

An evaluation of the effects of climate change on flood frequency in the Luvuvhu River Catchment, Limpopo Province, South Africa

P. M. Kundu, R. L. Singo, J. O. Odiyo & R. N. Nkuna

*Department of Hydrology and Water Resources,
University of Venda, South Africa*

Abstract

Climate change poses a great risk to natural resources and agriculture in the Luvuvhu River Catchment. The catchment has experienced floods resulting from higher than normal rainfall associated with the Intertropical Convergence Zone and cyclones, which have caused enormous damage to property and impacted negatively on fauna and flora and human livelihoods. In order to understand the dynamics involved in the effects of climate change in the catchment, annual maximum flow data was used to evaluate flood frequencies. The distribution models used in the study included the Generalized Extreme Value distributions, the Gumbel Extreme Value type I distribution, the Log-normal distribution and the Log Pearson type III distribution. The extreme value analysis showed that the Log-normal and Log-Pearson type III distributions provided the best fit, which could be used to derive the probability of occurrence of flood events. The results showed a general increase in the frequency of extreme events, accompanied by floods of higher magnitude.

Keywords: catchment, climate change, distribution models, flood frequency, rainfall.

1 Introduction

Luvuvhu River Catchment is one of the regions in South Africa where floods have caused enormous damage to both property and life and impacted negatively on infrastructure and development as shown in Figure 1, where people had to wade through water after the bridge was washed away in 2013 at Hoedspruit and Buckbuckridge area. Prolonged rainfall over large drainage basins is associated



with climate change and tropical cyclones or the intense depressions of mid-latitudes. They vary from semi-predictable seasonal rains over wide geographic areas, which give rise to the annual wet-season floods in tropical areas, to almost random convectional storms over small basins. Although the mean annual runoff is estimated at 333.6 million m³ under natural conditions, the risk of floods has been felt in the past, with the major flood occurring during the 1999/2000 rainy season.



Figure 1: People crossing the river after the bridge was washed away at Hoedspruit and Buckbuckridge area.

Duaibe [1] found that a number of human induced factors contributed to floods in South West Pacific. The factors included deforestation of catchment areas; increased population density along river banks; inadequate land use planning; zoning and control of floodplain for development; inadequate drainage, particularly in cities, and inadequate management of discharge from river. The South African National Disaster Management Center estimated that more than 100,000 people live along rivers, below water levels previously reached by floods. Therefore, flood frequency associated with higher magnitudes are of concern and require attention by water resources planners and hydraulic designers. Olofintoye *et al.* [2] used different frequency distribution models for hydrologic analysis to predict design floods in Luvuvhu River Catchment. Though several probability models have been developed to describe the frequency distribution of extreme hydrologic events, major problems arise when selecting the best method to use since there is no general agreement as to which distribution should be used. The statistical approach uses probabilistic methods to model flood events, while the deterministic approach requires the use of physically-based models. Of the two approaches, statistical probability methods remain the most widely used in the scope of hydrology [3].

2 The study area

Luvuvhu River Catchment shown in Figure 2 is located between latitudes $22^{\circ}17'57''\text{S}$ and $23^{\circ}17'31''\text{S}$ and longitudes $29^{\circ}49'16''\text{E}$ and $31^{\circ}23'02''\text{E}$ in Limpopo province of South Africa. It covers an area of approximately 5941 km², situated on a plateau of 1312 meters above sea level. The topography of the catchment is undulating with a few prominent ridges extending in an east-west direction, whereas the monotony of the area is relieved by separated hills and ranges. Luvuvhu River rises as a steep mountain stream in the southeasterly slopes of the Soutpansberg Mountain range, flows through Kruger National Park, and empties into the Limpopo River at the border with Mozambique and Zimbabwe.

