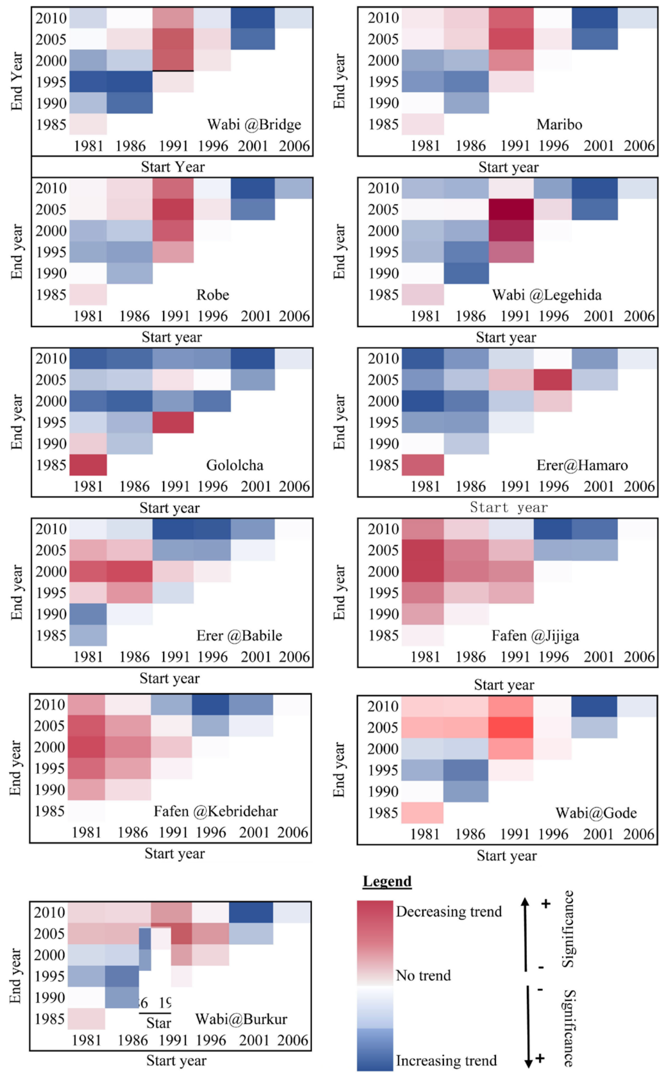


Figure 3. Multi-temporal trend analysis for the annual maximum discharge (AMAX) for river catchments in Wabi Shebele River Basin. Blue and red cells correspond to positive and negative tau values respectively (the darker the color the more significant the trend).

meaning that positive trends are more frequent than negative, especially for the most recent period since 1996 in all stations (**Figure 3**). Second, negative trends (red colors) appear more significantly in eastern Fafen catchments and lower stations of Wabi Shebele River Basin before 2000 (as shown on the Fafen @ Jijiga, Fafen @ Kebridehar, Wabi @ Gode and Wabi @ Burkur). Most of gauging stations in upper and middle catchments indicates weak to significant increasing trends and while stations in lower and eastern catchments showed weak to strong decreasing tendencies in annual maximum discharge during the past 30 years in Wabi Sheble River Basin. The years, 1980s and 2000s are the decades were weak to significant increasing flood discharges are occurred. For 1990s, most stations indicate decreasing trends in annual maximum discharge in the study area.

4.3. Flood Frequency and Seasonality

In Wabi Shebele Basin there are nine mean annual flood events, discharges greater than third quartile (3rd quartile). Some years (*i.e.*, 1981, 1982, 1983, 1990, 1991, 2006, 2007 and 2010) were particularly rich in flood events, with more than 14 events on average. In contrast, in some other years (*i.e.*, 1997, 2001, 2002 and 2004) only two to four events (POTF), was observed in the basin. The decades 1981-1990 and 2001-2010 are noticed as the decades of the richest flood events.

The averaged POTF in upper and middle catchments reveals increasing trends of flood events for longest time period, 1981-2010 (**Figure 4**). However average POTF in lower stations of Wabi Shebele Basin and Fafen catchments indicates decreasing trends in number of events. At sample gauging stations the result of MK trend test for the long-term period (1981-2010) reveals positive trends of POTF for 55% of gauging stations. Among these significant increasing trends in

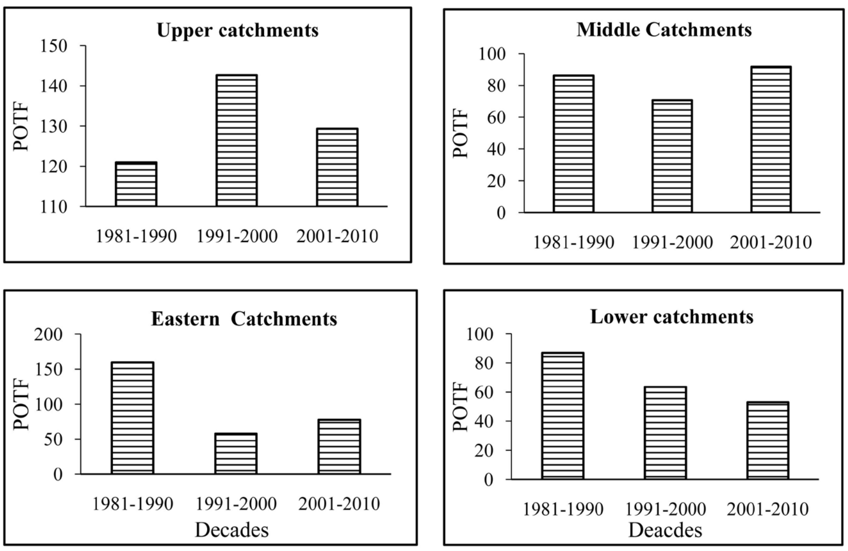


Figure 4. Decadal average POTF for periods 1981-2010 in different subbasins of Wabi Shebele.

POTF is observed over middle catchments (Erer watershed at Hamaro, and Gololcha) and the rest 45% of stations indicates negative trends, mainly present in Fafen watershed and Wabi Shebele River at Gode and Burkur.

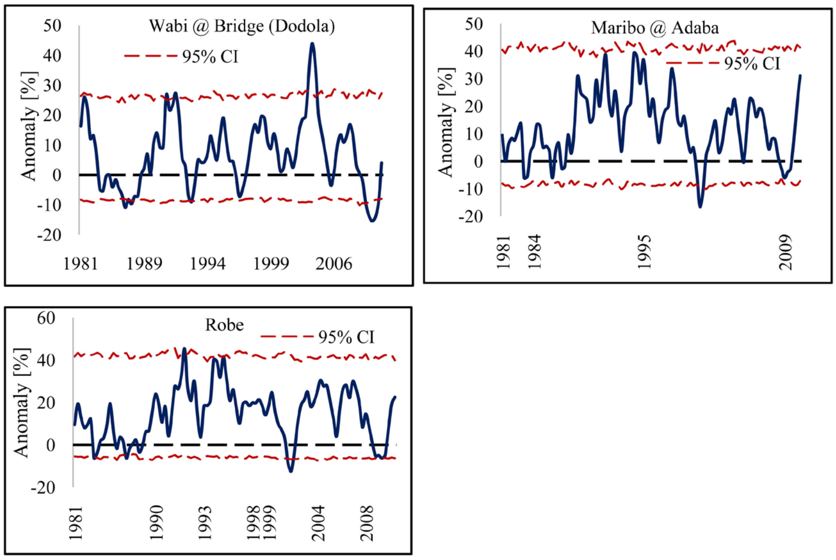
From **Figure 5**, the QPM analysis using peak over threshold (3rd quartile) indicates that most of extreme flows varies within a confidence with high oscillation patterns in the entire the basin. Extreme flows vary above reference line (above mean) in upper, middle and lower valley of Wabi Shebele River stations, whereas it varies below reference line in eastern catchments.

