Application of Extreme Value Theory in Predicting Floods in Region 3, Zimbabwe.

Abstract

The use of extreme value theory to predict the possibility of floods has advanced our understanding of the occurrence of this rare phenomenon. The main goal of this research was to use the Block Maxima technique of the Extreme Value Theory to predict the occurrence of floods by estimating the probability and intensity of future flooding events. The data used in this research to predict floods came from the Meteorological Services Department of Zimbabwe in Harare, Zimbabwe. It covered five "convenience sampled" provinces in the purposefully sampled natural Region 3, Zimbabwe. The regions include the Midlands, Mashonaland East, Manicaland, Masvingo and Matabeleland South. The Extreme Value Theory analysis was applied using the Block Maxima technique on monthly rainfall. We produced a summary of the estimated parameters for the generalised extreme value. We used the Maximum Likelihood Estimation Method, quantile, and return level plots for parameter estimation. The results showed that the generalised extreme value distribution provides a better theoretical justification for evaluating extreme precipitation. According to this study's recommendations, early warnings and preparedness, risk assessment, and management can help control or lessen flooding in the future while saving the environment and providing humanitarian aid in times of need.

Keywords: Extreme value theory, floods, Block Maxima, Generalised Extreme Value, Region 3 of Zimbabwe

1. Introduction

The Extreme Value Theory has been increasingly used to model and forecast rare and extreme flooding events. By focusing on the tail behaviour of flood data, extreme value theory provides a robust framework for estimating the probabilities of extreme events that traditional statistical methods may not capture. (Mosavi et al. ,2018). Davison and Huser (2015) studied the properties of an extreme value theory. They found that the Generalized Extreme Value (G.E.V) is a single parametric form that combines the Gumbel, Fréchet and Weibull distributions and that the distribution of the maxima should closely resemble one of these three forms. The block maxima and peak over threshold are two approaches used in Extreme Value Theory to model extreme events. The block maxima is suitable when data is scarce, while the Peak Over Threshold is better for large datasets. We fit the block maxima to a generalised extreme value, and the peak over threshold to a generalised Pareto distribution.

Zimbabwe is becoming more vulnerable to climatic changes, and local climatologists have predicted disastrous effects on the environment, agriculture and food security, health, water resources, economic activities, human migration, and physical structure (Mavhura et al., 2017). The climatic changes lead to notable shifts in the onset of rains, the increased frequency of heavy rainfall events and more tropical cyclones. Globally, disasters are increasing in frequency and intensity; they are often unforeseen; serious, cause threats, and may bring injury and death in the worst cases (Alfieri et al., 2018). The people have little idea what to expect or how to prepare for floods. Early warning systems could be one way to save lives and property as heavy rainfall becomes more unpredictable. This study investigated and predicted floods using the Extreme Value Theory, a tool that traditional models did not implement for shortterm predictions (Mosavi et al., 2018). In Zimbabwe, there were meteorological records of rainfall received in the past years, and thus there was sufficient data for statistical analysis. The observations listed and described the rainfall figures of 20 stations in Zimbabwe's natural and agricultural Region 3 for 10 years (twenty seasons) from 2000 to 2010.

2. Statement of the problem

In Zimbabwe's Region 3, flooding is a recurring natural disaster that seriously harms people's lives, property, and agriculture. The increasing frequency and severity of floods, potentially worsened by climate change, have created an urgent need for effective flood prediction and management strategies. Traditional flood prediction models in this region often lack precision, particularly in capturing the extreme values that characterize rare but catastrophic flood events. To increase the accuracy of forecasting flood events, the block maxima method, a statistical technique used in extreme value theory (EVT), offers a viable answer by concentrating on the highest values within predetermined timeframes. Nevertheless, nothing is known about its efficacy compared to other models applied in the flood forecast process when it comes to its application to region 3 flooding in Zimbabwe. Its usage for predicting floods in Zimbabwe's Region 3 has received little attention. Its efficacy in comparison to other models in the area is also unknown. This research aims to address this gap by applying the block maxima method to historical hydrological and meteorological data from region 3, Zimbabwe, to evaluate its potential to predict extreme flood events more accurately. The research aims to create a trustworthy flood prediction model that assists proactive flood management, reducing damage and protecting residents.

3. Purpose of the study

This study aims to come up with an Extreme Value Theory Model for predicting floods in region 3 in Zimbabwe and propose appropriate recommendations.

4. Objectives of the study

The objectives of the study are:

i. To collect and analyse historical rainfall data from region 3 (Zimbabwe).
ii. We aim to apply the block maxima approach of extreme value theory to model extreme rainfall events.
iii. We aim to estimate flood, return levels and return periods for various probability thresholds.
iv. To generate Extreme Value Theory predictions of floods in region 3 (Zimbabwe).

5. Significance of study

The study aims to improve flood risk management and mitigation activities by offering a strong statistical foundation for forecasting floods in region 3, Zimbabwe. The results will contribute to:

- Improved flood forecasting and early warning systems.
- Informed decision-making for floodplain management and development.
- Enhanced community resilience and adaptation to flood risks.

6. Literature Review

Extreme Value Theory is a branch of statistics that deals with extreme deviations from the median of probability distributions. The Extreme Value Theory originated in the 1920s, thanks to the pioneering contributions of Leonard Tippett with the help of Sir Ronald Fisher (Fisher & Tippet,1928). Their ideas on Extreme Value Theory have contributed to a growing demand for information from areas prone to extreme phenomena, compelling them to understand the underlying mechanisms responsible for the emergence of extreme events. (Gnedenko, 1943)

In another study, in the context of flood frequency analysis, Smith & Bate, (2020) conducted a study in the context of flood frequency analysis, where they identified homogeneous flood regions and selected an appropriate frequency distribution for these selected regions within a homogeneous region. It is necessary to use data from similar locations.