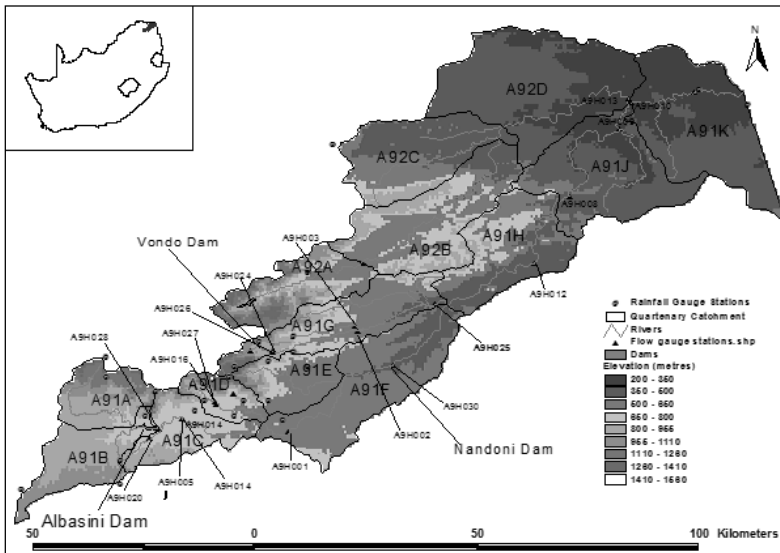


Figure 2: The study area.

3 Materials and methods

The magnitude of a flood can be predicted by using return period and probability distribution functions. Flood events are a succession of natural events which, as far as can be determined, do not fit any one specific known statistical distribution. The objective of frequency analysis is therefore to relate the magnitude of events to their frequency of occurrence through probability distribution. To define and make flood probabilities predictable, it is necessary to assign a distribution with the best goodness of fit. Expected peak flow requires the determination of the coefficient of variation as well as the coefficient of skew.

The probable maximum precipitation, which is the quantity of precipitation that is close to the physical upper limit for a given duration over a particular basin was determined. The mean annual rainfall was estimated by use of geostatistical

techniques involving interpolation and regression techniques within a GIS environment. Given that the bulk of the weather stations were located to the west of Vondo and Nandoni dams and the insignificant correlation of the intra-quaternary rainfall patterns, uncertainties were expected in the estimation of aerial rainfall. Due to the inadequacy of interpolation and regression techniques when the spatial distribution of stations is skewed, a “driver” station approach was selected for use. The flow data to determine the flood frequency included release data from the Albasini, Nandoni and Vondo dams and the discharge measurements recorded at Vredenburg weir (A9H025) on the Mutshindudi River, Weltevrede weir (A9H001) and at the Mhingwa weir (A9H012) on Luvuvhu River.

3.1 Flood frequency and magnitude

Flood peaks corresponding to return periods of 2, 5, 10, 50, 100 and 200 years were estimated. To assess the effect of floods within built-up areas, the degree of protection was considered for floods which occur once in 100 or 200 years. Each computed flood magnitude was determined at 95-percent confidence interval which contained the true flood magnitude for a particular exceedence probability. Annual maximum daily flood discharge series were extracted for each hydrological year. Moving-average filtering reduced the effects of random variations based on the premise that the systematic component of a time series exhibits autocorrelation while the random fluctuations were not autocorrelated. The coefficient of variation was applied in the data to measure the consistency and the steepness for the frequency curves in stream flow data. A probability plot was used for assessing competing distributions to select the one which provided the best fit. EasyFit software [4] was used to analyse flood frequency in the catchment. Four frequency distributions including the Generalized Extreme Value distributions (GEV), the Gumbel's Extreme Value type 1 (EV I) distribution, the Log-normal (LN) distribution and the Log-Pearson type III (LP3) distribution were compared for flood estimation.

3.2 Generalized Extreme Value (GEV) distributions model

Hosking and Wallis [5] approximated that the design flood at a desired return period is given by equation 1:

$$Q_t = u + \left(\frac{\alpha}{\kappa} \right) \left\{ 1 - \left(-\log \left(\frac{T-1}{T} \right) \right)^k \right\} \quad (1)$$

where, T = desired return period

$$u = \lambda_1 - \alpha \{ 1 - \Gamma(1+k) \} / k, \quad \alpha = \frac{\lambda_2 k}{(1-2^{-k})\Gamma(1+k)}$$

$$k = 7.8590 + 2.9554c^2, \quad c = \frac{2}{3 + \tau_3} - \frac{\ln 2}{\ln 3}$$



where: Γ = gamma function, τ_3 = L-moment ratio and λ_1 and λ_2 = first and second L-moment approaches.