Figure 6, illustrates the importance of time windows in analysis of seasonal flood trends. The EDA of spring and winter discharge does not present trends in the longest period (1981-2010) in most catchments of the basin. Almost in all sample catchments, seasonal maximum discharge indicates oscillation pattern at a decade interval. Like annual maximum discharge, maximum discharge in all seasons also indicates less than mean seasonal maximum discharge in 1990s.

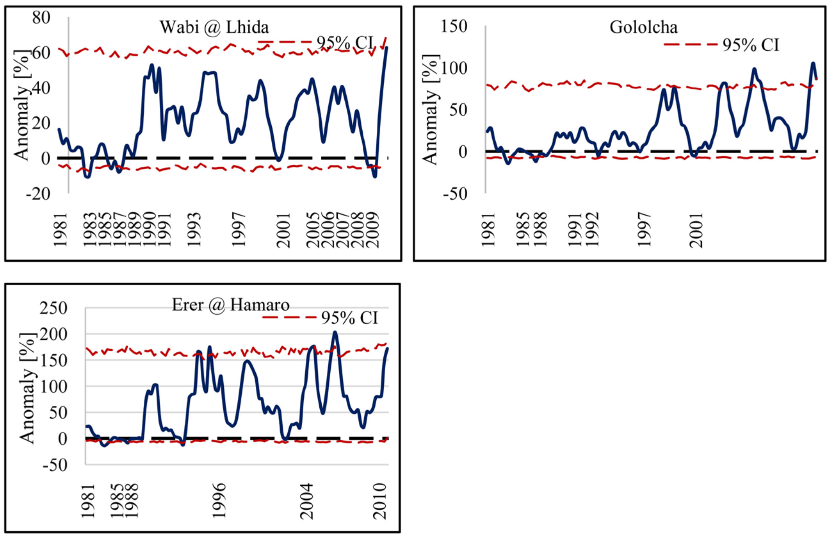
From **Figure 7**, the multi-temporal analysis in seasonal flood discharge indicates similar patterns in summer and winter and different in spring are observed in throughout the basin. In Fafen catchments spring flood discharge indicates weak increasing trend while indicates decreasing trend in summer and winter season for the last 30 years. Darker colors show statistically significant trends, and they are more frequent in summer and less in winter in eastern catchments. For 1990s, upper and lower Wabi Shebele Basin flood discharge indicates significant decreasing trends in all seasons and weak decreasing trend in eastern catchments. In all catchments flood discharge indicates increasing tendency in 2000s. This pattern is stronger in summer and winter when 72% of stations showed significant increasing trends for the period 2001-2010. The same result was previously observed in the Wabi Shebele River Basin, where a significant increase in spring, summer and winter floods was identified [22]. **Table 6**, illustrates seasonal significant anomalies in extreme discharges investigated using QPM.

Table 6. Characteristics of highest anomaly in seasonal extreme discharges.

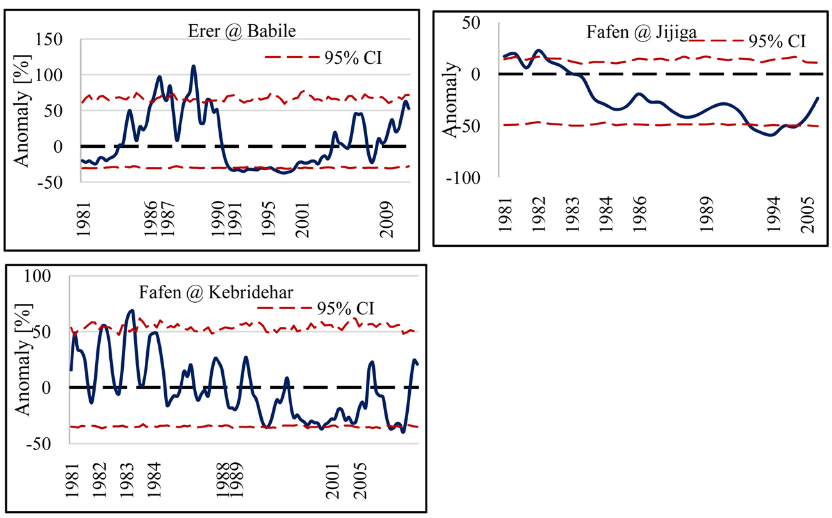
Station/River name	Highest +ve anomaly (%)	Season of occurrence	Year	Highest −ve anomaly (%)	Season of occurrence	Year
Wabi at Dodola	135	spring	1989	−29.7	summer	1983
Maribo	177	spring	1987	−60.5	spring	2004
Robe	193.2	spring	1987	−62.7	spring	2004
Wabi at Legehida	114.2	spring	1987	−49.1	spring	2004
Gololcha	116.5	spring	1987	−49.2	spring	2004
Erer at Hamaro	455.5	winter	2006	−57.4	spring	2004
Erer at Babile	296.8	winter	2006	−50.5	spring	1997
Fafen at Jijiga	177	spring	1981	−65.7	winter	1996
Fafen at Kebridehar	357.6	summer	1984	−65.2	spring	1996
Wabi at Gode	439.6	summer	1990	−61.5	summer	2001
Wabi at Burkur	438.7	summer	1990	−61.1	summer	2001


Wabi @ Bridge (Dodola)
95% CI
Anomaly [%]
Maribo @ Adaba
95% CI
Anomaly [%]
Robe
95% CI
Anomaly [%]

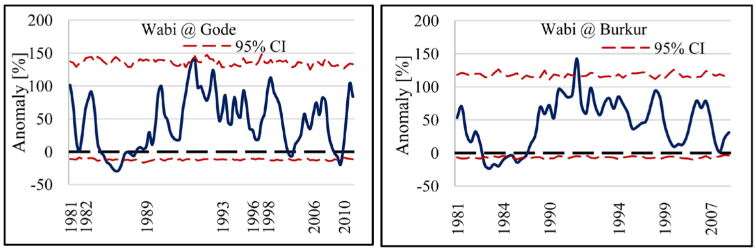
(a)


Wabi @ Lhida
95% CI
Anomaly [%]
Gololcha
95% CI
Anomaly [%]
Erer @ Hamaro
95% CI
Anomaly [%]

(b)


Erer @ Babile
95% CI
Anomaly [%]
Fafen @ Jijiga
95% CI
Anomaly
Fafen @ Kebridehar
95% CI
Anomaly [%]

(c)



(d)

Figure 5. Temporal variability in extreme flood discharge using QPM with 95% CI (Confidence Interval) in Wabi Shebele River Basin using four different categories: (a) Upper catchments, (b) Middle catchments, (c) Eastern catchments and (d) Lower catchments.

