PEAK-OVER-THRESHOLD ANALYSIS OF EXTREME RAINFALL IN THE NORTH WESTERN REGION OF RWANDA

Musayidizi Jean de Dieu

College of sciences and technology

School of Science

Master of Science in Atmospheric and Climate Science

2022

PEAK-OVER-THRESHOLD ANALYSIS OF EXTREME RAINFALL IN THE NORTH WESTERN REGION OF RWANDA

By

Musayidizi Jean De Dieu

Registration Number :219014456

A dissertation submitted in partial fulfillment of the requirements for the degree

Of

MASTER OF SCIENCE IN ATMOSPHERIC AND CLIMATE SCIENCE

at the College of Sciences and Technology

Supervisor: **Prof. Bonfils Safari**

March,2022

DECLARATION

I declare that this Dissertation contains my own work except where specifically acknowledged

Musayidizi Jean de Dieu

219014456

Signed:………………………………………..

Date:……………………………………………..

DEDICATION

This work is dedicated to my lovely wife Rugambwa Olga, sweet daughter Akariza Ola Gwyneth, my precious Mother, brother and sisters.

ACKNOWLEDGEMENT

To begin with and First, I would like to thank God for giving me life, protection, capacity and ability to carry out the research and my study.

This research would not have been possible without the assistance of my supervisor Prof Bonfils Safari whose strong help, guidance and constructive comments lead to the accomplishment of this thesis topic.

From my heart, I thank my family, mother, brother, sisters and friends for moral support, their words of encouragement kept a smile on my face and so gave me the courage and energy to work hard

I thank the entire university academic staff for the direction and awesome bolster given to me amid this period. My gratitude goes to the Rwanda government through higher education council (HEC) that fully supported my studies till the end.

I cannot disregard to thank my classmates whom we have been motivated to work and handle together different duties within the lesson until we are granted for completion of the program.

ABSTRACT

Extreme rainfall events over northwestern (N-W) of Rwanda causes floods and landslides which have a major impact on society and lead to loss of property and life. The main objective of this study is to study extreme rainfall in Northwestern of Rwanda using peak over threshold method. This research modelled and predicted extreme rainfall events in N-W using extreme value theory, where daily rainfall data from Rwanda Meteorological Agency for three stations Gisenyi, Ruhengeri and Rubengera from the years 1981 to 2018 representing the area of study were utilized.

Peak over threshold method (POT) was used to fit the generalized Pareto distribution, both the mean residual life plot and the parameter stability plot were chosen as the choice of threshold used in this study. Rainfall above threshold were selected with different threshold values which represents the 99 percentiles for each station. After threshold selection, the parameters such as scale σ and shape ξ were estimated using maximum likelihood estimator and these parameters are obtained numerically. To verify POT, we have made some diagnostic test such as quantile-quantile plot (Q-Q plots) and profile likelihood for all stations, almost all point lies to the line that means that our model is adequate.

For all stations confidence interval of 95% include the zero for the shape parameter that is, the data fits the exponential distribution except Gisenyi station which indicated negative for shape parameter as estimated. Furthermore, prediction for the return periods of 20, 50, 80, 100, 150 and 200 years for all stations were made using return level estimates and their corresponding confidence interval were presented. It was found that an increase in return periods lead to an increase in return levels and this indicates that the frequency of occurrence is increasing.

KEY WORDS

Extreme rainfall, peak over threshold, generalized Pareto distribution, extreme value theory, return level.

LIST OF SYMBOLS AND ACRONYMS

AMS: Annual Maximum Series

CI: Confidence Interval

ENSO: El Niño Southern Oscillation

EVT: Extreme Value Theory

GDP: Gross Domestic product

GEV: Generalized Extreme Value

GML: Generalized Maximum Likelihood

GoF: Goodness-of-Fit

GPD: Generalized Pareto Distribution

IOD: Indian Ocean Dipole

ITCZ: Inter-Tropical Converge Zone

MAM: March to May

MIDMAR: Ministry of Disaster Management of Rwanda

MIR: Minimum Ratio of Residual Correlation Coefficient

MLE: Maximum Likelihood Estimator

MRL: Mean Residual Life Plot

N-W: North Western

PDS: Partial Duration Series

POT: Peak Over Threshold

REMA: Rwanda Environment Management Authority

SOND: September to December

Table of Contents

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Background.

Extreme rainfall has been hazardous issue in Rwanda, for the reason that it causes major damage to infrastructure, agriculture, disturbance of human activities, injuries and loss of lives.
Most parts of Northwestern (N-W) of Rwanda numerous parts are affected by devastating rainfall associated floods and landslides. Identification of floods and landslides areas can lead to better understanding of disaster risk and putting in place measures for risk reduction (J.ean Baptiste & Habiyaremye, 2014).