The τ_3 , λ_1 and λ_2 were derived from the Probability Weighted Moments equations following Cunnane [7].

The location parameter (u) described the shift of the distribution on the horizontal axis; the scale parameter (α) described how spread out the distribution was, and defined where the bulk of the distribution lay. As the scale parameter increases, the distribution will become more spread out. The shape parameter (κ) strictly affects the shape of the distribution and governs the tail of each distribution. The shape parameter is derived from skewness, as it represents where the majority of the data lies, which creates the tail(s) of the distribution.

3.3 Gumbel's Extreme Value type 1 (EV1) distribution model

The probability of occurrence of an extreme event, equal or larger than a value, is given by Gumbel [6] as:

$$P(X \geq x_0) = 1 - e^{-e^{-y}} \quad (2)$$

where P = the probability of occurrence, X = the event of the hydrologic series, x_0 = the desired value of the event and y = is a dimensionless variate given as:

$$y = \alpha(x - a) \quad (3)$$

where X = the variate value, $a = \bar{x} - 0.45005\sigma_x$ and $\alpha = 1.2825/\sigma_x$, σ_x = standard deviation of variate X and \bar{x} = mean of the variate X .

The variate can thus be calculated as:

$$y = \frac{1.2825(x - \bar{x})}{\sigma_x} + 0.577 \quad (4)$$

where values 1.2825 and 0.577 are the constants for the reduced mean and reduced variate, respectively.

In practice, it is the X for a given P that is required hence, the equation $P(X \geq x_0) = 1 - e^{-e^{-y}}$ becomes, $-\ln[-\ln(1 - P)]$. Since the return period is given by the reciprocal of probability, ($T_r = \frac{1}{P}$), the value of the reduced variate y_T for a given T_r becomes:

$$y_T = -\ln \left[\ln \left(\frac{T_r}{T_r - 1} \right) \right] \quad (5)$$

3.4 Log-normal distribution model

In log-normal distribution, the data is said to be normally distributed with the mean and standard deviation. It works best for distributions where variables cannot be negative. Graphical estimates of either flood magnitudes or probabilities can be computed using standard normal distribution as:

$$Z = \frac{Y - \bar{Y}}{S_y} \quad (6)$$

where Z = standard normal distribution, Y = value of random variable, \bar{Y} = the mean and S_y = standard deviation.

3.5 Log-Pearson type III (LP3) distribution model

The log-Pearson Type III distribution uses log transformation of the data as a base method for flood flow frequency studies. The estimated discharge values X_T , for a given period can be evaluated using the logarithm of the design flood given as:

$$X_T = \log Q_T = X_{av} + K \sigma_x \quad (7)$$

where Q_T = the discharge for the estimated T-year return period, K = the probability factor based on n-years recurrence interval, X_{av} = the mean of the logarithms of annual peak flows at the stream flow-gauging station, X_T and σ_x = the standard deviation about the mean of the logarithms of annual peak flows.

The skewness coefficient G given in equation 8 is computed as an important hydrological characteristic which gives a measure of shape of a sampling distribution.

$$G = \frac{n \sum (X - \bar{X})^3}{(n-1)(n-2)S^3} \quad (8)$$

where X = annual peak flow or logarithm of annual peak flow, n = length of data set, S = standard deviation.

4 Results and discussion

The area receives one cycle of rainfall that extends from October of the previous year to April of the following year, while the dry season runs from May to September. The spatial distribution of mean annual rainfall for the rainy season was as shown in Figure 3. The estimation of areal rainfall in the catchment has many uncertainties arising from the spatial variation of precipitation. The estimation is complicated by the fact that the distribution of active rain gauges in the catchment was uneven with the majority located in the upper reaches of the catchment.

