Natural Resources, 2022, 13, 38-51
https://www.scirp.org/journal/nr
ISSN Online: 2158-7086
ISSN Print: 2158-706X

Flood Generation Mechanisms and Potential Drivers of Flood in 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: *belete.berhanu@aait.edu.et

How to cite this paper: Wudineh, F.A., Moges, S.A. and Kidanewold, B.B. (2022) Flood Generation Mechanisms and Potential Drivers of Flood in Wabi-Shebele River Basin, Ethiopia. *Natural Resources*, **13**, 38-51.
https://doi.org/10.4236/nr.2022.131003

Received: November 9, 2021
Accepted: January 26, 2022
Published: January 29, 2022

Abstract

Flood is a natural process generated by the interaction of various driving factors. Flood peak flows, flood frequency at different return periods, and potential driving forces are analyzed in this study. The peak flow of six gauging stations, with a catchment area ranging from 169 - 124,108 km^2 and sufficient observed streamflow data, was selected to develop threshold (3rd quartile) magnitude and frequency (POTF) that occurred over ten years of records. Sixteen Potential climatic, watershed and human driving factors of floods in the study area were identified and analyzed with GIS, Pearson's correlation, and Principal Correlation Analysis (PCA) to select the most influential factors. Eight of them (MAR, DA, BE, VS, sand, forest AGR, PD) are identified as the most significant variables in the flood formation of the basin. Moreover, mean annual rainfall (MAR), drainage area (DA), and lack of forest cover are explored as the principal driving factors for flood peak discharge in Wabi-Shebele River Basin. Finally, the study resulted in regression equations that helped plan and design different infrastructure works in the basin as ungauged catchment empirical equations to compute Q_{MPF}, Q_5, Q_{10}, Q_{50}, and Q_{100} using influential climate, watershed, and human driving factors. The results of these empirical equations are also statistically accepted with a high significance correlation ($R^2 > 0.9$).

Keywords

Flood Drivers, Climate Factors, Watershed Characteristics, Human Drivers, Principal Correlation Analysis (PCA), Multiple Regression Model

1. Introduction

Flood is a natural hazard that is most widespread around the globe both in terms

*Corresponding author.

of the occurrence and the resulting damages to human lives, environments, and properties [1]. Based on a combination of sources, causes, and impacts, floods categorize into river (or fluvial) floods, pluvial (or overland) floods, coastal floods, groundwater floods, or the failure of artificial water systems [2]. Therefore, the major causes of floods include intensity, duration, and spatial distribution of rainfall on catchments; steep slopes, deforestation, less soil infiltration capacity; failure of hydrologic structures, and sudden release of water from dams; and landslides [3]. Nied *et al.* [4] also describe physical controlling factors of flood include: hydrological pre-conditions (e.g., soil saturation, snow cover), meteorological conditions (e.g., amount, intensity, and spatial and temporal distribution of precipitation), runoff generation processes as well as river routing (e.g., superposition of flood waves in the main river and its tributaries). These multidimensional causes of the flood made it less predictable and aggravated its impacts worldwide [5].

Floods are mainly driven by climate, catchment, and river characteristics that determine the terrestrial conditions of water or runoff [6]. Climate is a critical driver on the fluvial flood hazard. And it is also highly affected by various features of atmospheric systems, including water content of the atmosphere, different precipitation characteristics (intensity, duration, total amount, timing, or phase), the antecedent precipitation index (API), large-scale circulation patterns [7]. It is true in Ethiopia also; the climate/weather characteristics, including torrential rainfall and summer thunderstorms, are strongly linked with flooding [8] [9]. Similarly, the catchment characteristics, variability in drainage area, very short changeable topography, and low infiltration capacity of the ground surface expose to high floods [3]. Although flood is a natural action, the human land base activities often involve clearing the natural vegetation (either for construction or agriculture), and altering the characteristics of the ground cover can increase runoff substantially, and the potential threat from flash floods and river floods [10].

However, the impact levels of flood drivers, the significance among the different elements of flood factors, and the relationship between peak discharges and potential drivers are still a critical knowledge gap in tropical river basins. Moreover, understanding the hydrological process of flooding in different regions and estimating the flood quintals are important limitations in the basin since most rivers are ungauged. Therefore, this study aimed to address the above knowledge gaps and development hindrance by identifying influential flood generations drivers and establishing relationships among drivers and peak flood indicators.

2. Flooding in Wabi-Shebele Basin

Floods that cause most damages in Wabi-Shebele River Basin are generated by a few days of heavy rainfalls with an average intensity of 10 - 200 mm/hr and a total sum of precipitation of a hundred millimeters [11]. In the basin, flood events have occurred regularly as flash floods in the lowland sections, as seen from river beds' state and sheet erosion evidence [12]. MoWR [13] summarizes floods formation mechanisms in Wabi-Shebele River Basin under three general categories: 1) High floods derived from a generalized runoff, occur after an average rainy phase of 10 days with a total rainfall depth exceeding 80 mm (e.g., northwest of the basin at upstream of Melka Wakena Hydropower); 2) Floods caused by intense rainfall and on impervious soils (e.g., middle basin between Melka Wakena and Hamaro Hedad); 3) Floods in alluvial plains (e.g., lower basin between Hamero Hedad and Somalia border); short and violent floods (e.g., floods in Fafen watershed). Historical flood events in the basin are summarized in Table 1.

3. Flood Discharge Characteristics in Wabi-Shebele River Basin

The peak flows over threshold (3rd quartile) magnitude and frequency (POTF) are analyzed. The analysis undertakes with a fixed time interval approach to ensure the time-series independencies of extreme values. The successive peaks within the Time intervals between 5 to 14 days are used in this study [14] [15]. A total of 89 events consider in this POTF analysis. The mean peak flow (Q_{MPF}) is expressed as the arithmetic mean value of peak over the threshold (3rd quartile) flows for the period of record.