We can pool the historical data to obtain efficient estimates of the distribution's parameters, leading to robust quantile estimates with small standard errors.

Researchers have extensively studied parameter estimation procedures for the three parameters using the generalised extreme value distribution and the generalised Pareto distribution (Okeke et al., 2020). The most used methods are the method of L-moments, the method of moments and the maximum likelihood estimation (Hossain et al., 2013). This study employed the maximum likelihood estimation.

The return period sometimes called the recurrence interval, is a statistical tool for estimating how frequently a flood event of a particular size will occur. Its definition is the typical amount of time between flood events that are at least that big. A flood with a 100-year return time has an average 1-in-100 chance of being matched or surpassed in any particular year. The predicted size, such as the flood's water level or flow rate that matches a specific return period is known as the return level. For example, a water level of 10 feet could be the return level for a flood that occurs once every 100 years. This implies that the likelihood of a flood rising to or exceeding 10 feet in any given year is 1% or 1 in 100. The return level helps with risk management, flood defence planning, and infrastructure design and guides disaster preparedness (Hossain et al., 2013).
In another study, Onwuegduche et al. (2019) modelled and predicted extreme rainfall events in Kenya using the Extreme Value Theory for rainfall data from 1901 to 2016. Meanwhile, Marian O (2018) also predicted the return periods of maximum rainfall and floods in Ghana, West Africa, using the Gumbel Extreme Value theory. Kumar et al. (2018) applied the extreme value theory to predict floods in India using the block maxima approach, demonstrating its effectiveness in capturing extreme events. Furthermore, Machado et al. (2015) also used the Extreme Value Theory to model floods in Brazil, highlighting the importance of considering extreme values in flood risk assessment. Zhang et al. (2019) applied the Extreme Value Theory to predict floods in China, showing that the Block Maxima Approach outperforms other Extreme Value methods. In Zimbabwe, Mavhura et al. (2017) used statistical models to predict floods, highlighting the need for more robust approaches and Mugweni et al. (2019) applied machine learning techniques to predict floods in Zimbabwe, demonstrating the potential of advanced methods.

Generalised Extreme Value Distribution

This is a flexible 3-parameter model that combines the Gumbel, Fréchet and Weibull maximum extreme value distributions (Davison and Huser,2015). People often use the generalised extreme value distribution as an approximation to model the maxima of long (finite) sequences of random variables. It uses all the data to estimate its parameters. Therefore, all the values in the study will be considered when developing the model. It has the following distribution function:

$$G(x) = e^{-\left[1+\xi\left(\frac{x-\theta}{\beta}\right)\right]^{-\frac{1}{\xi}}}, 1 + \xi\left(\frac{x-\theta}{\beta}\right) > 0$$

where θ is the location parameter, β is the scale parameter and ξ is the shape parameter.

The density function of the generalised extreme value is given by:

$$g(x) = \left[1 + \xi\left(\frac{x-\theta}{\beta}\right)\right]^{-\frac{1}{\xi}} e^{-\left[1+\xi\left(\frac{x-\theta}{\beta}\right)\right]^{-\frac{1}{\xi}}}$$

The Gumbel distribution is obtained by setting $\xi=0$, The Weibull distribution is obtained by taking $\xi < 0$ and the Frechet distribution is obtained by taking $\xi > 0$. We often use the Gumbel distribution to model the maximum or minimum of many independent, identically distributed random variables. It is also known as the Extreme Value Type 1 distribution. (Gumbel,1958) It is unbounded and has the following probability density function:

$$f(x) = \frac{1}{\sigma} e^{-z} e^{-e^{-z}} \; where \; z = \frac{x-\mu}{\sigma}$$

,μ is the location parameter and σ is the scale parameter ($\sigma > 0$).

People commonly use the Fréchet distribution to model extreme events, where the tail of the distribution decreases in a polynomial manner. It is also known as the Type 2 distribution and has the following probability density function.

f(x) = $\frac{\alpha}{\beta}\left(\frac{\beta}{x-\gamma}\right)$α+1exp {-($\frac{\beta}{x-\gamma}$): $\alpha, \gamma \geq 0$, where α is the shape parameter $\alpha > 0$, β is the scale parameter ($\beta > 0$) and γ is the location parameter ($\gamma > 0$. It is suitable for modelling data where extreme values occur more frequently than predicted by a normal distribution and relate to maxima (largest extreme value).

Weibull Distribution is versatile and can be applied to model a wide range of distributions, including those with increasing, decreasing, or constant failure rates. It is also known as the Extreme Value Type 3 distribution and has a bounded upper tail. (Weibull,1951) The Weibull distribution has the following probability density function.

f(x) = $\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}$exp {-$\left(\frac{x}{\beta}\right)^{\alpha}$, for $x > 0$ where α is the shape parameter ($\alpha > 0$) and β is the scale parameter, ($\beta > 0$). The Weibull distribution is flexible therefore it is widely used in many other areas. The Weibull model relates to minima (smallest extreme value).

7. Methodology

The Extreme Value Theory methodology offers a systematic approach to predicting floods by focusing on extreme events and their probabilities. The researchers used the Block Maxima of the Extreme Value Theory in hydrology to study how likely it was that Region 3 in Zimbabwe would experience an unusual flooding event.