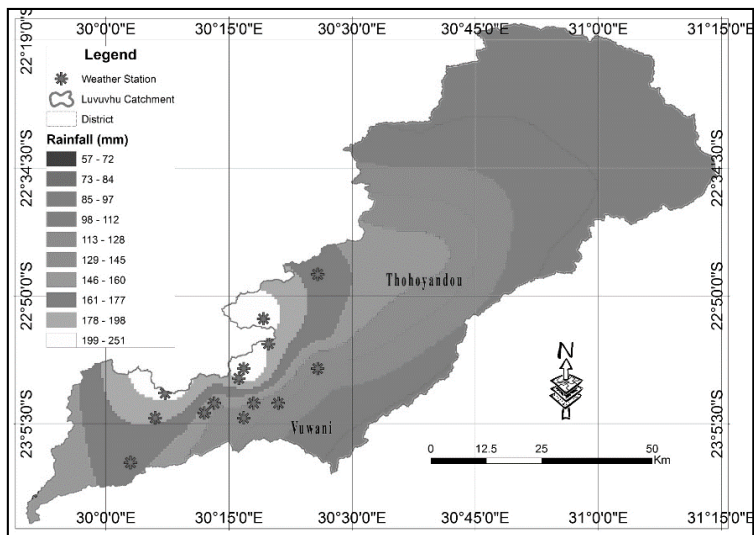


Figure 3: Mean annual rainfall.

The 50-year mean instantaneous flood flow was as shown in Figure 4. The highest measured flow of $106.72 \text{ m}^3/\text{s}$ was recorded during the 1976/77 hydrological year, while the lowest flood flow of $0.859 \text{ m}^3/\text{s}$ was recorded in 1991/92. A 5-year moving average was smoothed in the data to highlight significant changes in the trend. The smoothed trend showed significant hydrological conditions which may suggest that erratic rainfall associated with climate change may have caused flood peaks to increase. The coefficient of variation indicated that the distribution of flood flows was highly variable.

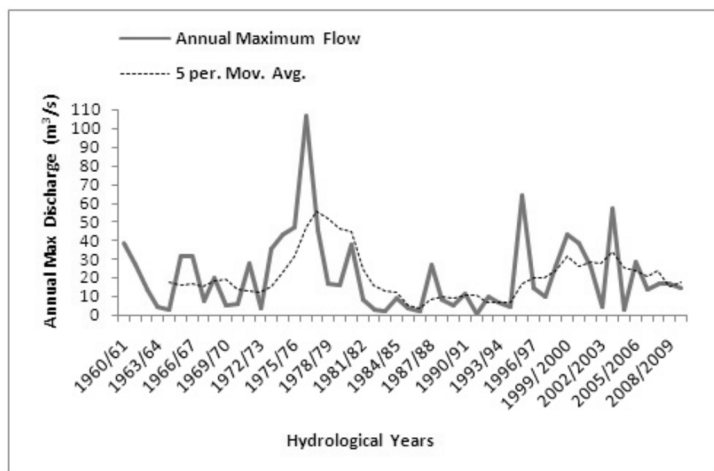


Figure 4: Annual maximum flow.

4.1 Goodness of fit measures

The goodness of fit tests were used to analyze how well the data fitted the given distributions. The tests described the differences between the observed values and calculated values from the distribution and could be used to reject possible distributions. The linearity of the probability plot indicated the goodness of fit as shown in Figure 5c, where the LN and LP3 showed more linear plots than GEV and EVI.

The histogram of annual maximum flood in Figure 5a showed a unimodal distribution which was skewed to the right. The probability density function showed that the LN and LP3 exhibited similar probability densities with higher frequencies than that of GEV and EV1 distributions. The cumulative distribution function in Figure 5b showed the non-exceedence probability for a given magnitude. The skew coefficient of the station record is sensitive to extreme events and is a good indicator of climate change. But it is difficult to obtain accurate skew estimates from small samples since it requires the use of at least 40 stations. The mean skew coefficient should be used as it provides the most accurate estimate. If the station skew is greater than +0.4, tests for high outliers are considered first.