Table 1. Historical flood events in Wabi Shebele River Basin, 1980-2019.

Date	Disaster event	Causality Reported	Source
July 1993	Flooding: 120,000 people affected	Heavy Rain	http://floodobservatory.colorado.edu/Archives/index.html
1995	Flooding: 89,902 people affected, 27 deaths	Heavy Rain	http://floodobservatory.colorado.edu/Archives/index.html
Oct. 1997 Feb. 1998	El Niño related flooding	Torrential rain	http://www.fao.org/docrep/004/w7832e/w7832e00.HTM
1999	Flooding: 85,789 people affected, 34 deaths.	Torrential rain	http://floodobservatory.colorado.edu/Archives/index.html
2003	Floods: 119 people died	Heavy Rain	https://go-api.ifrc.org/publicfile/download?path=/docs/appeals/05/&name=05me03001.pdf
2005	Flooding: 103,000 people affected, 177 deaths	Heavy Rain	https://go-api.ifrc.org/publicfile/download?path=/docs/appeals/05/&name=05me03001.pdf
2006	Flooding: 410,132 people affected, 132 deaths	Heavy Rain	http://floodobservatory.colorado.edu/Archives/index.html
May 2008	Flooding: 11 deaths, 52,000 people abandoned, 164 hectares farmland washed away.	Heavy seasonal rains	http://www.irinnews.org/report/81526/ethiopia
2010	Flooding: 16,000 people affected	Heavy Rain	http://floodobservatory.colorado.edu/Archives/index.html
2015	Flooding: 105,000 people affected	Heavy Rain	http://floodobservatory.colorado.edu/Archives/index.html

The sampled watersheds exhibit less variability in flood-peak discharges. From **Table 2**, the standard deviation in Q_{MPF} is less than 35% of the mean except at the Jijiga station (*i.e.*, a standard deviation related to 37% of the mean value). Studies [16] [17] indicated the higher standard deviation of flood discharges, indicating a potential for flash floods. Accordingly, only the northeastern part of the basin, in the Jijiga watershed, is identified as a potential flash flood area. Floods in other catchments are fall under riverine floods.

The Mann-Kendall test [18] [19], the common non-parametric trend detection, is used to detect trends in flood discharge. Most of the flood discharges indicate a significant trend ($p < 0.05$) in gauging stations located in the northwestern (*i.e.*, Wabi at Dodola and Maribo) and downstream part of the basin (*i.e.*, Gode). However, flood discharge at Robe and Erer discharge has no significant trend, as shown in **Table 2**. Both rivers locate in a northern highland part of the basin. Flood discharges at Robe River indicate a less increasing trend while flood discharge of Erer river at Babile shows decreasing tendency. Some studies also found similar upward trends of flood events with the consequence of increases in deaths, injuries, stress-related disorders in the study area in the past six decades [9] [20] [21].

4. Potential Flood Drivers in Wabi-Shebele River Basin

4.1. Climate Factors

Precipitation with its different characteristics; intensity, duration, total amount, timing, or phase (whether liquid or solid), are an essential climate variable in shaping flood hazard [3] [6] [20] [22] [23] [24]. However, due to the basin's data scares, the relationship between climate factors and Q_{MPF} was established based on annual and monthly rainfalls. The correlation analysis is performed between rainfall in the rainy seasons, *i.e.*, March to September (6 months), and flood discharge to see the impact of climate on flood events of the study area. The maximum water discharge in Wabi-Shebele River Basin moderately correlates with the total annual rainfall over the watersheds ($R^2 = 0.314$ on average; **Figure 1**).

Table 2. Statistics of log flood-peak (Q_{MPF}) records.

River/stations	Period of records	Number of events	Log Q		Mann Kendall trend test at 0.05 level
			Mean	St. dev.	
Maribo	1975-2008	18	0.87	0.17	2.45
Wabi at Dodola	1975-2015	13	1.21	0.15	2.42
Robe	1979-2006	18	0.48	0.34	0.13
Jijiga	1985-1996	14	0.40	0.37	3.71
Erer at Babile	1984-1999	16	0.77	0.20	−1.11
Wabi at Gode	1967-2002	10	2.44	0.20	2.76
	Total	89			