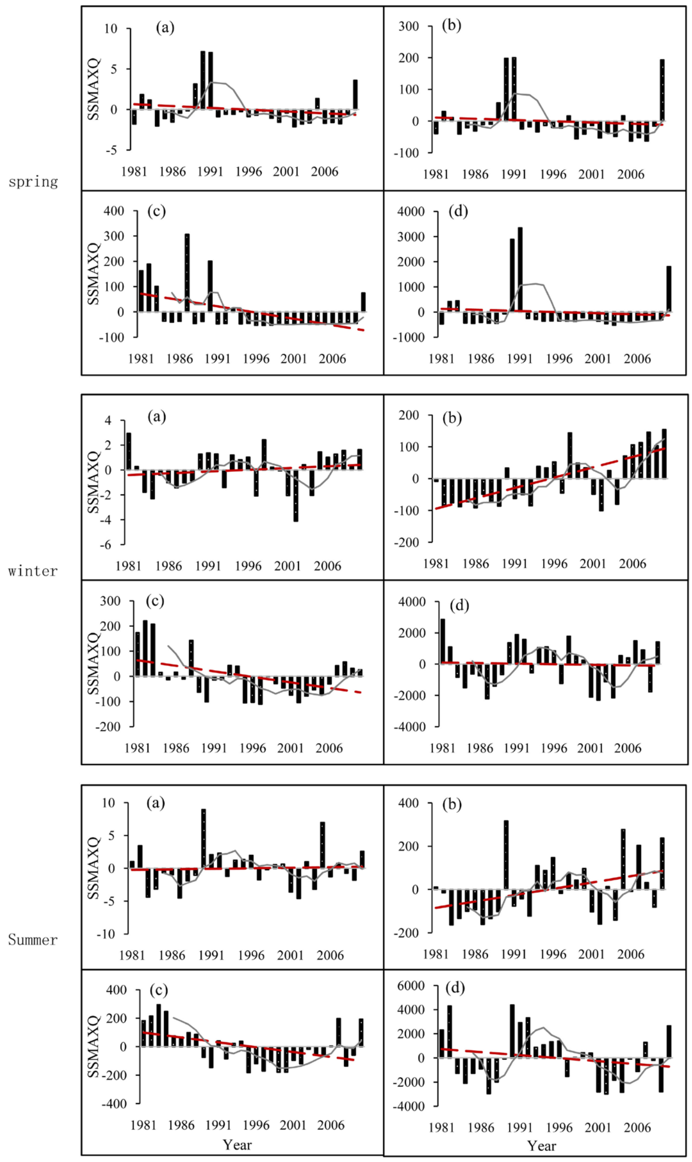


Figure 6. Standardized seasonal maximum discharge (SSMAXQ) averaged over the studied catchments for period 1981-2010: (a) Upper catchments, (b) Middle catchments, (c) Eastern catchments, (d) Lower catchments. red broken line is the linear trend and grey curve is the 5-year moving average.

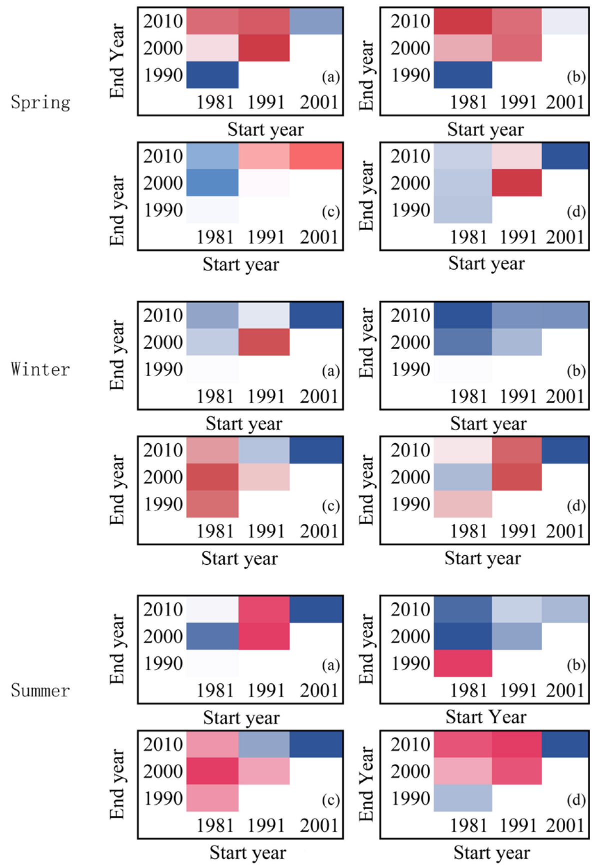


Figure 7. Multi-temporal trend analysis of seasonal maximum discharges for: (a) upper catchments, (b) middle catchments, (c) eastern catchments and (d) lower catchments. Legend is explained in Figure 3.

4.4. Summary of Flood Changes

The presented results revealed that there has been an increasing trend in the flood magnitude and frequency over early 21st century entire the basin. Similar result is reported in literatures that the magnitude and frequency of floods indicates increasing trend in Wabi Shebele River Basin since 2000 [10] [11] [22]. The positive Kendall's Z values indicates increasing trend within analysis period (Table 7). For the period 1981 to 2010 the annual maximum flood discharge shows upward trends in upper and middle catchments while downward trends in eastern and lower catchments of Wabi Shebele Basin. The annual maximum stream flow for middle catchments (*i.e.*, Erer at Hamaro and Gololcha at Wabi junction) shows a positive significant trend because the computed *p*-value is in both watersheds are lower than the significance level ($\alpha = 0.05$) in the region. However, significant decreasing trends in annual maxima are observed in Fafen watersheds at Jijiga and Kebridehar gauging stations. Seasonal trend analysis reveals similar trends and patterns with annual maximum stream flow almost in all stations during past 30 years.

Extreme discharge variability analysis using peak over threshold (3rd quartile) based on QPM showed significant increasing trend in early 1990s & 2000s and

Table 7. Mann Kendall trend test summary extreme flood discharges in Wabi Shebele River Basin.

Extreme Indices	Mann-Kendall Statistics	Upper catchments			Middle catchments			Eastern catchments			Lower catchments	
		Wabi @ Dodola	Maribo	Robe	Wabi @ Legehida	Erer @ Hamaro	Gololcha @ junction	Erer @ Babile	Fafen @ Jijiga	Fafen @ Kebridehar	Wabi @ Gode	Wabi @ Burkur
AMAX	Test statistcs (Z)	0.32	−0.18	−0.07	0.57	2.32	2.00	0.32	−2.64	−2.32	−0.82	0.41
	p-value (two-tailed)	0.75	0.86	0.94	0.57	0.02	0.05	0.75	0.01	0.02	0.41	0.45
	Mann-Kendall Stat (S)	19.0	−11.0	−5.0	33.0	131.0	113.0	19.0	−149.0	−131.0	−47.0	−43.0
	Kendall's tau	0.04	−0.03	0.09	0.08	0.30	0.26	0.04	−0.34	−0.30	−0.11	−0.10
	Alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	Trend	Not Significant	Not Significant	Not Significant	Not Significant	Increasing	Increasing	Not Significant	Decreasing	Decreasing	Not Significant	Not Significant
POTF	Test statistcs (Z)	1.471	0.969	1.149	0.997	2.41	2.046	−0.895	−3.207	−2.478	−1.208	−1.206
	p-value (two-tailed)	0.141	0.333	0.251	0.319	0.016	0.041	0.371	0.001	0.013	0.227	0.228
	Alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	Trend	Not Significant	Not Significant	Not Significant	Not Significant	Increasing	Increasing	Not Significant	Decreasing	Decreasing	Not Significant	Not Significant
SMW	Test statistcs (Z)	0.77	1.14	1.25	2.36	3.96	2.89	0.57	−2.36	−1.64	−0.04	−0.29
	p-value (two-tailed)	0.44	0.25	0.21	0.02	0.00	0.00	0.57	0.02	0.10	0.97	0.78
	Mann-Kendall Stat (S)	44.0	65.0	71.0	133.0	223.0	163.0	33.0	−133.0	−93.0	−3.0	−17.0
	Kendall's tau	0.10	0.15	0.16	0.31	0.51	0.38	0.08	−0.31	−0.21	−0.01	−0.04
	Alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	Trend	Not Significant	Not Significant	Not Significant	Increasing	Increasing	Increasing	Not Significant	Decreasing	Not Significant	Not Significant	Not Significant
SMSp	Test statistcs (Z)	0.82	−1.71	−1.82	−1.57	−1.61	−1.57	−2.61	−2.14	−2.53	0.57	0.57
	p-value (two-tailed)	0.41	0.09	0.07	0.12	0.11	0.12	0.01	0.03	0.01	0.57	0.57
	Mann-Kendall Stat (S)	47.0	−97.0	−103.0	−89.0	−91.0	−89.0	−147.0	−121.0	−143.0	33.0	33.0
	Kendall's tau	0.11	−0.22	−0.24	−0.21	−0.21	−0.21	−0.34	−0.28	−0.33	0.08	−0.08
	Alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	Trend	Not Significant	Not Significant	Not Significant	Not Significant	Not Significant	Not Significant	Decreasing	Decreasing	Decreasing	Not Significant	Not Significant
SMSu	Test statistcs (Z)	0.39	−0.18	0.00	0.54	2.21	1.68	0.00	−2.25	−2.21	−0.82	−0.89
	p-value (two-tailed)	0.70	0.86	1.00	0.59	0.03	0.09	1.00	0.02	0.03	0.41	0.37
	Mann-Kendall Stat (S)	23.0	−11.0	−1.0	31.0	125.0	95.0	−1.0	−127.0	−125.0	−47.0	−51.0
	Kendall's tau	0.05	−0.03	0.00	0.07	0.29	0.22	0.00	−0.30	−0.29	−0.11	−0.12
	Alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	Trend	Not Significant	Not Significant	Not Significant	Not Significant	Increasing	Not Significant	Not Significant	Decreasing	Decreasing	Not Significant	Not Significant