7.1 Area of Study

Region 3 in Zimbabwe typically experiences an unimodal rainfall pattern, with the rainy season usually occurring from November to March. The average annual rainfall in this region ranges from 600mm to 800mm. The temperatures in Region 3 of Zimbabwe are generally warm-to-hot throughout the year. During the dry season, winds are calm while during the rainy season, occasional gusts of wind accompany thunderstorms. Region 3 is susceptible to droughts, especially during periods of below-average rainfall. The region consists of undulating plains and some hilly areas. Region 3 has various catchment areas (Moyo S,2000, Vincent & Thomas,1961). The researchers chose Region 3 because it is prone to floods. The 2013 flood disasters claimed the lives of 159 people, caused damage worth US$53.2 billion, and claimed 33% of victims. This was the worst year affected by floods in Zimbabwe. The natural areas in region 3 in Zimbabwe with the amount of rainfall in millimetres received are shown in the map below:

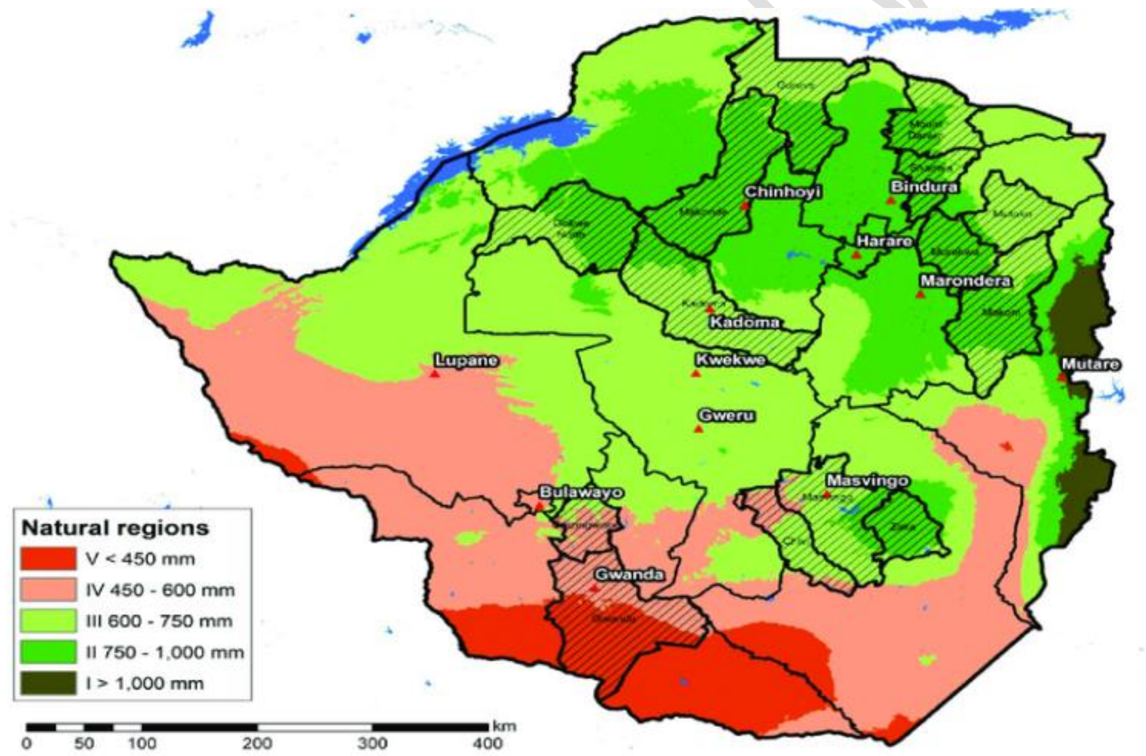

Figure 1: Region 3 area in Zimbabwe

Source: Adapted from Moyo, 2000; Vincent and Thomas, 1961.

7.2 Data used in research.

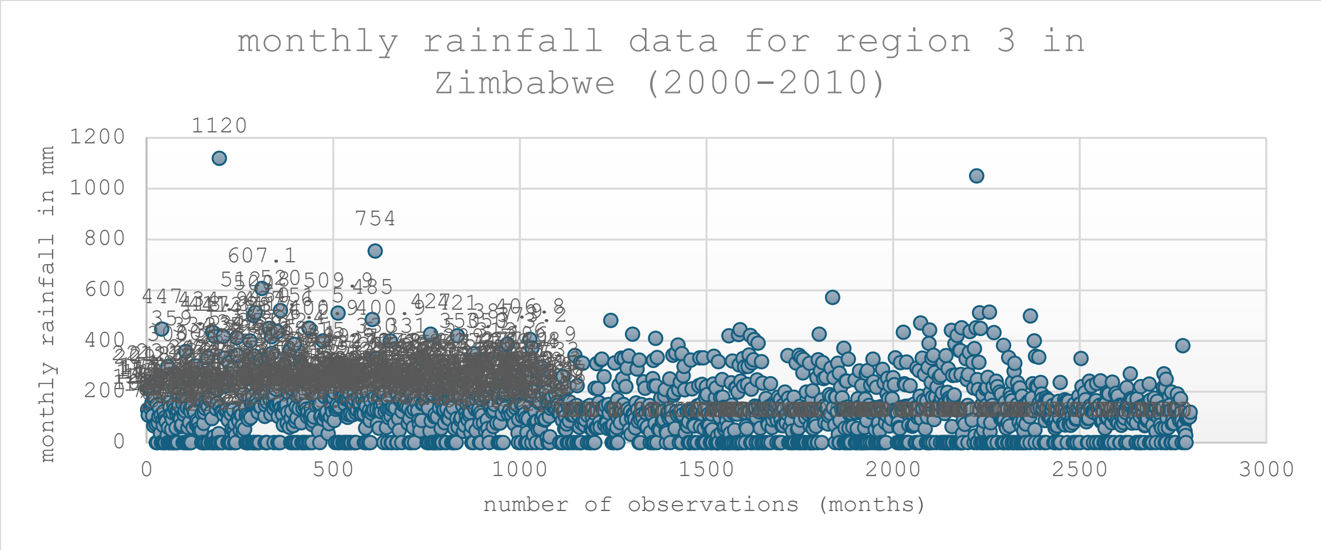

Figure 2: Monthly maximum rainfall in Region 3 in Zimbabwe from the period (October to March) 2000 to 2010.