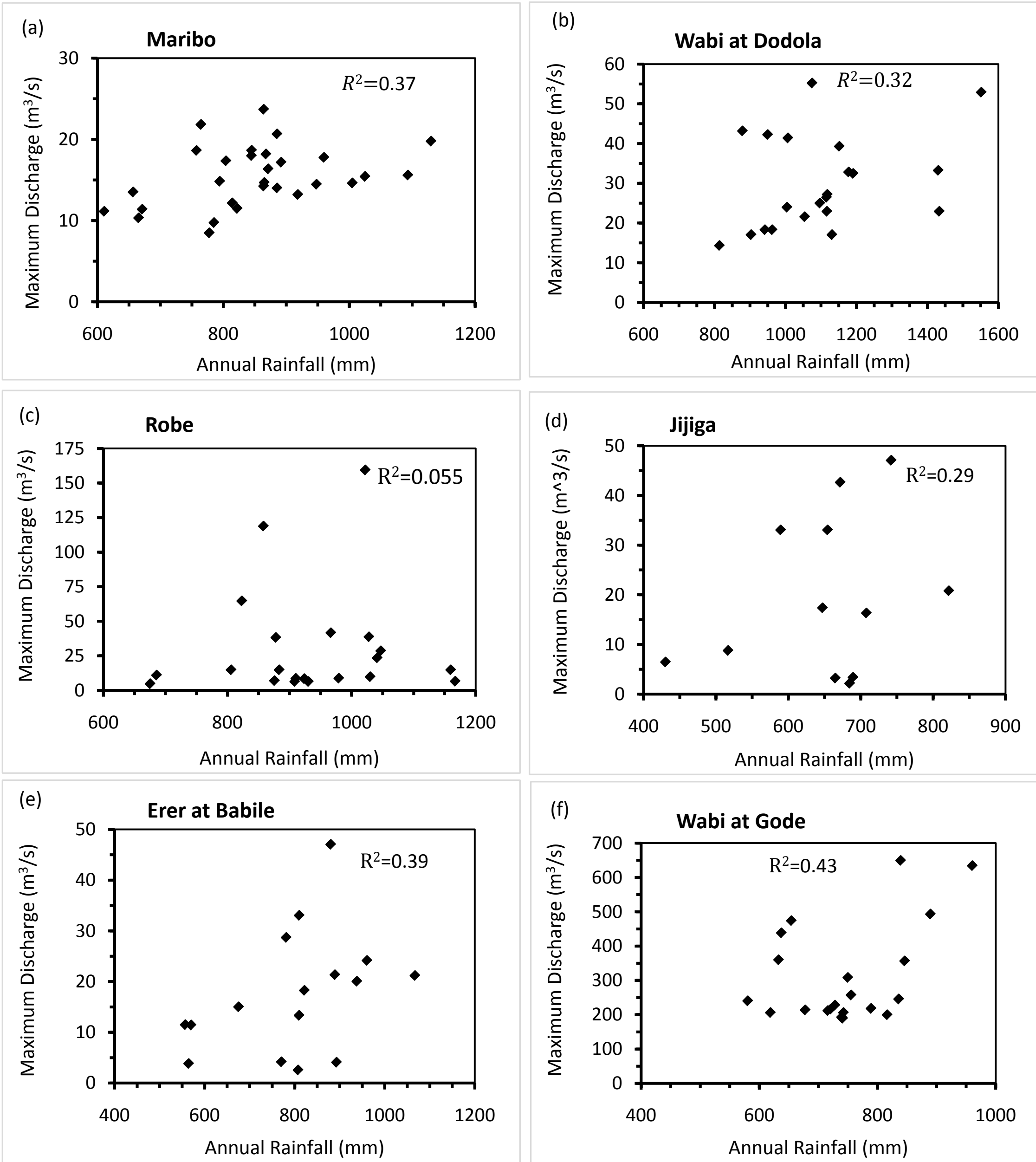

Figure 1. Relationship between annual rainfall (mm) and a yearly maximum of discharges (m^3/sec).

Although the time lag between maximum annual precipitation and the maximum annual flood was expected more than a day for such large size and complex land formation basins, the maximum number of consecutive wet days is positively correlated with peak daily discharges. Notably, on large sub-watersheds like Wabi at Dodola Bridge and Gode stations, its relation becomes significant with a correlation value of 0.49 and 0.7. It is noticed that climate factors than other drivers significantly influence the flooding in the Wabi-Shebele basin. Moreover, the rain-bearing clouds coverage over the Wabi-Shebele River Basin is less intense than other basins like the Abay basin in Ethiopia [13]. Therefore, flood events in the Wabi-Shebele basin are highly associated with the frequency of precipitation events.

4.2. Watershed Factors

The multiple catchment characteristics identified as vital variables affecting flows [17] [25] [26] [27] are considered watershed factors. These variables include Drainage area (DA km^2), Mean basin elevation (BE m), Basin slope (BS %), Basin perimeter (BP km), Basin shape factor (SF dimensionless), Drainage density (DD km/km^2), Valley slope (VS m/km), and the elongation ratio (ER dimensionless). The mean peak flow (Q_{MPF}) in Wabi-Shebele River Basin is positively correlated with variables: DA, BP, VL, and SF, where maximum correlation with DA, BP, and VL with a correlation coefficient of 0.92, 0.93, and 0.96, respectively. It is also evident from **Figure 1** above that the correlation value between flood discharge and annual rainfall increases with catchment size. Wabi watershed at Gode station has the largest catchment size showed maximum correlation value, and Robe watershed has the smallest catchment area in this study which exhibits minimum correlation value. Different studies also confirmed the impact of watersheds' size on peak flow. Rawas and Valeo [17] indicated that mean peak flow (Q_{MPF}) in arid watersheds is positively correlated with drainage size.

Similarly, the study conducted by Huang [28] showed that the drainage area affects not only the flow collection but also the time to peak flow. Moreover, the soil properties, mainly the soil infiltration rate, are sensitive variables for surface runoff generation. Coarse textured soils have big well-connected spaces and allow more water to infiltrate through them quite rapidly, while fine-grained soils dominated by clay have low infiltration rates due to their smaller-sized pore spaces [29]. Soils contain a large amount of sand and silt habit forming a crust and becoming more compacted, significantly reducing the infiltration rate. The mean peak flow (Q_{MPF}) in Wabi-Shebele River Basin is positively correlated with variables sand and loam.

4.3. Human Activities Factor

The land use and population density, and growth in the basin are considered human activity drivers for flooding [13] [22] [30]. The flood magnitude has a high positive correlation with cultivated land and population density and a strong negative correlation with forest cover.

4.4. Selection of Potential Flood Drivers in Wabi-Shebele

Using variables correlation matrix

The magnitude and type of correlation among the potential flood drivers from climate, watershed, and human variables (*i.e.*, MAR, DA, BS, VL, SF, DD, VS, ER, clay, sand, loam, forest, AGR, and PD) are estimated using the correlation matrix (**Table 3**) and scatter plot matrix (**Figure 2**). To identify significant predictors in watershed variables absolute value of correlation coefficient, R^2 exceeded 0.8, is selected.