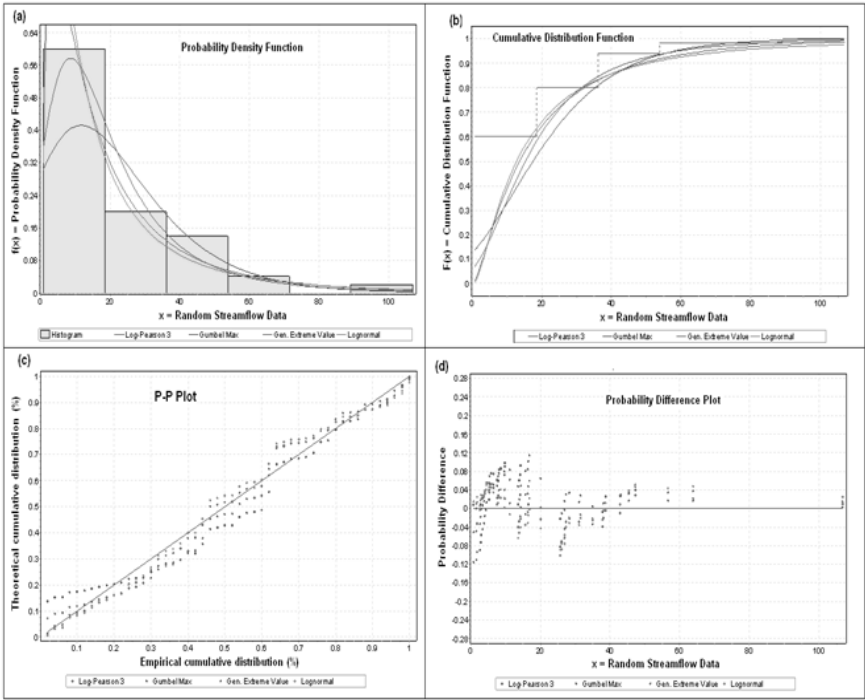


Figure 5: (a) Probability density functions, (b) cumulative distribution functions, (c) probability-probability plots and (d) probability difference plots.



If the station skew is less than -0.4, then tests for low outliers are considered first. Where the station skew is between -2 to -0.4, tests for both high and low outliers should be applied before eliminating any outliers from the data set. The skew in this study was found to be -0.4, which made it necessary for tests of low outliers to be considered first. Based on the probability plots, the distribution with the highest frequency was selected to ensure that designs would be based on the highest safety standard for the probable maximum flood. In this study, the LN and LP3 probability distributions were selected as the best valid distributions to model flood probability in the catchment.

The computed parameters for the four distributions were as shown in Table 1. In general, a distribution with a larger number of parameters would be able to model the input data more accurately than a distribution with a lesser number [7]. In this case, GEV had a positive (k) parameter which showed that it belonged to the EV3 distribution. The EVI and LN distributions had two parameters, (u) and (σ) which indicated that they had a k -value of zero.

Table 1: Estimated parameters for the four distributions.

	Distribution	Parameters
1	Gen. Extreme Value (GEV)	$K = 0.22411$ $\sigma = 11.52$ $\mu = 10.939$
2	Gumbel (EV I)	$\sigma = 15.753$ $\mu = 11.739$
3	Log-Pearson 3 (LP3)	$\alpha = 30.839$ $\beta = -0.19511$ $\gamma = 8.5669$
4	Log-Normal (LN)	$\sigma = 1.0726$ $\mu = 2.5497$

The estimated discharges using GEV, EV1, LN and LP3 distributions were as shown in Tables 2, 3, 4 and 5 respectively. The GEV model showed no significant difference in the discharges for all return periods. The LN model showed very large amounts for the 100 and 200 return periods.

Table 2: Estimated discharges for GEV distribution.

Return Period (Years)	Exceedence Probability (%)	Non-Exceedence probability (%)	Estimated Discharge (m ³ /s)
2	50	50	62.275
5	20	80	62.321
10	10	90	62.332
25	4	96	62.338
50	2	98	62.340
100	1	99	62.341
200	0.5	99.5	62.342

Table 3: Estimated discharges for EV1 distribution.