decreasing trends in 1980s particularly in upper and middle catchments (Table 8). Over eastern catchments 1980s is the decade in which significant increasing trends observed and decreasing trends in 1990s and 2000s. The lower Wabi Shebele river stations indicate general decreasing trend in analyses period, 1980-2010.

In seasonal extreme variability analysis, significant anomaly occurrence season varies with catchments as presented in Table 6. In upper and middle catchments (Wabi at Dodola, Maribo, robe Wabi at Legehida and Gololcha watersheds), spring season is the season in which highest extreme variabilities are occurred. Similarly, in eastern catchments (Erer watersheds) highest extreme discharge anomalies are occurred in winter season and in lower Wabi Shebele catchments (at Gode and Burkur stations) during summer season. The years; 2nd half of 1980s, 1st half of 1990s and 2nd half of 2000s are the years of positive anomalies (significant increasing trends) in flood discharges. Whereas, the years; 1st half of 1980s, 2nd half of 1990s and 1st half of 2000s are the years of significant negative anomalies occurrence in all season. It is known that, the Wabi Shebele Basin is characterized by two rainfall regimes [43]: the area characterized by a quasi-double maximum rainfalls pattern with a small peak in April and maximum peak in August which covers the west-east highland of the basin (bimodal type I); and the area dominated by double maximum rainfall pattern with peaks during April and October covers the south-eastern low-lying areas of the basin (bimodal type II).

Table 8. Summary of QPM analysis in annual extreme flow.

Upper catchments			Middle catchments			Eastern Catchments			Lower catchments		
Sub basin	Magnitude of highest anomaly (%)	Time	Sub basin	Magnitude of highest anomaly (%)	Time	Subbasin	Magnitude of highest anomaly (%)	Time	Sub basin	Magnitude of highest anomaly (%)	Time
Wabi at Dodola	−10.9	1986	**Wabi at Legehida**	−6	1986	**Erer at Babile**	+96.9	1986	Wabi at Gode	−29.4	1983
	+27.3	1991		−7.9	1987		+112.3	1988		−19.4	2009
	+43.9	2004		−10	2009		+65.1	1989	Wabi at Burkur	−23.4	1983
	−15.4	2009	**Gololcha**	−14.6	1983		−34.5	1992		−14.2	1985
Maribo	−16.5	2000		−11.8	1987		−36.7	1998		+142.3	1991
Robe	−6.03	1982		+81	2005	**Fafen at Jijiga**	+19.4	1981			
	−6.3	1987		+98.4	2006		+22.5	1982			
	+45.3	1991		+105.4	2010		−58.6	1994			
	−12.3	2001	**Erer at Hamaro**	−14.2	1983	**Fafen at Kebridehar**	+68.4	1983			
				−8.5	1987		−37.02	1998			
				+165.3	1994		−39.6	2009			
				+174	1995						
				+203.6	2006						

4.5. Potential Drivers of Flood Changes

4.5.1. Principal Component Analysis (PCA)

PCA of flood peak discharge with major driving forces like climate drivers, environmental background conditions and external factors is performed to highlight most important variables in Wabi Shebele Basin. Therefore, PCA is applied to eleven selected predictors including those which have no significant correlation with each other (less collinearity), *i.e.*, Drainage area (DA), Basin elevation (BE), Basin slope (BS), Shape factor (SF), valley slope (VS), clay, sand, loam, Mean Annual Rainfall (MAR), forest, Agricultural land (AGR) and Population Density (PD) to achieve uncorrelated six PCs. The eigenvalues represent the quantity of variability in the data and they are presented in **Table 9**. **Table 9** confirms that the first two PCs explain the maximum degree of variability of the data set with the proportion of 45.6% and 28.7%, respectively. The proportions of other PCs (PC3, PC4, PC5 and PC6) range 0.0% - 17.3%. Both PC1 and PC2 accounts for 74.3% variance, meaning more than 2/3 of variability in dataset is explained in the first two PCs. To explain at least 90% of variation in the data the first three components are used.

The coefficients in **Table 9** show the linear combinations variables that make each principal component. Absolute values near zero indicate that a variable

Table 9. Principal correlation analysis: Eigen analysis of the correlation matrix.

Name	PC1	PC2	PC3	PC4	PC5	PC6
Eigenvalue	5.931	3.727	2.250	0.727	0.365	0.000
Proportion	0.456	0.287	0.173	0.056	0.028	0.000
Cumulative proportion	0.456	0.743	0.916	0.972	1	1
Eigen Vectors (coefficient)						
Variables	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**
QMPF in m^3/s	0.327	0.270	−0.187	−0.047	0.205	0.033
DA in km^2	0.398	0.092	−0.099	0.000	0.122	0.149
BE in m	−0.398	−0.012	−0.062	−0.260	−0.063	−0.718
BS in m/km	−0.087	0.496	0.117	−0.013	0.141	−0.070
SF	−0.089	0.344	−0.282	0.668	0.165	−0.252
VS m/km	−0.368	0.027	0.178	0.351	−0.299	0.455
clay	−0.159	−0.412	−0.245	0.072	0.460	0.023
sand	0.295	−0.020	0.367	0.443	−0.318	−0.372
loam	−0.017	0.477	−0.026	−0.380	−0.352	0.080
MAR in mm	−0.295	0.268	−0.310	−0.007	0.068	0.086
forest	−0.391	−0.125	−0.064	0.074	−0.248	0.121
AGR	−0.260	0.263	0.346	0.026	0.448	0.127
PD in Pop/km^2	−0.075	−0.007	0.641	−0.093	0.313	−0.050

contributes little to the component, whereas larger absolute values indicate variables that contribute more to the component. In the analysis, first principal component has large negative associations with BE, VS, MAR and forest, so this component primarily measures the basin altitude difference and land cover. The second component has large positive associations with BS, SF and loam, so this component primarily measures the slope and shape of the catchment. The third component has large positive association with sand, AGR and PD, so this component primarily measures the basin farm land and population density.