Table 3. Correlation matrix in between variables.

Pearson's r	Q_{MPF}	DA	BE	BS	BP	VL	SF	DD	VS	ER	clay	sand	loam	MAR	forest	AGR	PD
Q_{MPF}	-																
DA	**0.999**	-															
BE	**−0.778**	−0.791	-														
BS	−0.018	−0.034	0.204	-													
BP	**0.999**	**0.999**	−0.803	−0.051	-												
VL	**1.000**	**1.000**	−0.791	−0.032	0.999	-											
SF	0.207	0.209	0.321	0.383	0.172	0.202	-										
DD	−0.594	−0.594	**0.859**	0.272	−0.624	−0.600	0.592	-									
VS	−0.560	−0.553	0.636	0.581	−0.583	−0.558	0.545	0.758	-								
ER	−0.295	−0.294	−0.223	−0.536	−0.257	−0.288	**−0.969**	−0.536	−0.517	-							
clay	−0.433	−0.413	0.257	−0.788	−0.414	−0.419	−0.173	0.314	−0.081	0.325	-						
sand	**0.613**	0.619	−0.754	0.468	0.615	0.619	−0.040	−0.537	−0.101	−0.146	−0.674	-					
loam	0.085	0.052	0.268	**0.681**	0.057	0.061	0.266	0.015	0.191	−0.319	−0.798	0.094	-				
MAR	−0.332	−0.348	**0.838**	0.412	−0.371	−0.349	0.738	**0.829**	0.624	−0.677	−0.058	−0.513	0.497	-			
forest	**−0.630**	−0.625	0.852	0.045	−0.648	−0.630	0.569	0.863	0.750	−0.408	0.390	−0.687	0.034	0.789	-		
AGR	−0.444	−0.446	0.347	**0.747**	−0.465	−0.448	0.093	0.499	0.655	−0.238	−0.257	0.259	0.135	0.250	0.154	-	
PD	−0.321	−0.317	−0.085	0.556	−0.318	−0.316	−0.361	−0.022	0.356	0.204	−0.313	0.530	−0.009	−0.297	−0.253	0.821	-

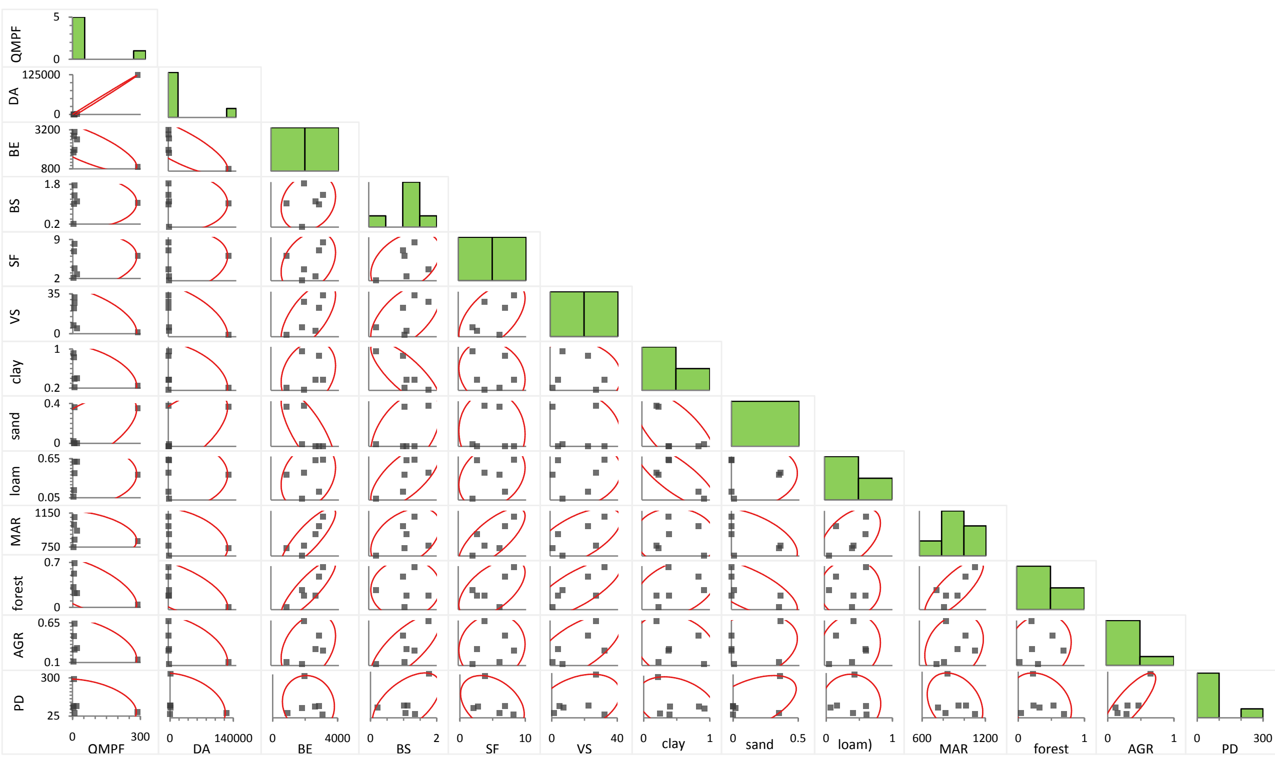

Figure 2. Scatter plot matrix for all pairs of variables. Note: Each plot shows the relationship between a pair of variables. The red ellipse contains the middle 75% of the neighborhoods and indicates whether the two variables are positively, negatively, or not correlated.

