

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Chapter

Fitting a Generalised Extreme Value Distribution to Four Candidate Annual Maximum Flood Heights Time Series Models in the Lower Limpopo River Basin of Mozambique

Daniel Maposa

Abstract

In this paper we fit a generalised extreme value (GEV) distribution to annual maxima flood heights time series models: annual daily maxima (AM1), annual maxima of 2 days (AM2), annual maxima of 5 days (AM5) and annual maxima of 10 days (AM10). The study is aimed at identifying suitable annual maxima moving sums that can be used to best model extreme flood heights in the lower Limpopo River basin of Mozambique, and hence construct flood frequency tables. The study established that models AM5 and AM10 were suitable annual maxima time series models for Chokwe and Sicacate, respectively. This study also revealed that the year 2000 flood height was a very rare extreme event. Flood frequency tables were constructed for the two sites Chokwe and Sicacate in the lower Limpopo River basin of Mozambique and these tables can be used to predict the return periods and their corresponding return levels at the sites and their neighbourhood. It is our hope that these long term forecasts will complement the short term flood forecasting and early warning systems in the basin in reducing the associated risk and mitigating the deleterious impacts of these floods on humans and property.

Keywords: moving sums, annual maxima, lower Limpopo River, generalised extreme value

1. Introduction

This chapter has its significance in disaster reduction in an economically challenged flood disaster prone developing country. The lower Limpopo River basin of Mozambique is one of the basins in Southern Africa that have not been deeply studied in terms of application of extreme value statistics. The basin suffers from extreme natural hazards that alternate between extreme floods and severe droughts. A lot of geoscientific work aimed at short-term forecasts and flood warning systems has been done in the basin. The present study is intended to complement the geoscientific work in the basin through shifting attention to long-term forecasts.

The combination of these two approaches is hoped to go a long way towards disaster reduction and mitigation efforts in the basin.

2. Background of the study

Hydrological extreme events such as floods and droughts have accompanied mankind throughout its entire history and these events are cyclic in nature. The twenty-first century, however, has been marked by an unusual number of natural disasters worldwide, among these events are the recent hurricane Matthew that devastated the Caribbean Islands of Haiti, parts of Jamaica and United States of America [1], the recent Nepal 2015 giant earthquake that killed more than 8000 people and injured more than 19,000 people [2], flooding and landslides in Brazil in 2011 and flooding in Mozambique and other parts of Southern Africa in 2000.

Natural disasters such as floods often pose an intolerable threat to society, hence a holistic approach is needed to understand such phenomena, predict such catastrophic events and mitigate the impact of these natural disasters. The lower Limpopo River in Mozambique has a history of worst floods and droughts than all other national and international rivers in Mozambique. The most catastrophic and expensive of these reported natural disasters in Mozambique were the year 2000 floods which killed a total of more than 700 people and caused economic damages estimated at US\$500 million. It is argued by the International Federation of Red Cross (IFRC) that aid money can buy more than seven times as much humanitarian impact if spent before a disaster rather than on post-disaster relief operations [3].

The chairman of the 2014 International Disaster and Risk Conference (IDRC) held in Davos, Switzerland, August 2014, Dr. Walter J. Ammann pointed out that the frequency and intensity of natural hazards such as floods and earthquakes are on the rise in these recent years [3]. In a separate study, a unique survey of 139 national meteorological and hydrological services carried out by the World Meteorological Organisation (WMO) in 2013 revealed that floods were the most frequently experienced extreme events worldwide over the course of the decade 2001–2010 [4]. Some studies have also shown that floods and droughts account for 90% of all the people that are affected by natural disasters [5]. According to Munich Re [6] the statistics of natural disasters for the year 2013 was dominated by floods that caused several billions of United States of America dollars in losses. Irina Bokova, the Director-General of UNESCO [7] stated that:

Every year, more than 200 million people are affected by natural hazards, and the risks are increasing – especially in developing countries, where a single major disaster can set back healthy economic growth for years. As a result, approximately one trillion dollars have been lost in the last decade alone. This is why disaster risk reduction is so essential. Mitigating disasters requires training, capacity building at all levels, and it calls for a change of thinking to shift from post-disaster reaction to pre-disaster action –this is UNESCO's position.