The researchers sampled ten seasons from ten years of the twenty stations under study. Figure 2 above shows that the maximum rainfall received was 1120 mm, and the minimum received was 15.7mm.

7.3 Construction of the Extreme Value Theory Model

Generally, there are two practical approaches to extreme value theory which, are the block maxima and the peak-over-threshold approach (Okeke & Ehiorobo, 2020). The researchers used the Block Maxima Approach. The appropriate block maxima values are selected. We estimated the parameters using the graphical method and fitted them to a generalised extreme value model. We conducted a likelihood ratio test to assess the model's appropriateness.

7.3.1 The Statistical Software Used in Analysing Data

We performed the result analysis using the R statistical package, version 4.3.2, and SPSS, version 23.0.

7.3.1.1 R Method

The R method is an integrated suite of software facilities for data manipulation, calculation, and graphical display (R Core Team, 2023). We stored rainfall data in Microsoft Excel and exported it to R for data analysis. We used R to plot and display the diagnostic plots for model evaluation, which included the return level plot, boxplot, density and quantile plots, profile likelihood plots, Weibull, and normal Q-Q plots.

7.3.1.2 SPSS

SPSS is a software program developed to manage and analyse statistical data. It enables the researchers to build, store data and validate predictive models using advanced statistical procedures (R Core Team, 2023). In this study, we stored the rainfall data in SPSS and exported it to R for analysis. We also conducted a likelihood ratio test to validate the Extreme Value Theory model.

7.4 THE BLOCK MAXIMA APPROACH

The block maxima approach is a statistical method used in extreme value theory to model the behaviour of the maximum values of a dataset within fixed intervals or "blocks" of time or space. This approach is commonly applied to analyse and predict extreme events, such as natural disasters, financial crashes, or rare occurrences (Katz et al., 2023).

7.4.1 Data collection and partitioning

The researchers used the seasonal phase's maxima for the areas under study since no rainfall was received off-season. Since the season phases spanned six months, we divided the rainfall data into non-overlapping blocks of six months (October to March). We partitioned the six months further into blocks of three months.

7.4.2 Extraction of Maximum

The maximum value was identified and extracted from a block of three months. These maxima were then used to form a new dataset consisting only of these extreme values.

7.4.3 Fitting an Extreme Value Distribution

The extracted block maxima were fitted to a generalised extreme value distribution, which is a good way to show how block maxima are spread out. The Likelihood Ratio Test for the validation of the model was carried out using the SPSS statistical package

7.4.4 Parameter Estimation

The maximum likelihood estimation in SPSS was used to guess where, how big, and what shape the generalised extreme value distribution would be.

7.4.5 Model Validation and Prediction

We formulated the null hypothesis, asserting that the parameters did not affect the model. We computed the likelihood ratio test value and the chi-square value in SPSS. The diagnostic plots were used to check how well the generalised extreme value distribution fit the observed extreme values. The plots included the Boxplot, Weibull Q-Q plot, Normal Q-Q plot, return period and return level plots.

8. **Results and findings**

The null hypothesis was that all parameters did not affect the model. The Likelihood Ratio Test showed a -2 log (likelihood) value of 447.278, therefore, the alternative model (the flood prediction model) is a big step up from the null model. This indicates strong evidence that the flood prediction model is appropriate for the data.

Table 1: The model fitting results: Likelihood Ratio Tests for Block Maxima

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	447.278	30.108	2.00	1.000

From the results in Table 1 above, a Chi-Square test with 2 degrees of freedom and a significance level of α=0.01 was found. The critical value for $\chi_2^2(0.01) = 9.21$.The calculated Chi-Square test statistic is 30.108. The value is greater than the critical value of 9.21. Since the test statistic, 30.108 is much greater than the critical value of 9.21. We reject the null hypothesis and conclude that there is strong evidence to suggest that the more complex model (alternative hypothesis) provides a significantly better fit to the data than the simpler model (null hypothesis. Hence, we **c**onsider adopting a complex model for future analyses, as it better captures the underlying data characteristics, such as extreme flood levels.

8.1 The Boxplot

The box plot is excellent for conveying location and variation information in data sets.

The rainfall data was plotted and produced the following box plot.