For examples, the records of the recent events have shown that on the 21st November 2000, Nyundo and Karambo in Rubavu district recorded over 200 houses damaged. In the September 2005, 35 deaths and 379 houses destruction were registered in the North-Western regions. In the years 2001, 2002, 2007, 2008 and 2012, devastating floods and landslides affected by heavy rainfall has been observed over most parts of Northern and Western region. 74 deaths, 22 injuries, 573 houses destroyed or damaged, and 656ha of affected land were recorded due to landslide from 2011 to 2013, (MIDIMAR, 2015; REMA, 2013).

Within the April 2012, heavy rains hit most parts of N-W region and caused flood that affected around 11,160 persons including deaths and injuries. Subsequently, property and houses with schools and churches have been extensively damaged (Red Cross, 2013).
The most affected region is Western region with more than half of the total deaths records (51%), followed by the Northern province records (38%). Rulindo, Burera, Karongi and Nyabihu districts experienced more deaths than other districts (REMA, 2013).

1.2 Statement of the Problem and justification.

Extreme rainfall events have resulted in massive destruction both on the society, environment and the economy. Rwanda is more vulnerable to extreme rainfall leading to numerous natural hazard such as flood and landslides. Several districts of the N-W regions of Rwanda such as Rubavu, Nyabihu, Karongi, Rutsiro, and Musanze, have been hit by extreme rainfall (MIDIMAR, 2015). A few of the casualties to be drowned in flood water, and others died after houses collapsed under the heavy rain. This can mostly affect economic growth of this region and also the country due to the landslides and floods which damage crops, houses and other infrastructures such as bridges, roads and schools and as well as loss of human and animal lives.

Extreme rainfall leads to flood and landslides, which can undermine human life and disturb transport. Northwestern area of Rwanda is one of the area mostly affected by extreme rainfall. However, extreme rainfall threshold was not documented over the region which could inform future forecasting in study area. For improved management of mitigation to effect of extreme rainfall in the northwestern area of Rwanda, identification of appropriate method for determination of extreme rainfall threshold and deriving extreme rainfall threshold adopted to the study area can lead to better forecasting of extreme and improve the impact based early warning system which will contribute for disaster risk reduction and putting in place emergency measures for risk reduction.

Here some figures showing floods and landslides risks.

Flood at Ecole d'Arts de Nyundo school, source: https://www.newtimes.co.rw, February 11,2021

Karongi landslide. source: http://ktpress.rw,February 12, 2021

Flood effect over Mukungwa river. source: https://www.newtimes.co.rw/ March 06,2017

Landslide over Karongi district. Source: https://climatechangenews.com/2019/04/09/

1.3 Objectives of the Study.

1.3.1 General objective:

The general objective of this study is to study extreme rainfall in Northwestern of Rwanda using peak over threshold method.

1.3.2 Specific objectives.

The specific objectives of this study are:

- ✖ To determine Peak over threshold of extreme rainfall.
- ✖ Understanding frequency and intensity of extreme rainfall over Northwestern
- ✖ Determine the return period corresponding to the return levels.

1.4 Area of Study

The area of interest for this study is North-western of Rwanda and is a landlocked nation situated in the southern hemisphere, it is bordered by Uganda to the North, Burundi to the South, Tanzania to the East and Democratic Republic of Congo (DRC) to the West. Rwanda has an area of 26,338 km^2 and lies between 1°4´and 2°51´south, and 28°53´and 30°53´ east (Muhire et al., 2015). however, N-W region is bolded Uganda and RDC, and has various topographic features such as Mountains, Lakes, Forests and Volcanoes. It is also known as the area where farming and livestock are dominant due to high fertility of the soils. Many types of crops are grown such as sweet potatoes, banana, sorghum, soya, coffee, beans, tomato, pepper and some vegetables. Due to beauty of the region, is also blessed with various tourist’s attraction which include lake Kivu, volcanoes and hot springs.