The present study considers floods in the lower Limpopo River basin of Mozambique. The lower Limpopo River basin is characterised by extreme natural climatic conditions alternating between extreme floods and severe droughts. Droughts affect the country on an average of 7–8-year cycle and are usually associated with the El Nino phenomenon which affects Southern Africa [8]. The provinces of Gaza and Inhambane, which house the lower Limpopo River basin,

are among the drought-prone regions due to the action of thermal anticyclones. Droughts usually result in increases in the prices of basic commodities and food aid due to food shortages. Additionally, the number of deaths and diseases increase during the periods of drought [8].

On the contrary, floods in the lower Limpopo River basin are mainly associated with tropical cyclones. The provinces of Gaza and Inhambane are, also, the most affected by extreme floods due to their low-lying nature. Floods result in heavy losses to human lives and damage to infrastructure including bridges, houses, roads and schools. Among the cyclones that occurred in the basin in the recent past are Claudete in 1976, Angela in 1978, Nadia in 1994, Eline, Gloria and Huday in early 2000s and Flavio in 2007. The year 2000 floods were due to cyclone Eline and were the most disastrous and expensive of all the floods in the basin [8].

While the characteristics of hydrologic extreme events depend on the regional climatic conditions and other factors, the magnitude, timing, duration, and frequency fall within predictable range and pattern over time. The annual maximum series (AMS), also known as block maxima [9], has long been employed in extreme value theory to estimate the distribution of extreme events such as flood flows, precipitation and wind speeds.

The main purpose of this paper is to identify suitable annual maxima flood heights moving sums time series models at Chokwe and Sicacate sites in the lower Limpopo River basin and to construct flood frequency tables for the basin at these sites in Mozambique. It is also our wish to suggest improvements in extreme flood frequency modelling of annual maximum flood heights in the basin.

The outline of the rest of the paper is such that Section 2 presents the research methodology, Section 3 presents the results and discussion of the findings, and finally Section 4 gives the concluding remarks.

3. Research methodology

In this section we present the data source and fundamental principles of extreme value theory [9], as well as a brief discussion of some goodness-of-fit tests.

3.1 Study sites, data and block maxima moving sums

Hydrometric data has been collected in Mozambique since the early 1930s. However, due to war and other external factors there were periods in which no data were collected at some stations. For this study we obtained hydrometric data for the lower Limpopo River for the sites Chokwe (1951–2010) and Sicacate (1952–2010) from the Mozambique National Directorate of Water (DNA), the authority responsible for water management in Mozambique. The data obtained were daily flood heights (in metres) and were time series in nature.

In statistics of extremes there are two fundamental realisations used in flood frequency analysis namely block maxima and partial duration series commonly known as peaks-over-threshold (POT) [9]. The approach used in this study is block maxima. In hydrological studies, when sample sizes are large it is natural to block observations by years [9, 10].

Since in flood frequency analysis the years are natural blocks, the flood heights data in this study were blocked by years. Sequential steps were taken to obtain annual maxima data from the daily flood heights data series. Further sequential steps were taken to obtain the annual moving sums of 2 days (AM2), 5 days (AM5) and 10 days (AM10). A generalised extreme value (GEV) distribution

was then fitted to the annual daily (AM1) flood heights and their corresponding moving sums.

3.2 Generalised extreme value model

The GEV distribution is a well-known distribution in statistics of extremes. Comprehensive details of probability framework of block maxima and the practical reasons for using block maxima over POT are given in [9, 11]. Dombry [12] proved the consistency of maximum likelihood (ML) estimators when using block maxima approach.

The GEV cumulative distribution function, G , is given in Eq.(1) as:

$$G(\mu, \sigma, \xi; x) = \begin{cases} \exp\left(-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi}\right), 1 + \xi \frac{x - \mu}{\sigma} > 0, \xi \neq 0, \\ \exp\left(-\exp\left(-\frac{x - \mu}{\sigma}\right)\right), x \in \mathfrak{R}, \xi = 0, \end{cases} \quad (1)$$

where μ , σ and ξ are the location, scale and shape parameters, respectively. The parameters of the GEV in (1) are estimated by the ML method [11].