MAR, the only climatic variable in the flood drivers, positively correlates with the variables with the watershed factors (BE, BS, SF, DD, VS, loam) and human factors (forest, and AGR), with maximum correlation coefficients. Either positive or negative, the analysis indicates that MAR has a strong relation with the flood indices (Q_{MPF}) and variables of both other factors, which push it to be one of the candidates for flood drivers in the basin. DA is correlated with BP (R^2 = 0.99) and VL (R^2 = 1.00). SF is negatively correlated with ER (R^2 = −0.97), and BE is significantly positively correlated with DD (R^2 = 0.86). Both VS and BS do not exhibit any significant correlation with any watershed characteristics. Given these, five of the main variables were selected as independent watershed variables to avoid information redundancy or multi-col-linearity problems in the multiple regression analysis: DA, SF, BE, VS, and BS. The soil and human variables (Sand, Clay, loam, forest, AGR, and PD) have less col-linearity, directly considered the member of the PCA analysis for further selection of independent variables.

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is used to see multivariate relationships between potential driving factors and mean peak flow discharge (Q_{MPF}). PCA is one of the multivariate statistical techniques that can deal with highly correlated variables in regression [31] [32] [33]. In this study, the PCA is further applied to select the most influential drivers among the twelve predictors (MAR, DA, SF, BE, VS, BS, Sand, Clay, loam, forest, AGR, and PD) sorted through correlation analysis to achieve uncorrelated six PCs.

The eigenvalues represent the quantity of variability in the data, and they are presented in **Table 4**. The first three PCs explain the maximum degree of variability of the data set with a proportion of 45%, 26%, and 19%, respectively. They indicate about 90% of the influence of the flood induces possible mange with the variables in these three PCs. Therefore the variables in the three PCs are taken to develop the multiple linear regression equations among the flood drivers and flood indices.

The coefficients in **Table 5** show the linear combinations of variables that make each principal component. The absolute values near zero indicate that a variable contributes little to the PCs, whereas larger absolute values indicate variables that contribute more to the element. In the analysis, the first principal component has high negative associations with BE, VS, MAR, and forest and a high positive association with DA and sand, so this component primarily measures the basin altitude difference and land cover. The second component has

Table 4. Principal correlation analysis: Eigen analysis of the correlation matrix.

Name	PC1	PC2	PC3	PC4	PC5	PC6
Eigenvalue	7.66	4.41	3.25	1.34	0.35	0.00
Proportion	0.45	0.26	0.19	0.08	0.02	0.00
Cumulative proportion	0.45	0.71	0.90	0.98	1	1

Table 5. Principal correlation analysis: Eigen analysis of the correlation matrix.

Variables	PC1	PC2	PC3	PC4	PC5	PC6
Q_{MPF}	0.316	0.184	−0.153	−0.060	−0.106	−0.121
DA	0.317	0.179	−0.154	−0.089	−0.074	−0.392
BE	−0.344	0.023	−0.089	0.195	−0.210	−0.139
BS	−0.081	0.395	0.277	0.092	−0.096	0.300
BP	0.323	0.166	−0.149	−0.064	−0.074	0.211
VL	0.318	0.178	−0.152	−0.079	−0.076	−0.134
SF	−0.101	0.357	−0.299	−0.217	0.139	−0.059
DD	−0.321	0.114	−0.121	−0.220	−0.346	−0.372
VS	−0.281	0.210	0.096	−0.221	0.550	−0.194
ER	0.068	−0.410	0.212	0.241	0.075	−0.526
clay	−0.117	−0.361	−0.218	−0.343	−0.166	0.325
sand	0.234	0.222	0.292	−0.231	0.203	−0.074
loam	−0.033	0.306	0.056	0.652	0.058	0.007
MAR	−0.262	0.242	−0.222	0.170	−0.230	0.089
forest	−0.317	0.041	−0.223	−0.080	0.397	−0.048
AGR	−0.179	0.177	0.382	−0.240	−0.435	−0.228
PD	−0.044	0.042	0.533	−0.199	−0.011	0.172

high positive associations with BS, SF, and loam, so this component primarily measures the slope and shape of the catchment. The third component has a high positive association with sand, AGR, and PD, so this component primarily measures the basin farmland and population density.

The loading plot in Figure 3 visually shows the results for the first two components. From the graph, DA and sand indicate a small angle (<90˚) from the Q_{MPF} line, meaning the variables positively correlated to Q_{MPF}. The variables: forest, PD, BE, VS, and forest indicate angles related to 180˚, meaning they are negatively correlated to Q_{MPF}. However, the variables: BS, SF, and loam have no significant correlation with Q_{MPF} in Wabi-Shebele River Basin.

4.5. Relationship Development among Drivers and Flood Magnitude

A significance level (p-value) for all drivers is examined (Table 6). The selection criterion is set to $p < = 0.1$ in regression analysis. Based on this criterion, DA, sand, MAR, and forest are found as the significant ones to be used in the development of regression equations to estimate the Q_{MPF}. Therefore, Q_{MPF} can well be estimated from Model 3 in Table 6, where adjusted R^2 has the highest value and p-value is significant (<0.05). The multiple regression equation is:

$$Q_{MPF} = 6.39\text{MAR} + 0.66\text{DA} + 0.35\text{sand} - 0.62\text{forest} - 19.64 \quad (1)$$

In the equation, climate factor (*i.e.*, MAR), catchment size (*i.e.*, DA), sand coefficient, and land use cover (*i.e.*, forest) are the most influential exploratory

Table 6. Selection of regression model.