Box-plot for Block Maxima

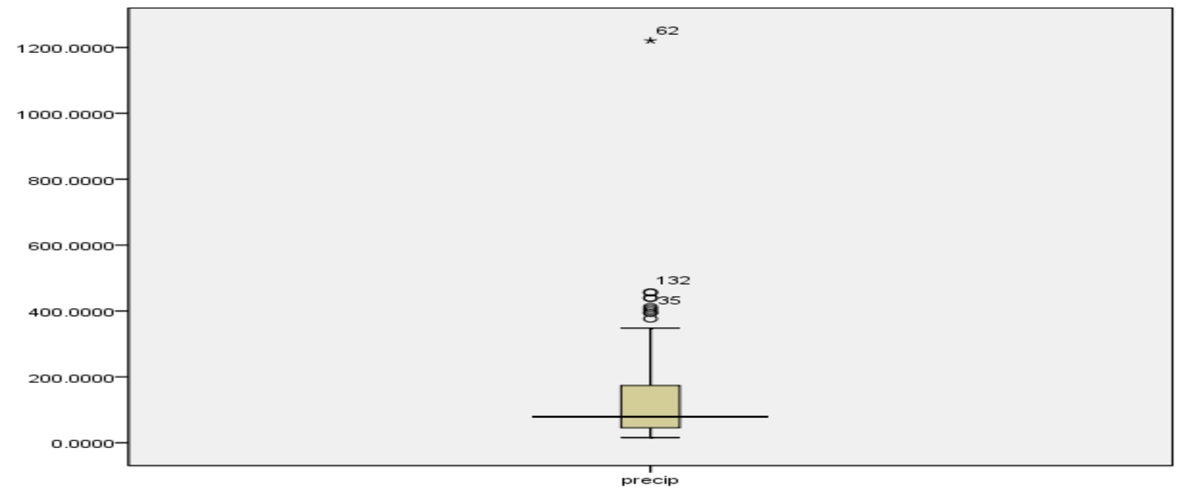

Figure 3: The box plot for the Block Maxima

The box plot for the Block Maxima method showed that the box is tall above the median of 79.1mm. This means that rainfall above this median was received in the areas under study. The graph shows three outliers which represent the heaviest rainfall of 1120mm, 776.5mm, and 457.1mm for three months.

8.2 The Weibull Q-Q plot

The Weibull Q-Q plot is a graphical technique for determining if a data set comes from a 2-parameter Weibull distribution. (the location parameter is assumed to be zero). The rainfall data plotted in this research produced the Weibull Q-Q plot below, which assisted the researchers in assessing the reliability of the results.

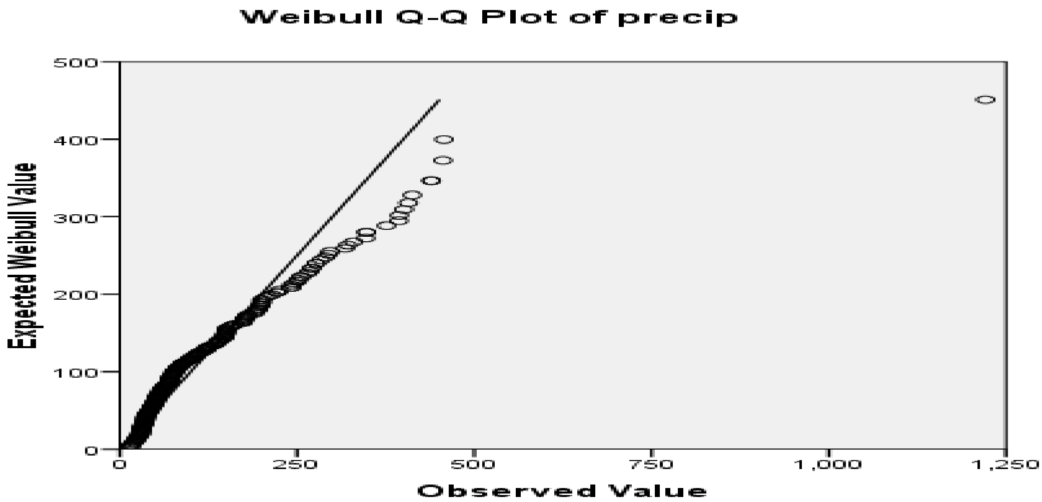

Figure 4: The Weibull Q-Q plot of precipitation

The Weibull QQ plots of the Block Maxima method showed the standard Weibull data (vertical axis) against the normal population (horizontal axis). The plot showed that most of the points lie on the unit diagonal line. Therefore, we conclude that the generated model provided a plausible fit to the observations.

8.3 The Normal-Q-Q plot

It is also called the Z-plot. The Z-plot below was produced.

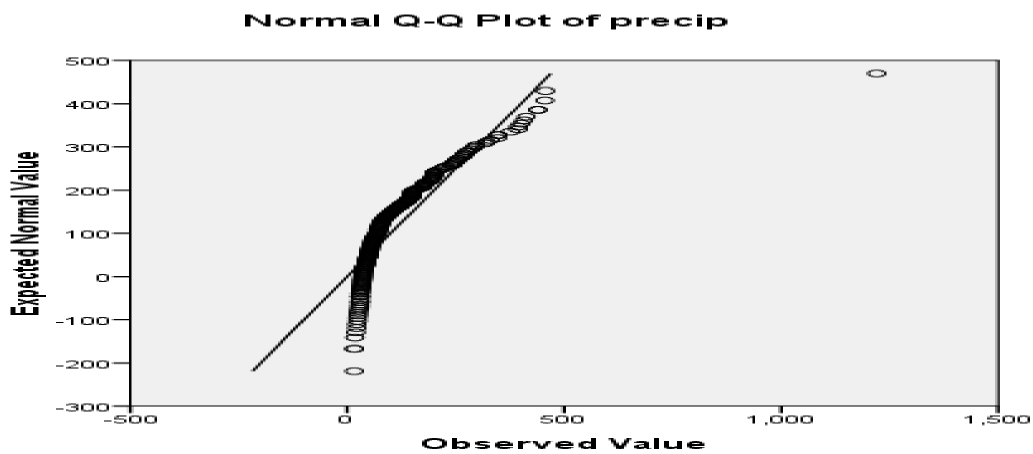

Figure 5: A Normal Q-Q plot of the precipitation

The normal QQ plot above showed the standard normal data (vertical axis) against the normal population (horizontal axis). From Figure 5 above, the expected normal distribution is a straight line, and the line of little circles represents the observed values from the data. The plot showed that the distribution deviated from normality at the lower end. The high end of the distribution was pretty much normal as the observations are closely around the straight line. The concave deviation from the straight line on the graph showed that the data fitted were heavy-tailed distributions, meaning that the data distributions were a family of the Extreme Value Theory distributions.