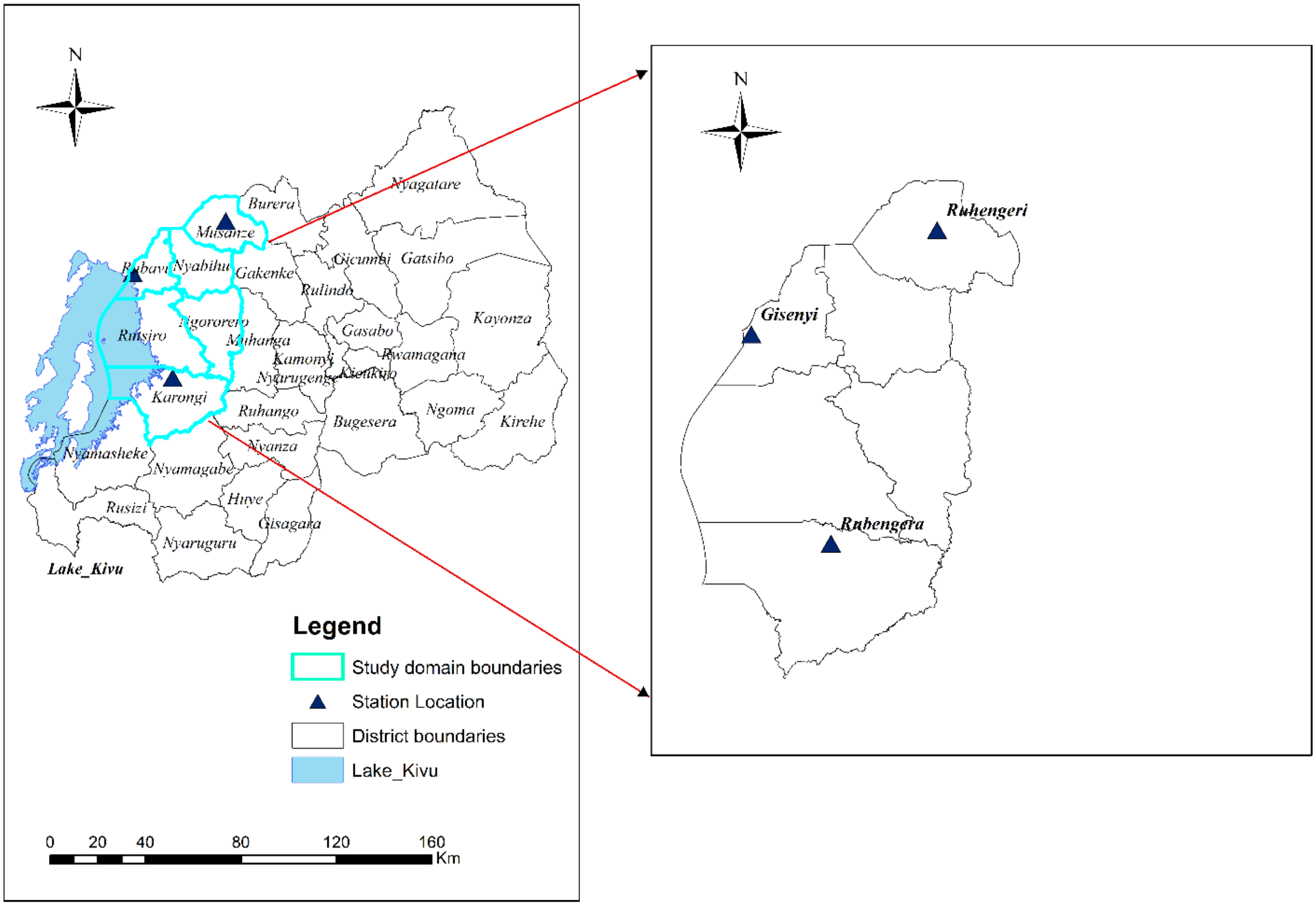


Figure 1:map of Rwanda (left)including districts and study domain (right) showing stations used

1.5 Climate characterization of study area

Rwanda is famous as "le Pays de mille collines" (Land of a thousand hills), its mountainous, with ridges and hills separated by a complex dense drainage network of rivers, lakes, and wetlands (USAID, 2019). (ITCZ) Intertropical convergence zone is defined as the region where the trade winds from northeast and southeast converge. Over East Africa, ITCZ is comprised of both the meridional arm formed by the convergence of easterly winds from the Indian ocean and most westerlies from the Atlantic Ocean and the Congo basin. And zonal arms formed by the convergence of south-easterly and north-easterly trade winds and these are responsible for both the long and short rainfall seasons when ITCZ is close to the equator (Okoola, 1999). Area near to the Equator tend to have two rainy seasons in MAM (long rains) and SOND (the short rains)(Okoola, 1999).this region has a significant rainfall in most months of the year but annual maximum rainfall occurs twice a year in April and November.

The flow of winds over our region are influenced by four major anticyclones which are : St. Helena high pressure system to the south west of Atlantic Ocean, the Mascarene high pressure system to the south east Indian ocean, The Arabian high in the middle east and Azores high pressure system; and they control the position and the movement of ITCZ as well as Congo air mass regime over the Central and south-east of Africa (Ngarukiyimana et al., 2018).

CHAPTER TWO:

LITERATURE REVIEW

2.1 Literature Review.

Extreme value theory has intense applications in several areas which are: flooding, rainfall, storms, precipitation, insurance claims, and price fluctuations (Friederichs, 2007). Several researchers around the world have applied EVT in different countries to come up with several estimates about extreme events.

The study of Lazoglou and Anagnostopoulou (2017) applied EVT in the Mediterranean region on ten stations data. Block Maxima and Peak Over Threshold (POT) with Maximum Likelihood, L-Moments, and Bayesian methods were employed in calculating the parameters of the extreme distributions. Results from the study indicated that GPD gives a good theoretical explanation compared to Generalized Extreme Value (GEV) distribution in the prediction of extreme precipitation. Bayesian method gave the most accurate estimate parameters for the highest precipitation levels in most of the stations. GEV distribution with Bayesian estimator was found to give the highest return levels for the western stations while GPD with Bayesian estimator gave the highest return levels for the eastern regions for 50, 150 and 300 return periods.