Q	Model	R^2	Adj.R^2	ΔAdj.R^2	p-value	Variables Used
Q_{MPF}	1	0.999	0.996		0.040	DA, BE, MAR, forest
	2	0.999	0.996	0.001	0.040	DA, MAR, AGR, PD
	3	**0.999**	**0.996**	**0.003**	**0.020**	**DA, sand, MAR, forest**
	4	0.995	0.989	−0.007	0.006	DA, MAR, forest
Q_5	1	0.999	0.999		0.010	DA, BE, VS, sand
	2	**0.999**	**0.999**	**0.000**	**0.010**	**DA, BE, VS, MAR**
	3	0.999	0.995	−0.004	0.040	DA, BE, VS, AGR
	4	0.994	0.984	−0.009	0.009	DA, MAR, forest
Q_{10}	**1**	**0.999**	**0.999**		**0.004**	**DA, BE, VS, sand**
	2	0.998	0.991	−0.008	0.060	DA, BE, MAR, AGR
	3	0.997	0.994	0.003	0.003	DA, MAR, AGR
	4	0.996	0.993	−0.001	0.000	DA, MAR
Q_{20}	1	0.998	0.991		0.060	DA, BE, VS, sand
	2	0.999	0.995	0.004	0.040	DA, BE, VS, MAR
	3	0.998	0.994	−0.001	0.050	DA, BE, MAR, AGR
	4	**0.998**	**0.996**	**0.002**	**0.001**	**DA, MAR, AGR**
Q_{50}	1	0.986	0.932		0.174	DA, BE, VS, MAR
	2	0.982	0.955	−0.023	0.099	BE, sand, MAR
	3	0.935	0.838	−0.006	0.090	BE, MAR, forest
	4	**0.899**	**0.833**	**0.008**	**0.030**	**BE, MAR**
Q_{100}	1	0.964	0.821		0.279	DA, BE, VS, sand
	2	0.976	0.881	0.060	0.228	DA, BE, VS, MAR
	3	0.912	0.781	−0.100	0.128	BE, MAR, AGR
	4	**0.907**	**0.845**	**0.024**	**0.020**	**DA, MAR**

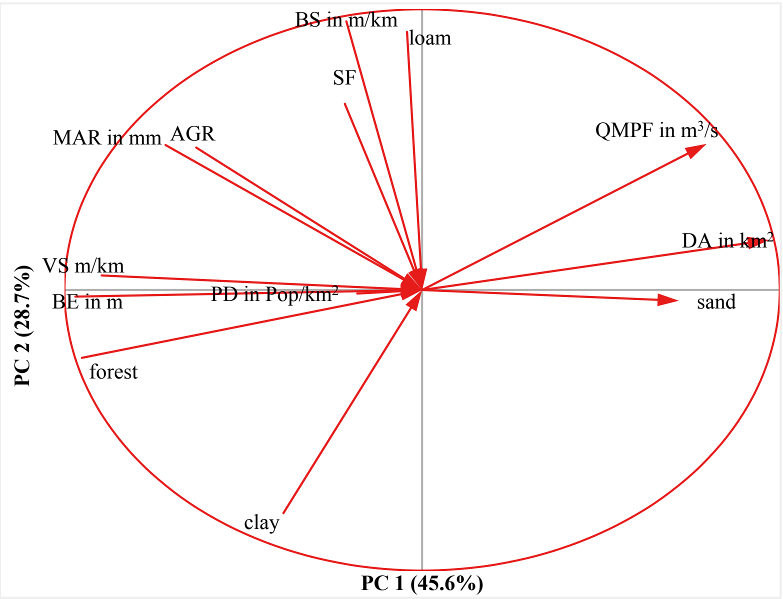

Figure 3. Two-dimensional correlation plot of coefficients of first two PCs (PC1 & PC2).

factors on flood quantiles, Q_{MPF} in the study area. Furthermore, forest function is negatively related to Q_{MPF}, meaning watersheds with high forest coverage yields less flood discharge than watersheds with less forest coverage. Similarly, the relationship between flood frequency and principal drivers is examined.

For Q_5, Model 2 with only four variables of DA, BE, VS, and MAR was selected as the best model to represent Q_5 estimation. It is noticed from Table 6 that watershed characteristics are the most influential factors of flood-peak frequency at 5 and 10-year return periods. On the other side, climate and human factors are most powerful in representing Q_{MPF} and flood-peak frequency at 20, 50, and 100-year return periods. The regression equations that describe the relationship between influential driving factors and different return periods flood-peak flows are:

$$Q_5 = 1.39\text{MAR} + 1.43\text{DA} + 2.98\text{BE} + 0.51\text{VS} - 17.67 \tag{2}$$

$$Q_{10} = 1.49\text{DA} + 2.44\text{BE} + 0.72\text{VS} - 0.59\text{sand} + 11.81 \tag{3}$$