The loading plot in Figure 8, visually shows the results for the first two components. From the graph, DA and sand indicates small angle (<90˚) from QMPF line, meaning the variables positively correlated to QMPF. The variables: forest, PD, BE and VS indicates angles related to 180˚, meaning they are significantly negatively correlated to QMPF. Whereas the variables: BS, SF and loam has no significant correlation with QMPF in Wabi Shebele River Basin.

4.5.2. Climate Drivers

Most of the rivers in Ethiopia exhibit typical characteristics of tropical rainfall-dependent flow regimes. According to previous studies, the spatial and temporal distribution of rainfall governs amount and intra- and inter-annual variability of water availability in Ethiopia [6]. General understanding of hydroclimatic variables such as precipitation, temperature and discharge is important for water resource planning and management. Among this temperature is a factor indirectly influencing streamflow as it is responsible for the evaporation and moisture, while precipitation is the major driving factor changes in streamflow especially in tropical river basins like Wabi Shebele River Basin. Table 10 shows that the correlation matrix in between peak flood discharge and driving forces. PCA conducted in this study reveals that rainfall factor is less related than other driving forces (Table 11). To see relationships of rainfall and floods in the basin,

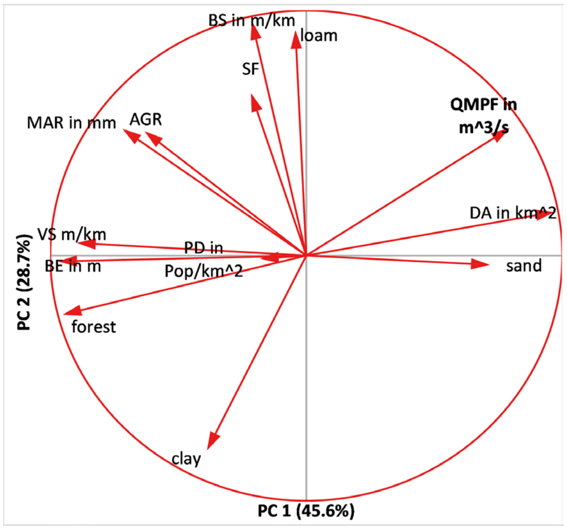


Figure 8. Two-dimensional correlation plot of the coefficients of the first two principal components (PC1 & PC2).

Table 10. Correlation matrix in between variables.

Pearson's r	QMPF	DA	BE	BS	BP	VL	SF	DD	VS	ER	clay	sand	loam	MAR	AGR	forest	PD
QMPF	-																
DA	**0.915**	-															
BE	**−0.754**	−0.934	-														
BS	0.292	−0.055	0.166	-													
BP	**0.928**	0.998	−0.919	−0.037	-												
VL	**0.959**	0.976	−0.893	0.079	0.982	-											
SF	0.282	−0.022	0.103	0.610	0.016	0.197	-										
DD	**−0.817**	−0.978	0.964	0.210	−0.972	−0.923	0.165	-									
VS	**−0.796**	−0.913	0.784	0.268	−0.905	−0.837	0.268	0.899	-								
ER	−0.282	0.022	−0.103	−0.610	−0.016	−0.197	−1.000	−0.165	−0.268	-							
clay	**−0.587**	−0.443	0.405	−0.721	−0.457	−0.484	−0.225	0.383	0.176	0.225	-						
sand	**0.538**	0.630	−0.785	0.275	0.610	0.621	0.017	−0.649	−0.333	−0.017	−0.739	-					
loam	0.450	0.126	0.086	0.860	0.163	0.215	0.417	0.005	0.006	−0.417	−0.792	0.188	-				
MAR	−0.166	−0.532	0.727	0.569	−0.491	−0.370	0.695	0.674	0.537	−0.695	0.050	−0.576	0.498	-			
AGR	−0.352	−0.580	0.538	0.734	−0.593	−0.505	0.294	0.666	0.689	−0.294	−0.273	0.033	0.395	0.486	-		
forest	**−0.877**	−0.963	0.929	−0.060	−0.948	−0.921	0.107	0.939	0.861	−0.107	0.559	−0.754	−0.180	0.597	0.390	-	
PD	−0.395	−0.307	0.098	0.212	−0.357	−0.389	−0.402	0.258	0.362	0.402	−0.226	0.383	−0.054	−0.315	0.657	0.049	-

QMPF = Mean peak flow (m^3/s), DA = Drainage area (Km^2), BE = Basin elevation (m), BS = Basin slope (m/km), BP = Basin perimeter (km), VL = Valley length (km), SF = Shape factor, DD = Drainage density, VS = Valley slope (m/km), clay = fraction of clay, sand = fraction of sand, loam = fraction of loam, MAR = Mean annual rainfall (mm), AGR = fraction of agricultural land, forest = fraction of forest coverage, PD = Population density (pop/km^2).

Table 11. Pearson correlation (r) computed between precipitation extremes and flood discharges for the studied period, 1980-2010.

River/watersheds	Correlation (Pearson's r)			
	Annual	Spring	Summer	Winter
Upper catchments				
Wabi @ Dodola Bridge	−0.012	0.155	−0.056	0.123
Maribo	−0.014	0.254	−0.003	0.106
Robe	0.005	0.257	0.024	0.132
Middle Catchments				
Wabi @ Legehida	0.003	0.09	0.057	0.432
Erer @ Hamaro	0.038	0.136	0.1	0.284
Gololcha @ Wabi junction	0.186	0.141	0.172	0.369
Eastern Catchments				
Erer @ Babile	−0.042	0.08	0.187	−0.136
Fafen @ Jijiga	0.2	0.166	0.083	0.201
Fafen @ Kebridehar	0.062	0.133	0.235	0.202
Lower Catchments				
Wabi @ Gode	−0.018	−0.034	0.016	−0.214
Wabi @ Burkur	−0.016	−0.034	0.008	−0.216

Pearson's correlation test is performed in between ground based gauged precipitation data and flood discharges at different catchments of Wabi Shebele River Basin. Maximum flow discharge at annual and seasonal aggregation levels are used to correlate with total rainfall at annual and seasonal levels. The result indicates that flood discharge in the basin is positively correlation values in most watersheds both at annual and seasonal aggregation levels.

4.5.3. Environmental Background Conditions

The environmental background condition factors, *i.e.*, drainage area (DA), elevation, slope, and soil type are analyzed in the study. The elevation varies greatly in Wabi Shebele River Basin, with moderate to low elevations in eastern and lower catchments (*i.e.*, Fafen watershed and lowlands of Wabi Shebele Basin), moderate and high elevations in the upper and middle catchments (*i.e.*, middle, northern and north west of the basin). Elevational differences of watersheds reflect differences in land cover type and associated hydrologic runoff response and geomorphic disturbance [44]. Less land cover results rapid response to rainfall, increased flood magnitudes, increased potential for debris flows, abundant erosion and consequently increase slope failures. Among the environmental background condition factors elevation and drainage size are among the most important factors in formation of high floods in terms of magnitude and frequency in Wabi Shebele River Basin (**Table 10**).

Watershed slope (WS; %) is the mean watershed slope, measured by calculating the maximum rate of change between each cell. It provides an indication of the steepness of the drainage area. As the slope decreases, catchment soils become more permeable and thus the effect of infiltration becomes more significant [45]. In Wabi Shebele Basin the east and downstream part of the basin is characterized by low slopes, while the west and upstream part of the basin is characterized by high slopes.

In surface runoff generation, soil infiltration rate is another sensitive variable. Course textured soils have large well-connected spaces and allow more water to infiltrate through it quite rapidly while fine grained soils dominated by clay have low infiltration rates due to their smaller sized pore spaces [46]. Soil containing large amount of the sand and silt tend to form crust and become compacted, which significantly reduces the infiltration rate. The amount of organic matter on soil surface can enhance infiltration because organic matter has more porous than mineral soil particles and it can hold much greater quantity of water. In Wabi Shebele Basin soil distributions vary spatially; loamy sand in the east and downstream part of the basin, clay in middle part of the basin, sandy loam in the center of the province, and silty clay in north west (**Figure 9**). Sand soil is identified as the most powerful variable among soil types in flood formation of Wabi Shebele Basin (**Figure 8** and **Table 10**).