8.4 Return Level Plot

The return level plot showed the expected return levels for each return period. The return level plot represents values expected to be exceeded on average once every n-year. Figure 6 illustrates the return level plot for the Block Maxima method, plotting the return period against the return level, along with an estimated 95% confidence interval line. The vertical axis in Figure 6 below displays 1000, 1500, 2000, and 2500 cubic feet per second, equivalent to 28.3 m^3, 42.5 m^3, 56.6 m^3, and 70.8 m^3, respectively, while the horizontal axis displays the return period in years, 1 year, 100 years, 200 years, and 300 years. Figure 6 below predicts an average rainfall of approximately 2400 cubic feet per second (68 m^3) every 100 years. The return level was approximately 1490 cubic feet per second (42.2 m^3) for a 1-year return period and approximately 1900 cubic feet per second (53.8 m^3) for 10 years. The graph demonstrated an increase in the maximum rainfall amount as the return period lengthened, indicating an increased likelihood of floods in the future. For example, the 2014 Tokwe-Mukorsi Dam flood received 850mm of rainfall, double its annual average (Mavhura et al., 2017). Figure 6 shows the probability, quantile, and density plots were linear. This meant that the model assumptions were correct for the data fit. All the points were within the two lines of the confidence limits; hence the data fit was valid.

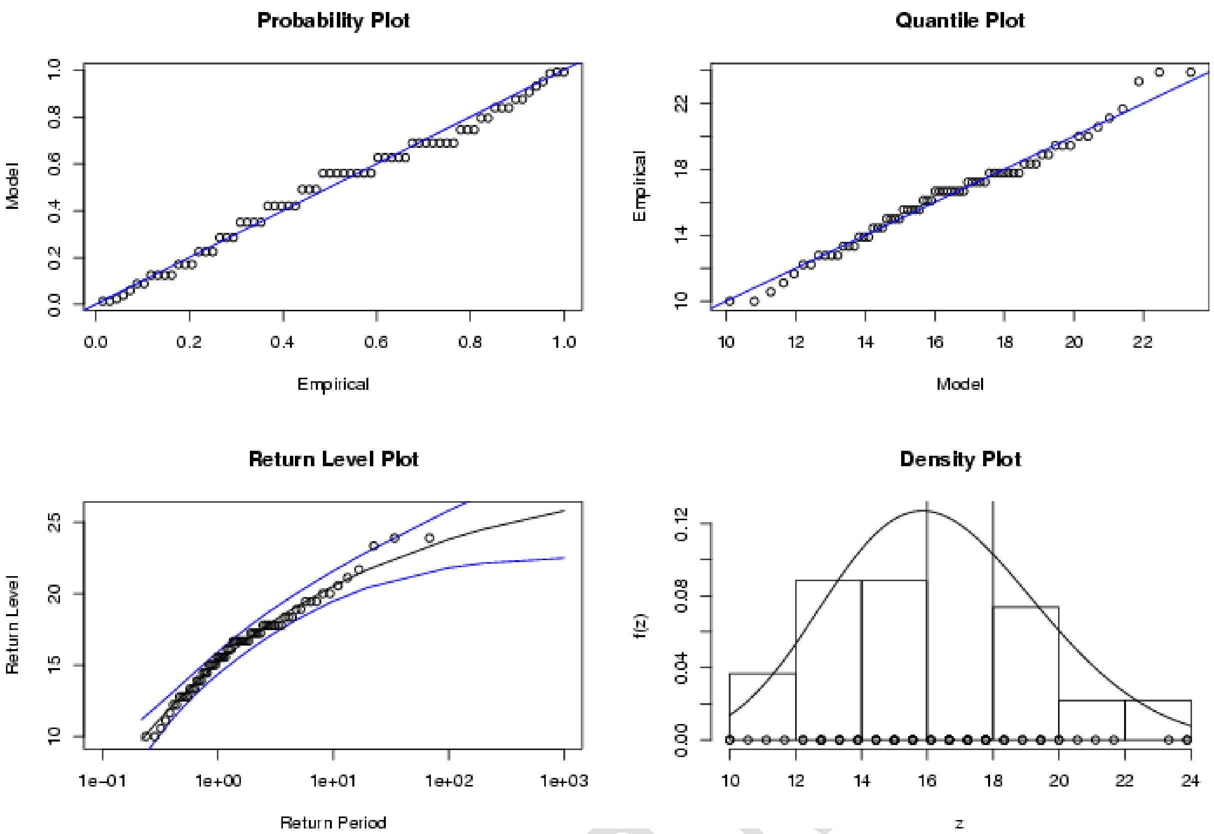

Figure 6: The Probability, Quantile, Return level and Density Plots

8.5 Return Periods

In hydrology, the size of a flood is often placed into perspective by referring to its expected recurrence period. A return period is also known as the recurrence interval or repeat interval. It is the inverse of the expected number of occurrences in a year. We determined the return period of the floods using $T = \frac{1}{p}$ for a 2,5 10 and 20-year period. Therefore, a 2-year return period has 50%, a 5-year return period has 20%, a 10-year return period has 10% and a 20-year return period has a 5% chance of being exceeded in any year.

8.6 Return levels.

The fitted Fréchet model was used to estimate the return levels for various return periods. For example, the 100-year return level (i.e., the flood level expected to be exceeded once in 100 years) was estimated to be significantly higher than the historical maximum observed flood level, underscoring the potential for extreme flood events beyond recorded history. These return level estimates are crucial for designing flood defences, setting insurance premiums, and forming emergency response plans.

9. The Flood Model

After estimating the parameters, the recommended model for the return level was expressed as follows:

$Z_p = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}((-\log(1-p))^{-\hat{\xi}} - 1)$ Where $\hat{\mu}$ is the estimated location parameter.