$$Q_{20} = 3.63\text{MAR} + 0.76\text{DA} + 0.09\text{AGR} - 11.59 \tag{4}$$

$$Q_{50} = 8.05\text{MAR} - 4.70\text{BE} - 6.59 \tag{5}$$

$$Q_{100} = 8.23\text{MAR} - 4.59\text{BE} - 7.42 \tag{6}$$

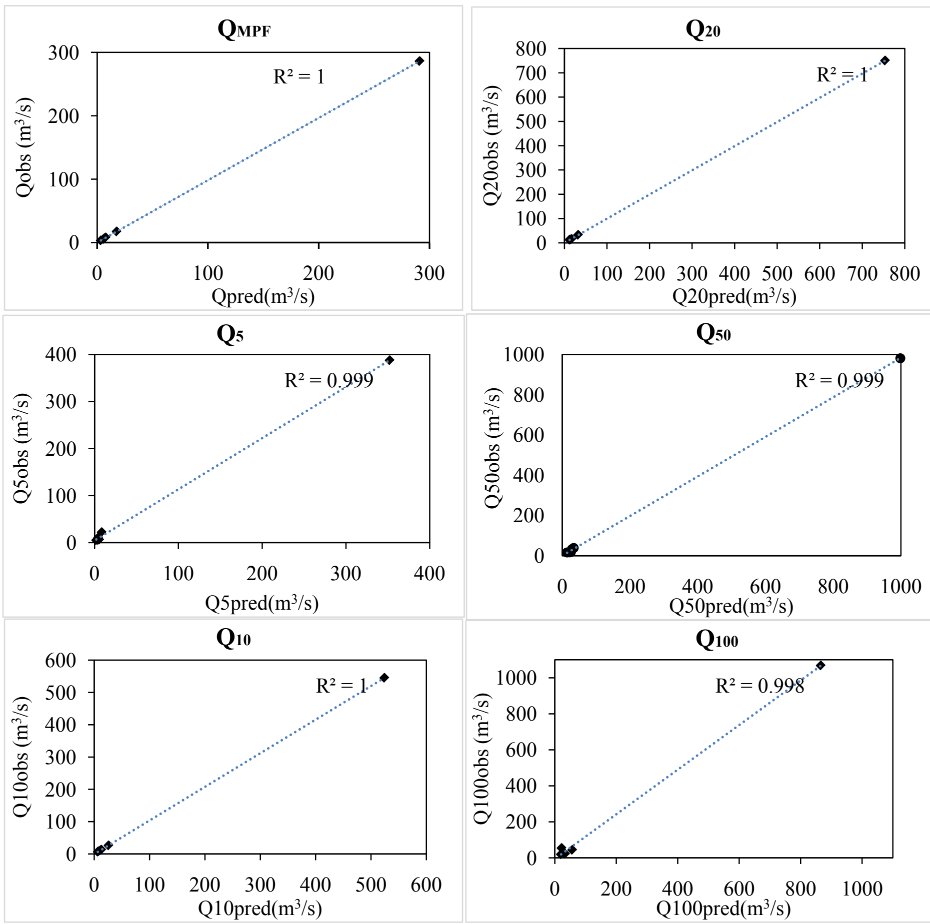

Figure 4. Comparison of observed and predicted flood quantiles of all quantiles and sample watersheds.

Table 6 summarizes the evaluation statistics from the regression model to MAE, NSE, RMSE, and R^2 based on observed and predicted flood values for all the six flood quantiles. A value close to zero is preferable for MAE as zero indicates no error in prediction. It is seen that all the MAE values for all quantiles lie between 1 and 44. The smallest value of MAE is found in the case of Q_{MPF} and Q_{20} estimations. It is noted that except for Q_5 and Q_{100}, most flood quantiles estimations are evaluated as good values. **Figure 4** shows plots of predicted quantiles over observed flood quantiles. These plots generally present a good agreement between the predicted and observed flood quantiles. Off courses in a few cases of underestimating the flood magnitude. For instance, the observed flows for Q_5 are range between 5.08 to 387.67 m^3/s, while the predicted values range from 2.04 to 352 m^3/s.

5. Conclusion

The major flood drivers and flood generation mechanisms in Wabi-shebele River basin were assessed using observed mean peak stream flow observed at six hydrological gauging stations in the basin. The six gauging stations have varied catchment areas with a range of between 169 to 124,108 km^2. The threshold (3^{rd} quartile) magnitude and frequency (POTF) that occurs over ten years of record, is used to build the flood dataset. Sixteen climatic, watershed and human factors were extracted and computed using GIS, Pearson's correlation analysis, Principal Correlation Analysis (PCA). Eight of them (MAR, DA, BE, VS, sand, forest AGR, PD) are identified as the most influential variables in flood formation of the basin. Moreover, mean annual rainfall (MAR), drainage area (DA), and lack of forest cover are explored as the principal driving factors for flood peak discharge in Wabi-Shebele River Basin. In other directions, watershed slope (BS), catchment shape factor (SF), fraction of loam and clay soil coverage are separated as less influential factors and the possibility of substituting of them by the most influential factors during quantification modeling ascertained. Moreover, larger watersheds with higher elevation and agricultural/farmlands lead to larger flood-peak flow in all investigated return periods. Finally, regression equations are developed to estimate flood quantiles using identified driving factors that are used for different planning and designing of infrastructures in the basin.

Acknowledgements

We would like to thank the Ethiopian National Meteorology Service Agency and Minis-try 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] Doocy, S., Daniels, A., Murray, S. and Kirsch, T.D. (2013) The Human Impact of Floods: A Historical Review of Events 1980-2009 and Systematic Literature Review.

PLoS Currents Disasters, **5**, 27 p.

[2] Jha, A.K., Bloch, R. and Lamond, J. (2012) Cities and Flooding: A Guide to Integrated Urban Flood Risk Management for the 21st Century. The World Bank. https://openknowledge.worldbank.org/handle/10986/2241

[3] Assefa, T.H. (2018) Flood Risk Assessment in Ethiopia. *Civil and Environmental Research*, **10**, 35-40.