4.5.4. External Factors

The results presented under Sections 4.5.2 and 4.5.3 indicates the relationships

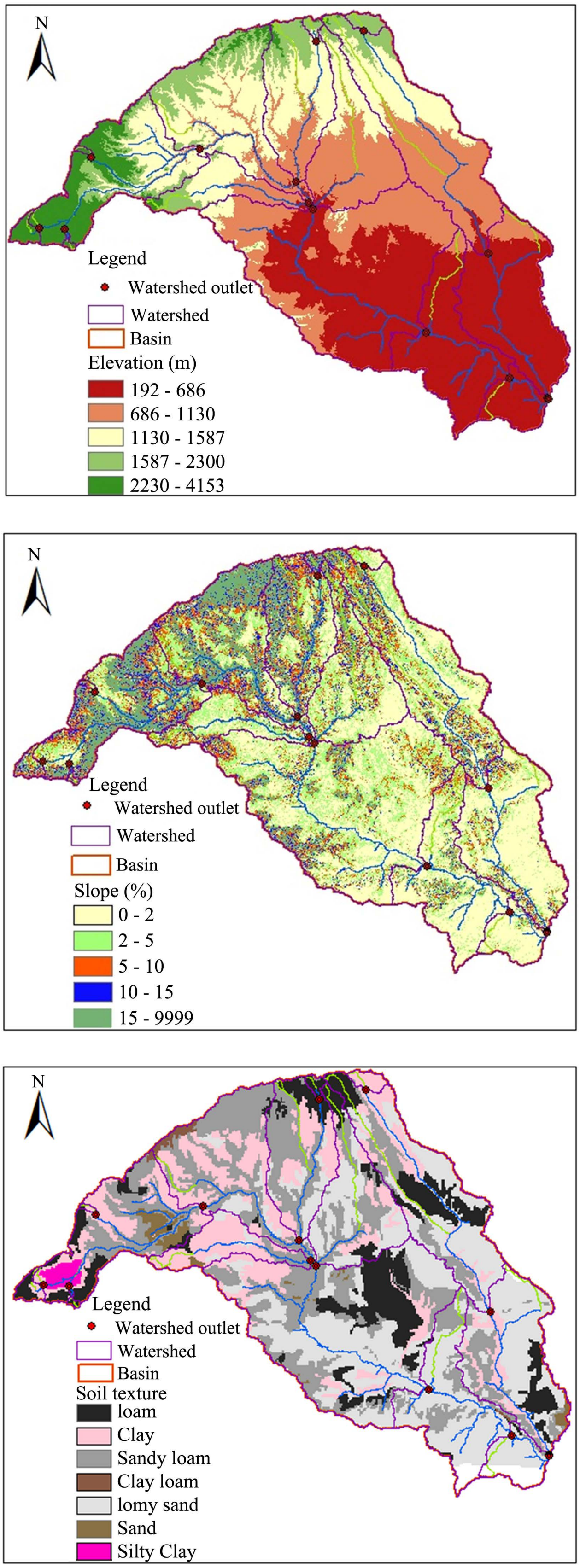


Figure 9. Elevation, slope and soil distribution of study area.

between flood trends and changes in some atmospheric variables (*i.e.*, meteorological) and environmental background conditions, but these attributes explain variability of flood discharge only at partial level. Therefore, non-climatic changes in catchment and river parameters must also be taken into account.

During the last four decades some environmental changes occurred in the studied catchments that influence the conditions of flood runoff. Land use and land cover has been changing in the northern part of Wabi Shebele River Basin (**Table 12**). In the basin the extent of shrublands indicates significant increasing trends while grassland and cultivated area showed decreasing trend from 1984 to 2004 [47]. Similarly, the extent of riparian woodland in the basin indicates decrement in this period interval. Shrubland class is the areas with extensive physical limitation; like very steep slopes, shallow soils, rock outcrops, series of deeply dissected gorges, dry and rugged areas. Due to human activities and pressures in large semi-arid areas (middle and eastern upper catchments) of the basin, most coverage of Riparian Woodland, Grassland and Perennial and seasonal Swamp and Marshland covers are changed to Shrubland in the basin.

Another potential driver of floods in watersheds is population density. The population growth and cultivated land density has strong correlation [48]. Increment of cultivated land results in accelerated runoff process, especially as a consequence of rapid development of gullies [17]. In Wabi Shebele Basin, human activities are concentrated in the west and eastern upper highland areas of the Basin. The cultivation land density has strong correlation value with population density with correlation value (r) of 0.657, while negatively correlated to flood discharge in catchments with correlation value of −0.395 (**Table 10**).

Table 12. Land use/land cover changes in Wabi Shebele Basin between 1984 and 2004.

S. No.	Land Use/Cover Type	1984	2004	Change (%)	
		Area (km^2)	Area (km^2)	Area (km^2)	%
1	Cultivated lands	26,989	23,507	−3482	−13%
2	Afro-Alpine and Sub-Afro-Alpine vegetation	167	397	230	138%
3	Forest lands	1691	1691	0	0%
4	Woodlands	4301	7409	3108	72%
5	Riparian wood lands	1080	241	−839	−78%
6	Shrub lands	51,406	138,396	86,990	169%
7	Grass lands	87,383	24,791	−62,592	−72%
8	Wet lands	1052	68	−984	−94%
9	Bare lands	27,954	6882	−21,072	−75%
10	Water bodies	39	32	−7	−18%
Total		202,220	202,220		

Source: [10].

4.6. Discussion

The peak flood discharge shows good agreement with some variables from climate, environmental background conditions and external factors in Wabi Shebele Basin. Climate factors, specifically precipitation indicates positive correlation with flood events in the basin. The correlation analysis indicates that annual maximum flow discharge has positively correlated to total rainfall in 55% of sample watersheds. Seasonal flood discharges also positively correlated to seasonal precipitations in most of sample watersheds (>78%). Therefore, our study identified the precipitation as one of driving forces of flood events in Wabi Shebele Basin, which confirms the findings of Ethiopian Ministry of Water, Irrigation and Energy (MoWIE) [10] during basin master plan study. This study reveals that severe hydrologic extremes in the basin especially in 1973, 1979, 1984-1985 is caused by natural atmospheric variability.

Watershed characteristics like drainage size, elevation, slope and soil types are identified as the most powerful variables in flood events of Wabi Shebele Basin. The mean peak flow (QMPF) is positively correlated with variables: drainage area (DA), basin perimeter (BP), valley length (VL), shape factors (SF), fraction of sand and loamy, where maximum correlation with DA, BP, VL and sand with correlation coefficient of 0.92, 0.93, 0.96 and 0.54 respectively. This means that larger watersheds are expected to have a higher mean peak flow. There are different studies confirms this result that the drainage area is a significant factor that positively affect peak discharge e.g., [45] [49]. Drainage size affects not only the flow collecting ability but also the time to peak discharge [45].

External factors like land use change and population density are also exhibit differences in hydrologic runoff response, which can be directly linked to flood events. Fraction of forest coverage and population density is found negatively correlated to flood discharges in Wabi Shebele Basin. Loss of land cover, thinner forest canopies, grass lands and reduced infiltration of rainfall result in rapid hydrologic response, increased flood magnitudes and frequency [49]. In Wabi Shebele River Basin, the magnitudes and frequency of flood events particularly in middle watersheds indicates increasing trend in recent decades (Sections 4.2 and 4.3). In mountainous zones, increased potential for intense convective storms, increase cultivated land and more highly confined river valleys results more rapid runoff response to precipitation and cause for extreme floods. In the basin, human activities are concentrated in the west and eastern upper highland areas of the Basin [10].