The log-likelihood function for the GEV in (1) is given in (2):

$$l(\mu, \sigma, \xi; x) = -k \log \sigma - (1/\xi + 1) \sum_{i=1}^k \log \left[1 + \xi \left(\frac{x - \mu}{\sigma}\right)\right]_+ - \sum_{i=1}^k \left[1 + \xi \left(\frac{x - \mu}{\sigma}\right)\right]_+^{-1/\xi}, \quad (2)$$

where k is the number of blocks (years) and annual maxima flood height $x = (x_1, x_2, \dots, x_k)$.

3.3 Anderson-Darling and Kolmogorov-Smirnov tests

The goodness-of-fit of the GEV model to the annual maxima flood heights moving sums time series models was verified using Anderson-Darling (A-D) and Kolmogorov-Smirnov (K-S) tests. The A-D test is sensitive to the tails of the distribution, while the K-S test is sensitive to the centre of the distribution [13]. The moving sums time series models were ranked from 1 to 4, with 1 being the best according to the particular test. A model that attains the lowest value of the total rank (sum of A-D rank and K-S rank) satisfies the criteria for the best annual maxima moving sums time series model. Where there is a tie in the total ranks for two or more models, then the rank value of the A-D test is used as a tie-breaker (with smaller value being best) since the main emphasis in extreme value theory is in fitting the tails.

4. Results and conclusion

This section presents the results of the study. **Tables 1** and **3** present the ML estimates of the parameters of the GEV distribution for Chokwe and Sicacate, respectively, for the models AM1, AM2, AM5 and AM10. **Tables 2** and **4** present results for the goodness-of-fit of the GEV distribution to the annual maxima moving sums time series models for Chokwe and Sicacate, respectively. **Table 5** presents the flood frequency tables of the return periods and their corresponding return levels for Chokwe and Sicacate based on the best fitting models.

Model	μ	σ	ξ
AM1	4.2671	1.8242	-0.11031
AM2	8.4538	3.4794	-0.09947
AM5	19.671	7.9968	-0.07243
AM10	35.889	14.8720	-0.05874

Table 1.
 ML estimates of the GEV distribution parameters for Chokwe (1951–2010).

Model	K-S	Cv	R	A-D	Cv	R	Total
AM1	0.054	0.172	3	0.237	2.50	4	7
AM2	0.051	0.172	1	0.234	2.50	3	4
AM5	0.052	0.172	2	0.185	2.50	1	3
AM10	0.062	0.172	4	0.218	2.50	2	6

Key: Cv stands for critical value, R stands for rank, K-S is Kolmogorov-Smirnov, A-D is Anderson-Darling.

Table 2.
 Goodness-of-fit tests of the GEV distribution for Chokwe (1951–2010).

Model	μ	σ	ξ
AM1	6.190	3.4587	-0.49005
AM2	12.267	6.8603	-0.48351
AM5	29.440	16.466	-0.45154
AM10	54.217	29.875	-0.38701

Table 3.
 ML estimates of the GEV distribution parameters for Siccate (1952–2010).

Model	K-S	Cv	R	A-D	Cv	R	Total
AM1	0.093	1.737	3	0.394	2.50	4	7
AM2	0.094	1.737	4	0.370	2.50	3	7
AM5	0.089	1.737	2	0.324	2.50	2	4
AM10	0.053	1.737	1	0.217	2.50	1	2

Key: Cv stands for critical value, R stands for rank, K-S is Kolmogorov-Smirnov, A-D is Anderson-Darling.

Table 4.
 Goodness-of-fit tests of the GEV distribution for Siccate (1952–2010).

Return period	20	50	100	200	500
Chokwe (AM5)	32.89	36.12	38.54	40.96	44.15
Siccate (AM10)	103.58	115.66	124.72	133.75	145.66

Table 5.
 Return periods (years) and their corresponding return levels (m) for the two sites.