The peak over threshold method was utilized as an alternative technique to the traditional analysis of annual discharge maxima of the Danube river. All mean daily discharges exceeding a defined threshold were considered within the POT analysis. Therefore, with respect to later regulations along the Danube channel bank the 40, 20, and 10-year period and the second analysed time data series were selected from the end of the 60-year period. The results suggest that the POT method can provide adequate and comparable estimates of N-year discharges for more stations with short temporal coverage (Bačová-Mitková & Onderka, 2010).

Assessment of frequency of extreme values of daily rainfall in the city of Sao Paulo, Brazil over the period 1933-2005, based on the peak over threshold (POT) and Generalized Pareto Distribution (GPD) approach. Therefore, there is strong evidence that high quantiles of daily rainfall in the city of Sao Paulo have been increasing in magnitude and frequency over time. For example, 0.99 quantiles of daily rainfall amount have increased by about 40mm between 1933 and 2005 (Manatsa et al., 2008).

Modelling of extremely high rainfall in Limpopo Province of South Africa, using dataset from 1960 to 2020.Daily and monthly rainfall were obtained from South Africa weather service. the parameters of the developed models were estimated using the maximum likelihood method. After the suitable model for data was chosen, the 50-year return level was estimated as 368mm, which means a probability of 0.02 exceeding 368mm in fifty years in the Thabazimbi area.(Sikhwari et al., 2022)

Over East African countries like Kenya and Tanzania, where (Onwuegbuche et al., 2019) applied EVT in predicting climate change induced extreme rainfall in Kenya, and the result showed that the increase in return periods leads to corresponding increase in return levels. And for Tanzania the study of modelling extreme maximum rainfall, the results showed that over the different years of return periods extreme rainfall has less than twenty years of return periods in the most stations across the country and conclude that the frequency of occurrence will increase (Ngailo et al., 2016). Also,

Iyamuremye et al. (2019) described “the change in extreme rainfall using rainfall indices related to extremes for Dodoma” in Tanzania and the research did not discover any linear trend in the extreme rainfall of 77 years. A few research has been done over Rwanda by (Mung’atu, 2018) whose applied EVT to Kigali monthly rainfall data, the best model was used to forecast and this was compared to observed data and checked if the estimated results were in agreement with reality. Estimates from return level showed that extreme rainfall will happen in 11 years.

Furthermore, researchers around the world used EVT to analyze extreme rainfall where they came up of some different results. (Mondal et al., 2015; Wang et al.,2015; Eduardo et al., 2001; Soheil et al., 2016; Bako et al., 2020).

CHAPTER THREE

DATA AND METHODOLOGY

3.0 Data

3.1 Station data

Daily rainfall data used in the study represent three stations of North-western region. Rainfall data ranging from 1981 to 2018 used under this research were collected from Rwanda Meteorology Agency (Meteo Rwanda) located at city of Kigali. Two synoptic stations and one agro-meteorological station across North-Western region of Rwanda were used in this study.