5. Conclusions

In this study variabilities in flood and relationships with major driving factors are investigated. Both exploratory data analysis (EDA) and non-parametric tests (*i.e.*, Mann-Kendal trend test and quantile perturbation (QPM) methods) are used to see temporal variabilities in flood discharge. The risk level used was 5 %. The *p*-value of the MK statistic S of sample data is used to measure the significance of trend value; if $p \leq 0.050$ (significance level), then the existing trend is assessed to be statistically significant. Finally, Principal Correlation Analysis (PCA) is performed in between flood discharge and potential driving factors to identify the most powerful drivers on flood changes. A larger absolute value of the Eigen coefficient indicates variables that contribute more to the component.

The multi-temporal trend analysis at 5, 10, 15, 15, 25 and 30-year intervals starting from 1980 in annual maximum discharge showed increasing trends for most recent periods (since 1996) in all sample stations while decreasing tendency of flood discharges is observed before the 2000s particularly in eastern and lower catchments in the basin. Peak over threshold (3rd quartile) values of discharge analysis using QPM show that there are significant positive anomalies (outside confidence interval) in the early 1990s & 2000s and negative anomalies in the 1980s, particularly in upper and middle catchments. However, in eastern catchments, the 1980s is the decade in which significant positive anomalies were observed and negative anomalies in the 1990s and 2000s.

The PCA reveals that drainage area (DA), watershed mean elevation (BE), valley slope (VS), fraction of sand coverage (sand), population density (PD), forest fraction (forest) and mean annual rainfall (MAR) were initially found to be the best predictors of flood peak discharge. Among these, drainage area (DA), mean annual rainfall (MAR) and the fraction of forest coverage (forest) were discovered as the principal driving factors for flood peak discharge in Wabi Shebele River Basin. However, fraction of clay soil, fraction of loam soil, basin shape factor (SF), basin slope (BS) and fraction of agricultural land (AGR) were found to be less important in flood peak discharge prediction of Wabi Shebele Basin watersheds.

Generally, the study tried to answer, the trends and variability status of extreme flood discharge and its relations ships with potential driving factors in the Wabi Shebele Basin. The study can provide information on which driving factors should be prioritized in mitigation measures to decrease extreme weather disasters in the basin. Accordingly, basin slope and drainage size from environmental background conditions and forest coverage and population density from external factors are the major driving factors one can improve to minimize the hydrological disasters in the area.

Acknowledgements

We would like to thank the Ethiopian National Meteorology Service Agency and Ministry of Water, Irrigation, and Electricity for providing the necessary data.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Kay, A.L., Crooks, S.M., Pall, P. and Stone, D.A. (2011) Attribution of Autumn/Winter

2000 Flood Risk in England to Anthropogenic Climate Change: A Catchment-Based Study. *Journal of Hydrology*, **406**, 97-112. https://doi.org/10.1016/j.jhydrol.2011.06.006

[2] Kundzewicz, Z.W. and Stoffel, M. (2016) Anatomy of Flood Risk. In: Kundzewicz, Z., Stoffel, M., Niedźwiedź, T. and Wyżga, B., Eds., *Flood Risk in the Upper Vistula Basin*, Springer, Cham, 39-52. https://doi.org/10.1007/978-3-319-41923-7_3

[3] Bissolli, P., Friedrich, K., Rapp, J. and Ziese, M. (2011) Flooding in Eastern Central Europe in May 2010—Reasons, Evolution and Climatological Assessment. *Weather*, **66**, 147-153. https://doi.org/10.1002/wea.759

[4] Li, Z., Liu, D., Li, X.-Y., Wu, H., Li, G.-Y. and Li, Y.-T. (2016) Runoff Coefficient Characteristics and Its Dominant Influencing Factors in a Riparian Grassland in the Qinghai Lake Watershed, NE Qinghai-Tibet Plateau. *Arabian Journal of Geosciences*, **9**, Article No. 397. https://doi.org/10.1007/s12517-016-2404-z

[5] Williams, A.P. and Funk, C. (2011) A Westward Extension of the Warm Pool Leads to a Westward Extension of the Walker Circulation, Drying Eastern Africa. *Climate Dynamics*, **37**, 2417-2435. https://doi.org/10.1007/s00382-010-0984-y

[6] Ethiopian Panel on Climate Change, EPCC Ethiopian Panel on Climate Change (2015) First Assessment Report, Summary of Reports for Policy Makers. Ethiopian Academy of Sciences, Addis Ababa.

[7] Hall, J., Arheimer, B., Borga, M., *et al.* (2014) Understanding Flood Regime Changes in Europe: A State-of-the-Art Assessment. *Hydrology and Earth System Sciences*, **18**, 2735-2772. https://doi.org/10.5194/hess-18-2735-2014

[8] Akola, J., Binala, J. and Ochwo, J. (2018) Guiding Developments in Flood-Prone Areas: Challenges and Opportunities in Dire Dawa City, Ethiopia. *Jàmbá*: *Journal of Disaster Risk Studies*, **11**, a704. https://doi.org/10.4102/jamba.v11i3.704

[9] Mamo, S., Berhanu, B. and Melesse, A.M. (2019) Historical Flood Events and Hydrological Extremes in Ethiopia. In: *Extreme Hydrology and Climate Variability*: *Monitoring*, *Modelling*, *Adaptation and Mitigation*, Elsevier, 379-384. https://doi.org/10.1016/B978-0-12-815998-9.00029-4

[10] Ministry of Water Resources (MoWR) (2003) Wabi Shebele River Basin Integrated Master Plan Study Project. Vol. VII Water Resources, Part 2 Hydrology, Part. Ethiopia. Part; Ethiopia.

[11] Tadesse, T., Haigh, T., Wall, N., *et al.* (2016) Linking Seasonal Predictions to Decision-Making and Disaster Management in the Greater Horn of Africa. *Bulletin of the American Meteorological Society*, **97**, ES89-ES92. https://doi.org/10.1175/BAMS-D-15-00269.1

[12] UNDP (2005) Drought and Floods Stress Livelihoods and Food Security in the Ethiopian Somali Region—Ethiopia. https://reliefweb.int/report/ethiopia/drought-and-floods-stress-livelihoods-and-food-security-ethiopian-somali-region

[13] Adnan, N.A. (2010) Quantifying the Impacts of Climate and Land Use Changes on the Hydrological Response of a Monsoonal Catchment. University of Southampton, Southampton.

[14] Taye, M.T. and Willems, P. (2012) Temporal Variability of Hydroclimatic Extremes in the Blue Nile Basin. *Water Resources Research*, **48**, W03513. https://doi.org/10.1029/2011WR011466

[15] Allamano, P., Claps, P. and Laio, F. (2009) An Analytical Model of the Effects of Catchment Elevation on the Flood Frequency Distribution. *Water Resources Research*, **45**, W01402. https://doi.org/10.1029/2007WR006658

[16] Bates, B.C., Kundzewicz, Z., Palutikof, J. and Shaohong, W. (2008) World Meteorological Organisation (WMO); Intergovernmental Panel on Climate Change Climate Change and Water [Electronic Resource]: IPCC Technical Paper VI, IPCC Secretariat, Geneva.

[17] Kundzewicz, Z.W., Stoffel, M., Niedźwiedź, T. and Wyżga, B. (2018) Flood Risk in the Upper Vistula Basin. Springer International Publishing AG, Switzerland.