Return Period (Years)	Exceedence Probability (%)	Non-Exceedence probability (%)	k-value	Estimated Discharge (m ³ /s)
2	50	50	-0.156	17.679
5	20	80	0.819	37.374
10	10	90	1.466	50.443
25	4	96	2.276	66.805
50	2	98	2.888	79.168
100	1	99	3.498	91.489
200	0.5	99.5	4.093	103.509

Table 4: Estimated discharges for log-normal distribution.

Return Period (Years)	Exceedence Probability (%)	Non-Exceedence probability (%)	Z-value	Estimated Discharge (m ³ /s)
2	50	50	0.0000	12.794
5	20	80	0.8416	31.871
10	10	90	1.2816	51.361
25	4	96	1.7507	85.424
50	2	98	2.0538	118.669
100	1	99	2.3264	159.490
200	0.5	99.5	2.576	209.072

Table 5: Computed discharges for LP3 distribution.

Return Period (Years)	Exceedence Probability (%)	Non-Exceedence probability (%)	K-value	Estimated Discharge (m ³ /s)
2	50	50	0.066	13.75
5	20	80	0.855	32.33
10	10	90	1.231	48.60
25	4	96	1.606	72.96
50	2	98	1.834	93.41
100	1	99	2.029	115.37
200	0.5	99.5	2.201	139.03



The four models showed increasing discharges at higher probabilities of exceedence for all return periods, which could be associated with the effects of climate change. The EVI and LP3 models showed the highest discharges which could be taken as the estimated limiting values for design purposes, especially for the 100 and 200 year return periods.

Based on the results, the computations showed that an increase in the peak discharges was to be expected, especially for the discharge range corresponding to small and middle floods. Given the rapid land use changes in the catchment, a significant rise in water levels would lead to an increase in potential flood damages, particularly for flood events of lower to medium extremity, reducing flood security for existing protection facilities. The effect would however be less for events with a lower probability of occurrence involving large floods.

5 Conclusion

The flood frequency analysis showed increasing discharges at higher probabilities of exceedence for all return periods. The results showed that an increase in the peak discharges was to be expected, especially for small and middle floods. This could be associated with the effects of climate change and anthropogenic factors. Given the rapid land use changes in the catchment, a significant rise in water levels would lead to an increase in potential flood damages, particularly for flood events of lower to medium extremity, reducing flood security for existing protection facilities. However, a distribution with the best goodness of fit could be a powerful tool for estimating and calculating the return period for floods anywhere along rivers in the catchment. The EVI and LP3 models showed the highest discharges which could be taken as the estimated limiting values for design purposes, especially for the 100 and 200 year return periods.

Acknowledgements

The authors would like to acknowledge the support of the South African Research Foundation (NRF) and the University of Venda for funding the study. We are grateful to relevant authorities at the Department of Water Affairs (DWA) for providing valuable historical data.

References

- [1] Duaibe, K. (2009) Human activities and flood hazards and risks in the south west Pacific: A case study of the Navua catchment area, Fiji Islands. Unpublished MSc thesis, University of Wellington, New Zealand.
- [2] Olofintoye, O.O., Sule, B.F., & Salami, A.W. (2009). Best-fit Probability distribution model for peak daily rainfall of selected Cities in Nigeria. *New York Science Journal*, 2(3), 1-12.



- [3] NOAA (2011). Flood Frequency Estimates for New England River Restoration Projects: Considering Climate Change in Project Design. US.DOC/NOAA fisheries services, USA.
- [4] MathWave Technologies (2011). EasyFit Software Version 5.5, USA.
- [5] Hosking JRM and Wallis JR (1995). A comparison of unbiased and plotting-position estimators of L-moments. *Water Resour. Res.* 31 (8) 2019-2025.
- [6] Gumbel, E.J. (1941). The return period of flood flows. *Ann. Math. Statist.* 12(2), 163-190.
- [7] Cunnane, C. (1989). Statistical Distributions for Flood Frequency Analysis. Operational Hydrology Report no. 33, World Meteorological Organization.