[4] Nied, M., Hundecha, Y. and Merz, B. (2013) Flood-Initiating Catchment Conditions: A Spatio-Temporal Analysis of Large-Scale Soil Moisture Patterns in the Elbe River Basin. *Hydrology and Earth System Sciences Discussions*, **9**, 10053-10094. https://doi.org/10.5194/hessd-9-10053-2012

[5] Erena, S.H. and Worku, H. (2018) Flood Risk Analysis: Causes and Landscape Based Mitigation Strategies in Dire Dawa City, Ethiopia. *Geoenvironmental Disasters*, **5**, Article No. 16. https://doi.org/10.1186/s40677-018-0110-8

[6] Kundzewicz, Z.W. and Robson, A.J. (2004) Change Detection in River Flow Records—A Review of the Methodology. *Hydrological Sciences Journal*, **49**, 7-19. https://doi.org/10.1623/hysj.49.1.7.53993

[7] 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**, 1-28. https://doi.org/10.1080/02626667.2013.857411

[8] NDRMC (2018) Annual Report, Federal Democratic Republic of Ethiopia National Disaster Risk Management Commission, Early Warning and Emergency Response Report Directorate.

[9] Tadesse, T., Haigh, T., Wall, N., Shiferaw, A., Zaitchik, B., Beyene, S., Berhan, G. and Petr, J. (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

[10] Broxton, P., Troch, P.A., Schaffner, M., Unkrich, C. and Goodrich, D. (2014) An All-Season Flash Flood Forecasting System For Real-Time Operations, *Bulletin of the American Meteorological Society*, **95**, 399-407. https://doi.org/10.1175/BAMS-D-12-00212.1

[11] ERCS (2005) Ethiopia: Heavy Rains and Floods. Ethiopian Red Cross Society. Information Bulletin No. 2/2005, Addis Ababa.

[12] IWMI (2015) Drivers of Hydrological Dynamics in the Bale Eco-Region. SHARE Bale Eco-Region Research Report Series No. 7.

[13] Ministry of Water Resources (MoWR) (2004) Wabi Shebele River Basin Integrated Master Plan Study Project, Water Resources, Part 2: Hydrology, Land Use/Cover Study.

[14] Keast, D. and Ellison, J. (2013) Magnitude Frequency Analysis of Small Floods Using the Annual and Partial Series. *Water*, **5**, 1816-1829. https://doi.org/10.3390/w5041816

[15] Moes, D. (2015) Power-Law flood Frequency Analysis of Selected Queensland Stream Gauges. Geography Corpus ID: 132395939.

[16] Enzel, Y., Ely, L.L., House, P.K., Baker, V.R. and Webb, R.H. (1993) Paleoflood Evidence for a Natural Upper Bound to Flood Magnitudes in the Colorado River Basin. *Water Resources Research*, **29**, 2287-2297. https://doi.org/10.1029/93WR00411

[17] Al-Rawas, G.A. and Valeo, C. (2010) Relationship between Wadi Drainage Characteristics and Peak-Flood Flows in Arid Northern Oman. *Hydrological Sciences Jour-*

nal, **55**, 377-393. https://doi.org/10.1080/02626661003718318

[18] Kendall, M.G. (1975) Rank Correlation Methads. Charles Griffin & Company, London.

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

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

[21] Woldegebrael, S.M., Berhanu, B. and Melesse, A.M. (2019) Historical Flood Events and Hydrological Extremes in Ethiopia. In: Melesse, A.M., Abtew, W. and Senay, G., Eds., *Extreme Hydrology and Climate Variability, Monitoring, Modeling, Adaptation and Mitigation*, Elsevier, Amsterdam, 379-384.

[22] Kundzewicz, Z.W. and Stoffel, M. (2016) Anatomy of Flood Risk. In: Kundzewicz, Z.W., 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

[23] 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

[24] Hall, J., Arheimer, B., Borga, M., Brázdil, R., Claps, P., Kiss, A., Kjeldsen, T.R., Kriauciuniene, J., Kundzewicz, Z.W., Lang, 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

[25] Stoffel, M., Wyżga, B., Niedźwiedź, T., Ruiz-Villanueva, V., Ballesteros-Cánovas, J.A. and Kundzewicz, Z.W. (2016) Floods in Mountain Basins. In: Kundzewicz, Z.W., Stoffel, M., Niedźwiedź, T. and Wyżga, B., Eds., *Flood Risk in the Upper Vistula Basin*, Springer, Cham, 23-37. https://doi.org/10.1007/978-3-319-41923-7_2

[26] Patra, K.C. (2010) Hydrology and Water Resources Engineering. Alpha Science International Ltd., Oxford.

[27] Newson, M.D. (2005) Hydrology and the River Environment. Clarendon Press, Oxford.

[28] Huang, P.C. (2020) Analysis of Hydrograph Shape Affected by Flow-Direction Assumptions in Rainfall-Runoff Models. *Water*, **12**, Article No. 452. https://doi.org/10.3390/w12020452

[29] Wudineh, F., Moges, S. and Kidanewold, 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

[30] Liu, Y.S., Yuan, X.M., Guo, L., Huang, Y.H. and Zhang, X.L. (2017) Driving Force Analysis of the Temporal and Spatial Distribution of Flash Floods in Sichuan Province. *Sustainability*, **9**, Article No. 1527. https://doi.org/10.3390/su9091527

[31] Li, Q.H., Shang, L.H., Gao, T.J., Zhang, L., Ou, T., *et al.* (2014) Use of Principal Component Scores in Multiple Linear Regression Models for Simulation of Chlorophyll-a and Phytoplankton Abundance at a Karst Deep Reservoir, Southwest of China. *Acta Ecologica Sinica*, **34**, 72-78. https://doi.org/10.1016/j.chnaes.2013.11.009

[32] 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*, **12**, Article No. 781. https://doi.org/10.3390/w12030781

[33] Elsheikh, R.F.A., Ouerghi, S. and Elhag, A.R. (2015) Flood Risk Map Based on GIS, and Multi Criteria Techniques (Case Study Terengganu Malaysia). *Journal of Geographic Information System*, **7**, 348-357. https://doi.org/10.4236/jgis.2015.74027