$\hat{\sigma} > 0$ Is the estimated scale parameter.
$\hat{\xi}$ >0 Is the estimated shape parameter.

p = $\frac{1}{Return\ Period}$, the model is Z_p,

9.1 The Flood Model Results

We fitted the data to a generalized extreme value distribution using the Block Maxima method and the SPSS statistical package.

The table below displays the estimated parameters.

Table 2: Model Fitting results for the Block Maxima

#	Distribution	Parameter
1	Fréchet	α =1,4815 β = 59,557
2	Fréchet 3P	α = 1,690 β = 76,588 γ = 14,512
3	Generalised Extreme Value	**ξ =0,3009 σ = 56,436 $\mu = 69.159$**
4	Gumbel max	σ = 96,034 μ = 69,897
5	Gumbel min	σ = 96,034 μ = 180,76
6	Weibull	α = 1,4602 β = 129,02
7	Weibull 3P	γ = 15 α = 0,98949 β = 110,19

According to Table 2 above, the estimated parameters were μ (location parameter) = 69.159, σ (scale parameter) = 56.436, and ξ (shape parameter) = 0.3-009. The location parameter is the central value that governs the distribution of the extreme daily rainfall quantities. Here, the result of 69.159 indicates that this is around the normal extreme daily rainfall quantity that could cause floods. On a scale of 56.436, the severe rainfall levels show a large variation. This implies that real rainfall could vary significantly, even though the average extreme rainfall might be about 69.159 mm. The distribution is Fréchet, which means it has a heavy tail, according to the shape parameter ξ = 0. 3009. This indicates that there is a greater chance of having extreme rainfall quantities much above 69.159 mm. Therefore, even though the central tendency is approximately 69.159 mm, there is a considerable likelihood of intense rainfall events with significantly larger amounts that could cause serious flooding.

Letting p = $\frac{1}{Return\ Period} = \frac{!}{T}$,

Since $\hat{\xi}$ >0, the distribution is of a Fréchet type, the recommended model for the return level using the Fréchet 3P parameters as per the researchers' results (α (shape parameter) = 1,690, β (scale parameter) = 76,588 and γ $(location\ parameter)$ = 14,512) was:

Z_p = 14,512 + $(\frac{76{,}588}{1{,}69})((-\log(1-p))^{-1{,}69} - 1)$.

9.1.1. Calculating the return period and the return level using the estimated model.

P=0.3

Return period = $\frac{1}{p} = \frac{1}{0.3}$ = 3.3years

Return Level = Z_p = 14,512 +$(\frac{76,588}{1,69})$ $((-\log(1-0.3))^{-1,69}-1)$ = 1028,6359 cubic feet per second.

9.1.2. Probability Estimation

The probability estimation formula is $P_e = 1 - \left[1-\left(\frac{1}{T}\right)\right]^n$ where T is the return period of a given storm threshold and n is the number of years of the return period.

The Probability estimation for Block Maxima $=1 - \left[1-\left(\frac{1}{3}\right)\right]^3$ = 0.7037=70.4%.

According to the above calculations, the Block Maxima analysis indicates a 70.4% probability of a return level of 1028.6359 cubic feet per second of rainfall in approximately 3 years. Hence this suggests a high chance of receiving rainfall with a high flow in 3 years. In 2013 and 2014 heavy rainfall in some areas, like Murehwa, which received 77mm, Mucheke received 66 mm, and Masvingo received 50mm was received in region 3, Zimbabwe.

10. Conclusion

The Block Maxima Approach can be used to determine the likelihood of floods. The uncertainties in the parameter estimates should be considered when the probability associated with a flood event is assessed. The results showed that the Fréchet distribution from the Generalized Extreme Value family is suitable for predicting the highest rainfall in Zimbabwe's Region 3. Modelling extreme flood events using the Block Maxima approach has provided a more realistic and nuanced understanding of flood risk in the study area. The heavy-tailed nature of the Fréchet distribution appropriately reflects the probability of large flood events, which are often underestimated by traditional modelling techniques. This research demonstrates that relying on the Fréchet distribution can significantly enhance flood prediction accuracy and improve risk management strategies, particularly considering growing concerns over climate change and increased variability in weather patterns. The results showed that after 3 years, the return level was 1028.6359 cubic feet per second. We could employ the presented model to calculate the return levels or duration periods for Region 3. The average extreme rainfall event is approximately 69.159 mm, with the heavy-tailed nature of the distribution, there is a real possibility of much larger events, underscoring the need for robust flood management and preparedness plans.

The fitted Fréchet model might indicate a heavy-tailed distribution, suggesting a non-negligible probability of experiencing discharge levels much higher than the historical maximum. The Extreme Value Theory provides efficient estimators for extreme value parameters, enhancing better estimation accuracy even with limited data on extreme events. However, there exist limitations in this study, which include loss of valuable data and potentially important patterns or trends present in the dataset since only maximum values were selected within each block and the other information in the block was discarded. In the future, researchers should focus on expanding this study to include more flood-prone areas, adding climate change scenarios to the modelling framework, and creating a complete method for extreme value modelling that uses more than one data source and method.

In the future, researchers should focus on expanding this study to include more flood-prone areas, adding climate change scenarios to the modelling framework, and creating a more complete method for extreme value modelling that uses more than one data source and method.