4.1 Time series plots of the data

Figure 1 gives the time series plot of the annual daily maximum (AM1) flood heights data at sites, Chokwe, Combomune and Sicacate. The AM1 series for Chokwe is for the period 1951–2010, for Combomune is for the period 1966–2010, and for Sicacate is for the period 1952–2010. Since the original data series for Chokwe and Sicacate are comparable in terms of the starting period, only the two sites were considered for further analysis in this chapter. That is, the series for Combomune site was dropped for further analysis. Therefore, the results for this study are based on the two sites, Chokwe and Sicacate.

Furthermore, since the fundamental approach of this study is on the use of extreme value statistics using the block maxima, only the annual maximum value for each year was recorded and plotted in **Figure 1**. It is assumed that these AM1 data series are independent and identically distributed (iid) since they are blocked by years [9–12]. Similarly, it is also assumed that the annual maxima moving sums are also iid.

Since extreme value analysis is used in this chapter the data used in this chapter was not divided into training period, test period and forecasting period as these are not necessary in extreme value theory (EVT). In EVT special attention is paid to the estimation of the extreme quantiles and their corresponding return periods.

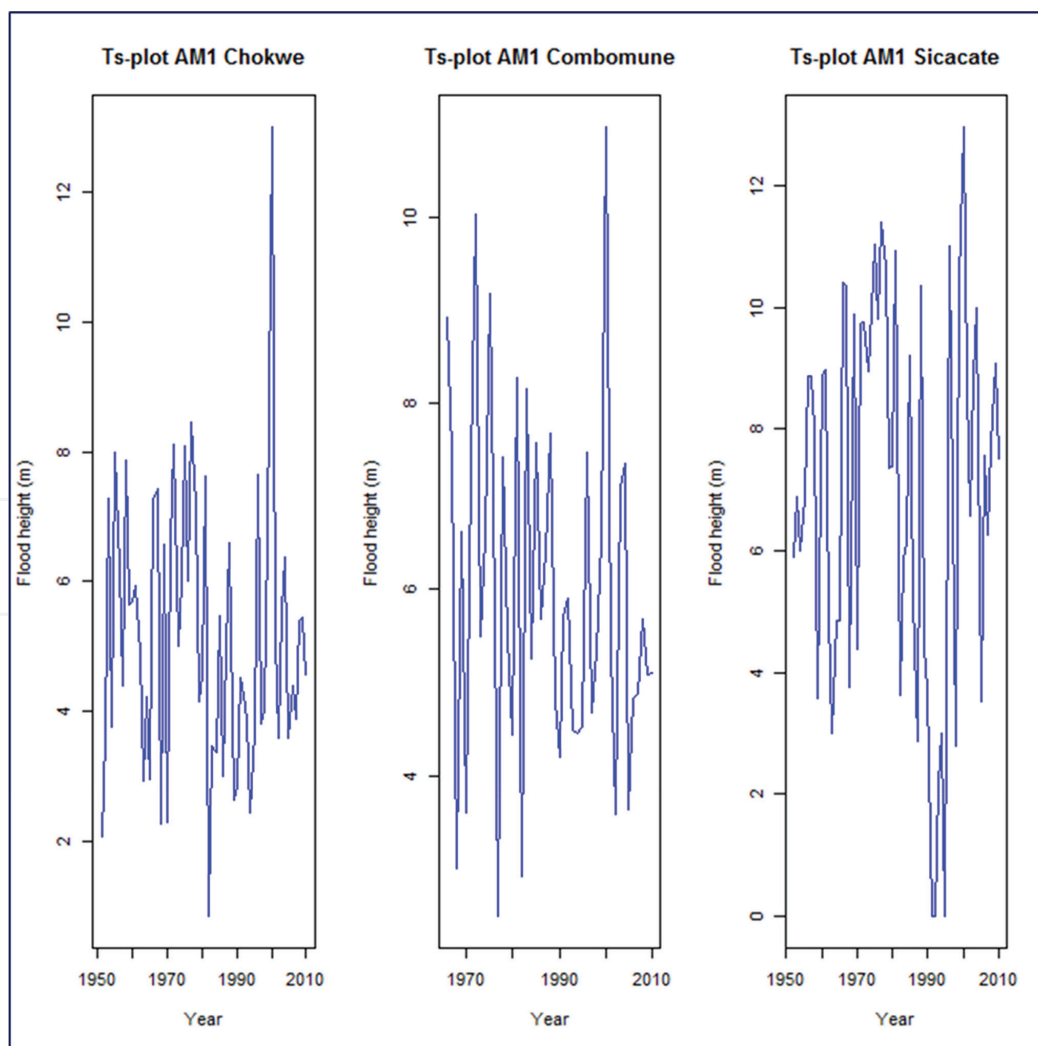


Figure 1. Time series plots of annual daily maximum (AM1) flood heights (in metres) at the three sites: Chokwe (1951–2010), Combomune (1966–2010) and Sicacate (1952–2010) along lower Limpopo River of Mozambique.

4.2 Chokwe models

The results in **Table 1** show that the shape parameter, ξ , is negative for all the models. This reveals that the distribution of floods in the lower Limpopo River basin at Chokwe site follows a short-tailed Weibull family of distributions. Further analysis on the values of the shape parameter at Chokwe site indicated that the distribution of floods at the site also belong to the Gumbel family of distributions since these values are not significantly different from zero (p-value >0.05) for all the models particularly AM2, AM5 and AM10.

Results in **Table 2** show that model AM5 had the lowest total rank of 3 suggesting that it is the best annual maxima moving sums time series model at the site of Chokwe. Consequently, forecasting at the site in this chapter is based on model AM5 for Chokwe.

4.3 Sicacate models

Results in **Table 3** show that the shape parameter, ξ , is negative for all the models at the Sicacate site. This suggests that the distribution of floods in the lower Limpopo River basin at Sicacate site follows a short-tailed Weibull family of distributions.