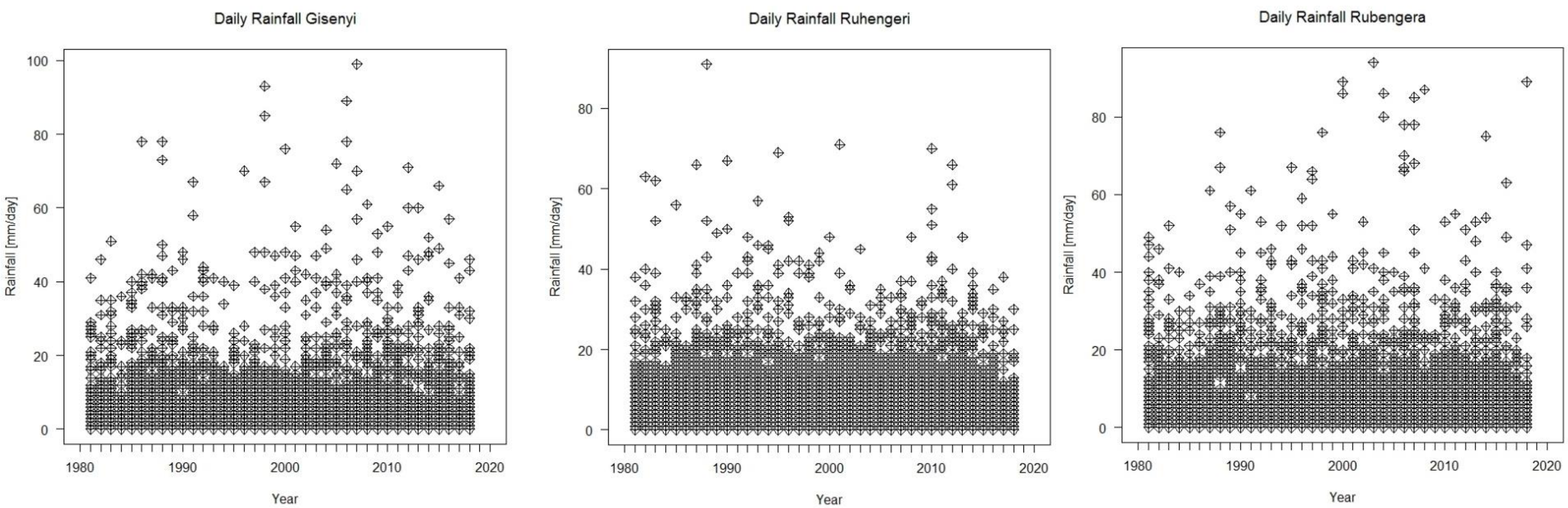


Figure 2:Daily rainfall (mm/day) at Gisenyi, Ruhengeri and Rubengera stations for period of 1981-2018

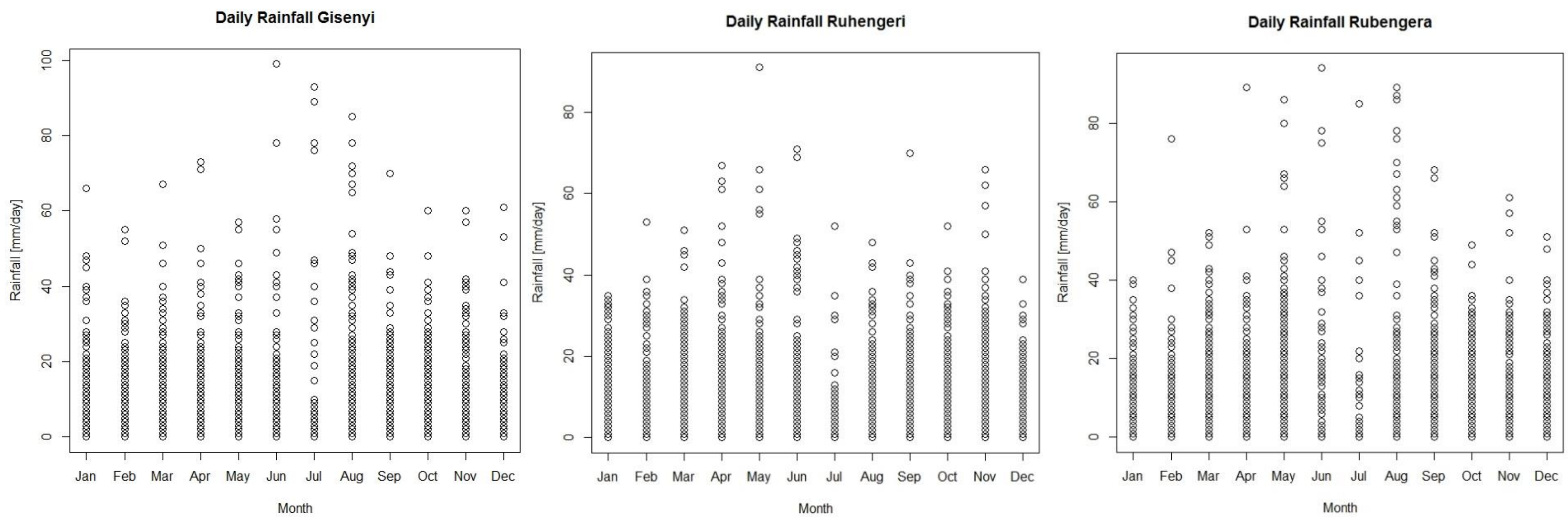


Figure 3:Daily rainfall (mm/day) at Gisenyi, Ruhengeri and Rubengera stations for the period of 1981-2018

3.1.1 Organization and analysis of data.

1. Peak over threshold approach using generalize Pareto distribution.
    - Threshold value and selection
    - Using R packages developed by R development core team (2003).
    - Testing the goodness fitting of model.
2. return period and return level

3.2. Methodology

3.2.1 Extreme value theory (EVT)

Friederichs (2007) has expressed that extreme value theory focus on extreme and rare events, whereas (Goldstein et al., 2003) defined an extreme event as a phenomenon with a very high or low value which has a low probability of occurrence although devastating consequences.
EVT has been applied in many different fields like climatology, finance, and hydrology (Bako et al., 2020). In order to describe the behavior of extreme rainfall at a particular area, it is necessary to identify the distributions which is fitted the data. EVT has two methods namely Peak over threshold (POT) and block maxima method. In this study POT is applied because it provides more efficient usage of the extreme value information.