11. Way Forward

In Zimbabwe, flood forecasting is essential for preventing fatalities, safeguarding property and infrastructure, reducing financial losses, maintaining the environment, encouraging sustainable development, and enabling efficient humanitarian responses during floods. The researchers recommend the following course of action:

- **Early Warning and Preparedness**: Early warning systems must be implemented to notify other relevant parties and the local population of impending floods. Early warning systems save lives and lessen casualties during flood catastrophes by enabling the prompt evacuation of populations that are at risk.
- **Risk Assessment and Management**: Authorities should oversee risk assessments to determine locations susceptible to floods. Developing emergency response plans, installing flood protection measures, and constructing resilient infrastructure are just a few of the effective flood risk management techniques that can be used.
- **Reducing Economic Losses**: The government ought to reduce economic losses. Due to damage to property, crops, infrastructure, and business interruptions, flooding can cause large economic losses. Flood predictions aid in estimating possible economic effects, enabling improved insurance coverage, financial planning, and resource allocation for recovery activities following a flood.
- **Environmental Protection**: The authorities ought to safeguard the environment. Floods can harm the ecosystem by eroding soil, contaminating water, destroying habitats, and reducing biodiversity. Flood prediction facilitates the evaluation of environmental hazards and the execution of preservation strategies to minimise ecological harm.
- **Sustainable Development**: Legislators and urban planners ought to implement plans for sustainable development. Urban planners and legislators may make well-informed choices about infrastructure development, land use planning, and catastrophe risk reduction programs by accurately forecasting floods. This supports sustainable development strategies that give resilience to calamities like floods.
- **Humanitarian Aid**: Funding for humanitarian aid should come from the government and humanitarian organizations. Government agencies can efficiently coordinate emergency response activities, preposition relief supplies, and deploy rescue teams when they have access to timely flood forecasts. This guarantees that aid reaches impacted communities quickly, lessening suffering and effectively meeting humanitarian needs.

Funding
This research was conducted without external funding. The researchers covered all study costs as part of a self-funded initiative.

Informed Consent Statement

The study did not require the involvement of human participants or the collection of personal data, so obtaining informed consent was unnecessary. All the data used in the research were sourced from the meteorological and hydrological records from Meteorological Services of Zimbabwe, to ensure adherence to ethical guidelines and data usage policies.

Data Availability Statement

The data used in this study are from the Meteorological Services Department of Zimbabwe.

Conflicts of Interest
The authors declare no conflict of interest.

12. References

1. Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., … Feyen, L. (2018). Global projections of river flood risk in a warmer world. *Earth's Future,* 6(5), 583-600.
2. Davison, A. C., & Huser, R. (2015). Statistics of extremes. *Annual Review of Statistics and Its Application*, 2, 203-235.
3. Fisher, R.A., Tippett, L.H.C. (1928),*' Limiting forms of the frequency distribution of the largest and smallest member of a sample.* 'Proc. Cambridge Philosophical Society 24:180-190.
4. Gnedenko, B. (1943) *Sur la distribution limite du terme maximum d`une serie aleatoire. Annals of mathematics*,423-453.https://doi.org/10.2307/1968974.
5. Gumbel, E.J. (1958).*'Statistics of Extremes'*. Columbia University Press. New York.
6. Hossain, M. A., Reinsel, G. C., & Davison, A. C. (2013). Maximum Likelihood Estimation for Generalized Extreme Value Distributions. *Technometrics*, 55(4), 513-523.IBM Corp. (2023). *IBM SPSS Statistics for Windows, Version 29.0*. Armonk, NY: IBM Corp.
7. Katz, R. W., Parlange, M. B., & Naveau, P. (2023). "Statistics of extremes in hydrology". *Water Resources Research*, 59(4), e2022WR031914. doi:10.1029/2022WR031914.
8. Kumar, R., Jain, S.K., &Kumar, V. (2018) Application of Extreme Value for Flood Frequency Analysis: A Case Study of the Mahanadi River Basin, India: *Journal of Hydrology*.
9. Machado, M.J., Machado, M.A., & Franches. (2015) Extreme Value Theory Applied to Flood Risk Analysis: A Case Study of the Sao Francisco River Basin, Brazil (2015).
10. Marian O. (2018) Estimation of the return periods of maxima rainfall and floods at the Pra River Catchment, Ghana, West Africa, using the Gumbel Extreme Value Theory, *Journal of Hydrology.*
11. Mavhura, E., Mwamuka, L., & Mabaya, M. (2017). Application of Extreme Value Theory for Flood Frequency Analysis in the Save River Catchment, Zimbabwe. Journal *of Water Resource and Protection.*
12. Mosavi A., Ozturk P., Chau K. (2018). Flood Prediction Using Machine Learning Models: A literature review. *Water* 10(11),1536,2018.
13. Moyo S (2000). Zimbabwe environmental dilemma: balancing resource inequities. Harare, Zimbabwe environmental research organization, P. 161.
14. Mugweni, T., Mavhura, E., & Hoko, Z. (2019). Flood Risk Assessment Using Extreme Value Theory and Machine Learning Techniques: A Case Study of the Mazowe River Catchment, Zimbabwe: *Journal of Flood Risk Management*.
15. Okeke O., Ehiorobo J., (2020). Packages "extremes". *Extremes*.18:1
16. Onwuegduche, C., F., Affognon, S., B., Kenyatta, A., and Enock Akinade, M.O., E. (2019) Application of Extreme Value Theory in Predicting Climate Change-Induced Extreme Rainfall in Kenya. *International journal of statistics and probability.*
17. R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

18. Smith, M. J., & Bates, P. D. (2020**).** Regional flood frequency analysis: Defining homogeneous regions based on geomorphological catchment properties. *Hydrology and Earth System Sciences*, 24, 5551-5567.
19. Vincent V, Thomas RG (1961). An agricultural survey of Southern
Rhodesia. Part 1 - Agro-ecological survey. Government Printers,
Salisbury.
20. Weibull, W. (1951).*'A statistical distribution function of wide applicability*. Appl.Mech., Trans.ASME 18(3), 293-297.
21. Zhang, Q., Li, J., & Sigh, V.P. (2019) A Comparative Study of Extreme Value Theory and Hydrological Modeling for Flood Risk Assessment: A Case Study of the Yangtze River Basin, China. Journal *of Hydrology*.