Results in **Table 4** show that model AM10 has the lowest total rank value of 2 which implies that it is the best annual maxima moving sums time series model for Sicacate site. Therefore, forecasting at Sicacate site in this chapter are based on model AM10.

4.4 Return level analysis

Flood heights (peaks) corresponding to return periods of 20, 50, 100, 200 and 500 years were estimated for flood disaster risk reduction in the basin (see **Table 5**). The best fitting annual maxima moving sums time series models AM5 and AM10 for Chokwe and Sicacate, respectively, were used to predict return levels for the selected return periods.

Table 5 presents results for selected return periods and their corresponding return levels for Chokwe and Sicacate sites based on the best fitting models AM5 and AM10, respectively. The predicted return levels were compared with the observed annual maxima moving sums. It was found that only one annual extreme flood peak, the year 2000 flood height of 124.89 m, exceeded the 100-year flood level for Sicacate site. As for Chokwe site, the year 2000 flood height of 50.44 m, exceeded the 500-year flood level which explains why it was very destructive at the site. The 100-year flood level for Chokwe was exceeded by three observed extreme events with two coming in the mid-1970s and the third one being the year 2000 flood peak. These findings suggest that the return periods of extreme flood heights in the lower Limpopo River basin of Mozambique can be used as a proxy for the return periods of the summer flood intensity. Results in **Table 5** can also be used to construct flood frequency curves [14].

5. Concluding remarks

The principal aim of this study was to identify suitable annual maxima moving sums time series models for the lower Limpopo River basin and to construct flood frequency tables for the basin at the sites of Chokwe and Sicacate in Mozambique. This study has successfully identified the prevailing models at the two sites Chokwe

and Sicacate in the lower Limpopo River basin. Annual maxima moving sums of 5 days (AM5) and 10 days (AM10) were identified as suitable time series models for Chokwe and Sicacate, respectively. It was also found in this study that the year 2000 flood height was a very rare extreme event. Flood frequency tables were constructed based on the identified models. It can be concluded that the findings in this study are promising for this vulnerable part of the basin. The findings obtained in this study are aimed to make contributions in the long term forecasts of floods in the basin to complement the much established and sponsored short term forecasts in the region.

Future research studies in the lower Limpopo River basin of Mozambique may advance this study through hierarchical modelling for spatial extremes and Markov chain Monte Carlo methods to the location and scale parameters in a changing climate.