3.2.1.1 Generalized Pareto distribution (GPD)

This is common approach which deals with modelling the data that exceed the threshold, it has two parameter namely scale parameter (σ) and shape parmeter (ξ) (Friedrichs, 2007). POT method is utilized for numerous areas to distinguish extreme events such as floods, wind velocities, etc. (Andersson et al., 2018). The data which exceed the threshold are modelled according to the GPD.

$$H(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{\frac{-1}{\xi}} \qquad (1)$$

This is called The Generalized Pareto Distribution (GPD), the parameters are estimated using maximum likelihood estimator (MLE)
It has three type of distribution depending on shape parameter ξ known as:

1. Gumbel $\xi = 0$, exponential tail
2. Frechet $\xi > 0$, Polynomial tail behavior
3. Weibull $\xi < 0$, has upper end point

The shape parameter ξ has significant effects of identifying qualitative behavior of Generalized Pareto distribution (Coles, 2015).

3.3 Peak over threshold

The POT approach models excess values of a sample over a given (high) threshold within a time period, and estimates the tail behavior using the conditional distribution of these exceedances series contains all peak flows that are greater than a given threshold μ flow. The rainfall episodes with amounts higher than 99% were chosen as extreme. The percentage of 99% is in agreement with Anagnostopoulou and Tolika (2017) who proved that this is the most appropriate threshold for extreme rainfalls in Europe.

Advantage: more efficient use of data. Consider several large values instead of only the largest one.

Given a threshold μ , the distribution function of extreme values of X over μ is,

$$F_\mu(x) = P(X - \mu \leq x | X > \mu) = \frac{F(x+\mu) - F(\mu)}{1 - F(\mu)} \tag{2}$$

Where $F_\mu(x)$ is called the conditional excess distribution which represents the probability that the value of X exceeds the threshold μ by most amount where $x = X - \mu$.

3.3.1 Threshold selection

They exist many different methods to obtain threshold, among of them we choose two. If the chosen is too high, and if the threshold is too low the larger amount of measurements will be taken into account, therefore leading to bias. The two chosen threshold used in this study is mean residual life plot and parameter stability plot.

3.3.1.1 Mean residual life plot (MRL)

Graphic method for threshold selection has been suggested by Smith and Davison (1990), this is based on the mean of GPD. If Y has a generalized Pareto distribution with parameters σ and ξ ,then

$$E(Y) = \begin{cases} \frac{\sigma}{1-\xi} \\ \infty \end{cases} \tag{3}$$

Suppose that the GPD is valid as a model for excesses of a certain threshold μ_0 generated by a series , $X_{1,\dots,}X_n$ and denoted by X

$$E\ (X - \mu_0 \backslash X > \mu_0) = \frac{\sigma\mu_0}{1-\xi}\ \ ,\ \xi < 0 \tag{4}$$

$$E\ (X-\mu \backslash X>\mu)=\frac{\sigma\mu}{1-\xi}=\frac{\sigma\mu_0+\xi_\mu}{1-\xi} \tag{5}$$

$$\left\{\left(\mu,\frac{1}{\eta_\mu}\sum_{i=1}^{\eta_\mu}\left(x_{(i)}-\mu\ \right)\right);\ \mu\ <x_{max}\right\} \tag{6}$$

3.3.1.2 The parameter stability plot

the parameter stability plot also called threshold stability plot is another graphical method which is commonly used to determine the threshold μ (Scarrot and Mac-Donald,2012). if the exceedances of a high threshold μ_0 follow a GPD with parameters ξ and $\sigma_{\mu 0}$, then for any threshold μ, $\mu > \mu_0$, the exceedances still follow a GPD with $\xi_\mu = \xi$ and $\sigma_\mu = \sigma_{\mu 0} + \xi(\mu - \mu_0)$

σ^*=$\mu_{0-\xi_\mu\mu}$ The new parameterization does not depend on μ any longer, given that μ_0 is a reasonably high threshold.

The plot is defined by the focus of points

$\{(\mu, \sigma^*);\ \mu < x_{max}\}$ and

$$\{(\mu,\xi_\mu);\mu<x_{max}\} \tag{7}$$

Note that the threshold should be chosen at the value where the shape and scale parameters remain constant (Bommier, 2014)

3.4 Model diagnostics

The reason for fitting a statistical model to data is to make conclusions about some aspect of the population from which the data were drawn. Such conclusions can be sensitive to the accuracy of the fitted model; therefore, it is essential to check if the model fits well.