[18] Ruiz-Villanueva, V., Wyżga, B., Kundzewicz, Z.W., *et al.* (2016) Variability of Flood Frequency and Magnitude during the Late 20th and Early 21st Centuries in the Northern Foreland of the Tatra Mountains. In: *Flood Risk in the Upper Vistula Basin*, Springer International Publishing AG, Switzerland, 231-256.

[19] IPCC (2010) Expert Meeting on Detection and Attribution Related to Anthropogenic Climate Change: The World Meteorological Organization, Geneva, Switzerland, 14-16 September 2009: Meeting Report; Stocker, T., Intergovernmental Panel on Climate Change, Working Group I, The Physical Science Basis, Eds.; IPCC Working Group I Technical Support Unit: Bern, Switzerland.

[20] Awass, A.A. (2009) Hydrological Drought Analysis—Occurrence, Severity, Risks: The Case of Wabi Shebele River Basin, Ethiopia. Seigen University, Seigen.

[21] BCEOM_ORSTOM_EDF (1973) Hydrological Survey of the Wabi Shebelle Basin. http://www.hydrosciences.fr/sierem/Bibliotheque/biblio/Wabi/Wabi-20-20hydrological-20survey-20-20Part-20I.pdf

[22] International Water Management Institute, IWMI (2015) Drivers of Hydrological Dynamics in the Bale Eco-Region. https://phe-ethiopia.org/pdf/7.%20Drivers-20-of-hydrological-20dynamics-%2020in-20the-20BER.pdf

[23] Food and Agricultural Organization of United Nation, FAO (1961) Soil Maps and Soil Databases. http://www.fao.org/soils-portal/soil-survey/soil-maps-and-database/FAO-UNESCO-soil-map-of-the-world/en

[24] Neitsch, S.L., Grassland, S. and W.R.L. (2005) Blackland Research Center Soil and Water Assessment Tool: Theoretical Documentation: Version 2005; Grassland, Soil and Water Research Laboratory; Blackland Research Center: Temple (Texas).

[25] Ven Te Chow, C., Maidment, D.R. and Mays, L.W. (1963) Applied Hydrology; Technical Correspondence School, N.Z. Department of Education, Wellington.

[26] 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**, 885-900. https://doi.org/10.13031/2013.23153

[27] Schuol, J. and Abbaspour, K.C. (2007) Using Monthly Weather Statistics to Generate Daily Data in a SWAT Model Application to West Africa. *Ecological Modelling*, **201**, 301-311. https://doi.org/10.1016/j.ecolmodel.2006.09.028

[28] Monteith, J.L. (1965) Evaporation and Environment. *Symposia of the Society for Experimental Biology*, **19**, 205-234.

[29] Priestley, C.H.B. and Taylor, R.J. (1972) On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Monthly Weather Review*, **100**, 81-82. https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

[30] Hargreaves, G.H. and Samani, Z.A. (1985) Reference Crop Evapotranspiration from Temperature. *Applied Engineering in Agriculture*, **1**, 96-99. https://doi.org/10.13031/2013.26773

[31] Keast, D. and Ellison, J. (2013) Magnitude Frequency Analysis of Small Floods Us-

ing the Annual and Partial Series. *Water*, **5**, 1816-1829. https://doi.org/10.3390/w5041816

[32] Malamud, B.D. and Turcotte, D.L. (2006) The Applicability of Power-Law Frequency Statistics to Floods. *Journal of Hydrology*, **322**, 168-180. https://doi.org/10.1016/j.jhydrol.2005.02.032

[33] Kendall, M.G. (1975) Rank Correlation Methods. Charles Griffin, London.

[34] Mann, H.B. (1945) Nonparametric Tests against Trend. *Econometrica*, **13**, 245-259. https://doi.org/10.2307/1907187

[35] Ntegeka, V. and Willems, P. (2008) Trends and Multidecadal Oscillations in Rainfall Extremes, Based on a More than 100-Year Time Series of 10 Min Rainfall Intensities at Uccle, Belgium. *WRCR Water Resources Research*, **44**, W07402. https://doi.org/10.1029/2007WR006471

[36] Onyutha, C. (2016) Statistical Analyses of Potential Evapotranspiration Changes over the Period 1930-2012 in the Nile River Riparian Countries. *Agricultural and Forest Meteorology*, **226-227**, 80-95. https://doi.org/10.1016/j.agrformet.2016.05.015

[37] Tabari, H., AghaKouchak, A. and Willems, P.A. (2014) Perturbation Approach for Assessing Trends in Precipitation Extremes across Iran. *Journal of Hydrology*, **519**, 1420-1427. https://doi.org/10.1016/j.jhydrol.2014.09.019

[38] Merz, B., Vorogushyn, S., Uhlemann, S., Delgado, J. and Hundecha, Y. (2012) HESS Opinions "More Efforts and Scientific Rigour Are Needed to Attribute Trends in Flood Time Series". *Hydrology and Earth System Sciences*, **16**, 1379-1387. https://doi.org/10.5194/hess-16-1379-2012

[39] Rahman, A.S. and Rahman, A. (2020) Application of Principal Component Analysis and Cluster Analysis in Regional Flood Frequency Analysis: A Case Study in New South Wales, Australia. *Water* (*Switzerland*), **12**, 781. https://doi.org/10.3390/w12030781

[40] Çamdevýren, H., Demýr, N., Kanik, A. and Keskýn, S. (2005) Use of Principal Component Scores in Multiple Linear Regression Models for Prediction of *Chlorophyll*-a in Reservoirs. *Ecological Modelling*, **181**, 581-589. https://doi.org/10.1016/j.ecolmodel.2004.06.043

[41] Kerr, R.A. (1992) Unmasking a Shifty Climate System. *Science*, **255**, 1508-1510. https://doi.org/10.1126/science.255.5051.1508

[42] Nash, J.E. and Sutcliffe, J.V. (1970) River Flow Forecasting through Conceptual Models, Part I—A Discussion of Principles. *Journal of Hydrology*, **10**, 282-290. https://doi.org/10.1016/0022-1694(70)90255-6

[43] NMA (1996) Climatic and Agroclimatic Resources of Ethiopia. NMSA Meteorological Research Report Series, Ethiopia.

[44] Sutfin, N.A. and Wohl, E. (2019) Elevational Differences in Hydrogeomorphic Disturbance Regime Influence Sediment Residence Times within Mountain River Corridors. *Nature Communications*, **10**, Article No. 2221. https://doi.org/10.1038/s41467-019-09864-w

[45] Al-Rawas, G.A. and Valeo, C. (2010) Relationship between Wadi Drainage Characteristics and Peak-Flood Flows in Arid Northern Oman. *Hydrological Sciences Journal*, **55**, 377-393. https://doi.org/10.1080/02626661003718318

[46] Prachansri, S. (2007) Analysis of Soil and Land Cover Parameters for Flood Hazard Assessment; 106. https://webapps.itc.utwente.nl/librarywww/papers_2007/msc/aes/prachansri.pdf

[47] Ministry of Water Resources (MoWR) (2004) Wabi Shebele River Basin Integrated Development Master Plan, Land Use/Land Cover Study, Part. Ethiopia. Part; Ethiopia.

[48] Liu, Y., Yuan, X., Guo, L., Huang, Y. and Zhang, X. (2017) Driving Force Analysis of the Temporal and Spatial Distribution of Flash Floods in Sichuan Province. *Sustainability*, **9**, 1527. https://doi.org/10.3390/su9091527

[49] Huang, P.-C. (2020) Analysis of Hydrograph Shape Affected by Flow-Direction Assumptions in Rainfall-Runoff Models. *Water* (*Switzerland*), **12**, 452. https://doi.org/10.3390/w12020452