Acknowledgements

We thank the Mozambique National Directorate of Water (DNA) in particular, Mr. Isac Filimone of DNA, who provided us with all the necessary data used in this study. We are also indebted to United Nations Office for the Coordination of Humanitarian Affairs-Southern Africa (OCHA) for providing us with weekly update reports of floods in Southern Africa, particularly for the lower Limpopo River basin of Mozambique. Finally we thank the University of Limpopo for providing the resources that facilitated the production of this article.

Conflict of interest

The authors declare that there is no 'conflict of interest' in this study.

Author details

Daniel Maposa
Department of Statistics and Operations Research, University of Limpopo,
Polokwane, South Africa

*Address all correspondence to: danmaposa@gmail.com

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] CNN. Hurricane Matthew devastates the Caribbean Islands of Haiti and parts of Jamaica and United States of America. CNN News; 2016
- [2] CNN. At least 66 dead after another powerful earthquake hits Nepal. 2015. Available at: <http://edition.cnn.com/2015/05/12/asia/nepal-earthquake/>
- [3] IDRC Davos. Integrative Risk Management-The role of science, technology & practice. In: 5th IDRC Davos 2014, Programme & Short Abstracts, 24-28 August 2014, The IDRC Davos 2014 Conference Proceedings. Davos, Switzerland; 2014. Available at: <http://idrc.info/outcomes/conference-proceedings/> [Access: 10 December 2017]
- [4] WMO. Global Climate 2001-2010: A decade of climate extremes – Summary Report, WMO no. 119. Geneva, Switzerland: World Meteorological Organization; 2013
- [5] Smakhtin V. Managing floods and droughts through innovative water storage solutions. In: 5th IDRC Davos 2014, The IDRC Davos 2014 Conference Proceedings. 2014. Available at: <http://idrc.info/outcomes/conference-proceedings/> [Access: 10 March 2015]
- [6] Munich R. Floods dominate natural catastrophe statistics in first half of 2013. Press Release, 9 July 2013. Munich: Munich Reinsurance; 2013
- [7] Arya AS, Boen T, Ishiyama Y. Guidelines for Earthquake Resistant Non-engineered Construction. Paris: UNESCO; 2014
- [8] Maposa D, Cochran JJ, Lesaoana M. Fighting flooding in Mozambique. INFORMS: ORMS Today. 2015;42(2) April 2015 Edition
- [9] Ferreira A, de Haan L. On the block maxima method in extreme value theory: PWM estimators. The Annals of Statistics. 2015;43(1):276-298. DOI: 10.1214/14-AOS1280. Available at: <http://arxiv.org/pdf/1310.3222.pdf> [Access: 10 December 2017]
- [10] Maposa D, Cochran JJ, Lesaoana M. Modelling nonstationary extremes in the lower Limpopo River basin of Mozambique. Jamba: Journal of Disaster Risk Studies. 2016;8(1):a185. DOI: 10.4102/jamba.v8i1.185
- [11] Coles S. An Introduction to Statistical Modelling of Extreme Values. London: Springer-Verlag; 2001
- [12] Dombry C. Maximum likelihood estimators for the extreme value index based on the block maxima method. In: ArXiv: 1301.5611v1 [math.PR]. 2013. Available at: <http://arxiv.org/pdf/1301.5611.pdf> [Access: 15 December 2017]
- [13] Maposa D, Cochran JJ, Lesaoana M. Investigating the goodness-of-fit of ten candidate distributions and estimating high quantiles of extreme floods in the lower Limpopo River basin, Mozambique. Journal of Statistics and Management Systems. 2014;17(3):265-283. DOI: 10.1080/09720510.2014.927602
- [14] Maposa D, Cochran JJ, Lesaoana M. Construction of flood frequency curves in the lower Limpopo River basin of Mozambique using Bayesian and Markov chain Monte Carlo methods. In: Proceedings of the 60th International Statistical Institute (ISI) World Statistics Congress, 26-31 July 2015, Rio de Janeiro, Brazil. 2015. ISBN: 978-90-73592-35-3. Copyright ISI 2015. Available at: www.isi-web.org