3.4.1 Quantile-Quantile plot(Q-Q)

A quantile plot is a useful tool used to check if the assumed distribution fits well the studied data set. the plot should show linearity i.e. almost all points should lie to the line

$$\left\{\left(\hat{F}^{-1}\left(\frac{i}{n+1}\right)\right), x_{(i)}\right) : \quad i = 1, \dots n\right\} \quad (9)$$

3.5 Estimation of the parameters

Shape and scale parameters have been estimated using MLE, and shape parameter is a helpful tool to check if the chosen distribution is appropriate. ξ varies between -0.5 and 0.5

$$l(\sigma, \xi) = -q log(\sigma) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{q} log\left(1 + \frac{\xi x_i}{\sigma}\right) \quad (11)$$

$$l(\sigma) = -q log(\sigma) - \left(\frac{1}{\sigma}\right) \sum_{i=1}^{q} x_i \quad (12)$$

3.6 Goodness-of-fit test (GoF)

This application is essential for checking the adequacy of probability. It is provided that GPD distribution can characterize the extreme rainfall behavior with high accuracy.

3.7 Return level

Return level plot consists of the locus of point. It is usual to plot the return level curve on a logarithmic scale to emphasize the effect of extrapolation, and also to add confidence bounds and empirical estimates of the return levels (Coles, 2015). For the GPD model, the return level is given by x_q which defines the extreme level that exceeded the average once every q observations (Reuder et al., 2016).

$$x_{q=} \begin{cases} \mu + \frac{\sigma}{\xi}\left[\xi_{\mu} - 1\right], & \xi \neq 0 \\ \mu + \sigma log\left(q\xi_{\mu}\right), & \xi = 0 \end{cases} \quad (12)$$

CHAPTER FOUR

RESULTS AND DISCUSSION

Extreme value analysis using peak over threshold approach of the daily rainfall data in North-Western of Rwanda and discussion of the results obtained are based on the methodologies described above. The R-statistical software is used to conduct the analysis.

4.1 Threshold selection

Thresholds appropriate chosen based on mean residual life plots and parameter stability plots for three stations, Ruhengeri, Gisenyi and Rubengera. When threshold was chosen diagnostics plots were analyzed in order to see if generalized Pareto distribution was a suitable distribution for the exceeding values. thresholds selection is represented in the addendum one.

We select one constant threshold for every station. The MRL plot is plotted for Ruhengeri, Gisenyi and Rubengera stations. The possible threshold is the point when the mean excess plot shows linearity, and estimated parameters look stable at different threshold. the threshold for Ruhengeri was chosen at 30.as can be seen in the mean residual life plot there is some evidence for linearity above $\mu = 30$ as shown in figure 5. Parameter threshold stability plot of the modified scale parameter for threshold up to 50mm. Dots represent the estimated modified scale parameter for a GPD model for some threshold. The lines represent a 95% confidence interval. The plot was made with tcplot from the evt package in R as represented in figure 4. Same procedure was made for all other two stations. Figure 9 strengthens that generalized Pareto distribution is a reasonable distribution for the exceeding values.

4.2 Model diagnostic: Quantile- Quantile plot (Q-Q)

The QQ-plots shows that all the points are approximately linearly distributed along the unit diagonal. Quantile plot for all three stations GPD function provides a good fit between empirical and model quantile**.**

4.3 parameter estimation for extreme rainfall

Maximum likelihood estimation method (MLE) is used to estimate the parameters of generalized Pareto distribution with 95% confidence interval (CI). they were estimated using profile likelihood. This method gives the result for the various confidence intervals which are high and low CI as show in table 2. The profile likelihood estimates are indicated in the figure 6.

Gisenyi station has a negative shape which implies that the distribution belongs to the Weibull distribution. the confidence interval of the station includes positive value as well so we cannot ignore the fact that it might belong to one of other distributions. In the case when the MLE of the shape parameter $\xi < 0$ this implies that the support of the GPD with upper endpoint.

Ruhengeri and Rubengera stations have positive shape parameters $\xi > 0$, and when we are looking at the GPD confidence interval values are opposite. This implies that any assumption cannot be made. For all stations the confidence interval of ξcontains zero, thus, the exponential distribution fits the data well.

4.4 Cumulative distribution GPD

Figure 8 shows that all stations representing N-W of Rwanda, the empirical represented in black and model in blue are reasonably linear.

4.5 Return level for extreme rainfall

The estimation of the precipitation return levels, offers a common way to estimate the climatic risk, usually based on historical data. In this study, the return levels of extreme rainfalls were calculated for GPD using MLE (table3). According to table 3, Rubengera presents the same amount 112.55mm of return level over the period of 50 and 80 years, and also the maximum return level of 131.52mm over the period of 200 years whereas Ruhengeri station presents minimum rainfall of 75.41mm over the period of 20 years. it can be seen from table 3 the results that increase in return periods lead to a corresponding increase in return levels.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

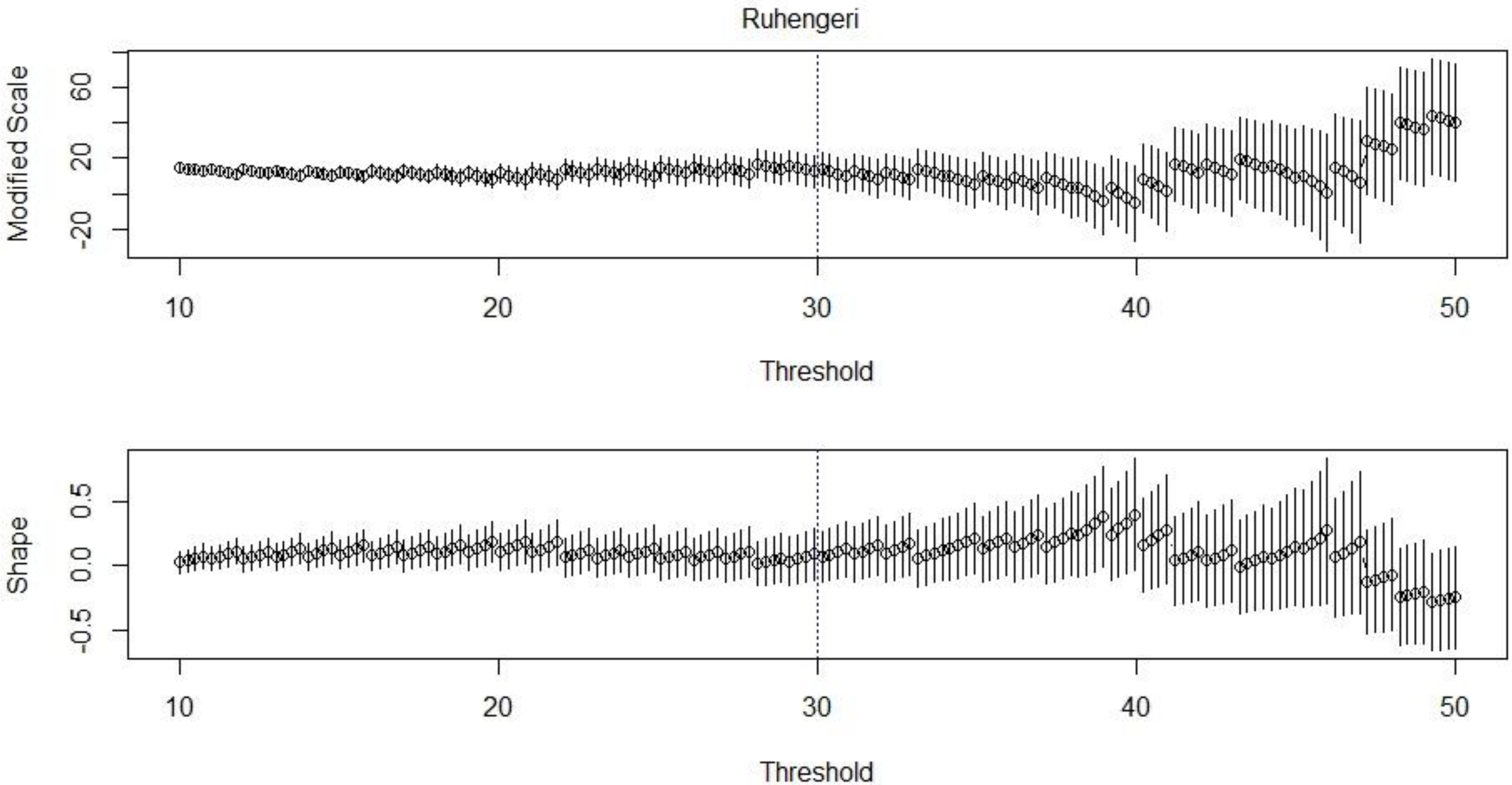
This research demonstrates the application and significance of EVT at describing extreme rainfall events in Northwestern of Rwanda. The GPD model is considered for daily rainfall data in N-W from 1981 to 2018. The model parameters were estimated using the Maximum Likelihood Estimation, peak Over Threshold method was used to fit the Generalized Pareto Distribution. Our results reveal that the exponential and Weibull distribution is the optimal model from the GPD family for daily rainfall data over the threshold value of 31, 30 and 33(mm) for Gisenyi, Ruhengeri and Rubengera stations. The results showed that increase in return periods leads to a corresponding increase in return levels. When comparing the return levels for the GPD of all stations, our results show that the GPD gives higher return levels for 200 years over Rubengera station compared to the rest. However, the lower return periods indicate 20 years over Ruhengeri. The model diagnostics showed that the models were reasonable for modelling the rainfall data. This study will help decision makers in Northwestern of Rwanda with knowledge about extreme rainfall events in the return periods considered, to enable them make appropriate decisions to reduce damage to crops, infrastructure and lives that is caused by extreme rainfall. As climate change persists, continuous preparedness and adaptation measures are essential for the Northwestern communities. Thus, this research will be useful in coming up with flood risk early warning, management, preparedness, response and mitigation

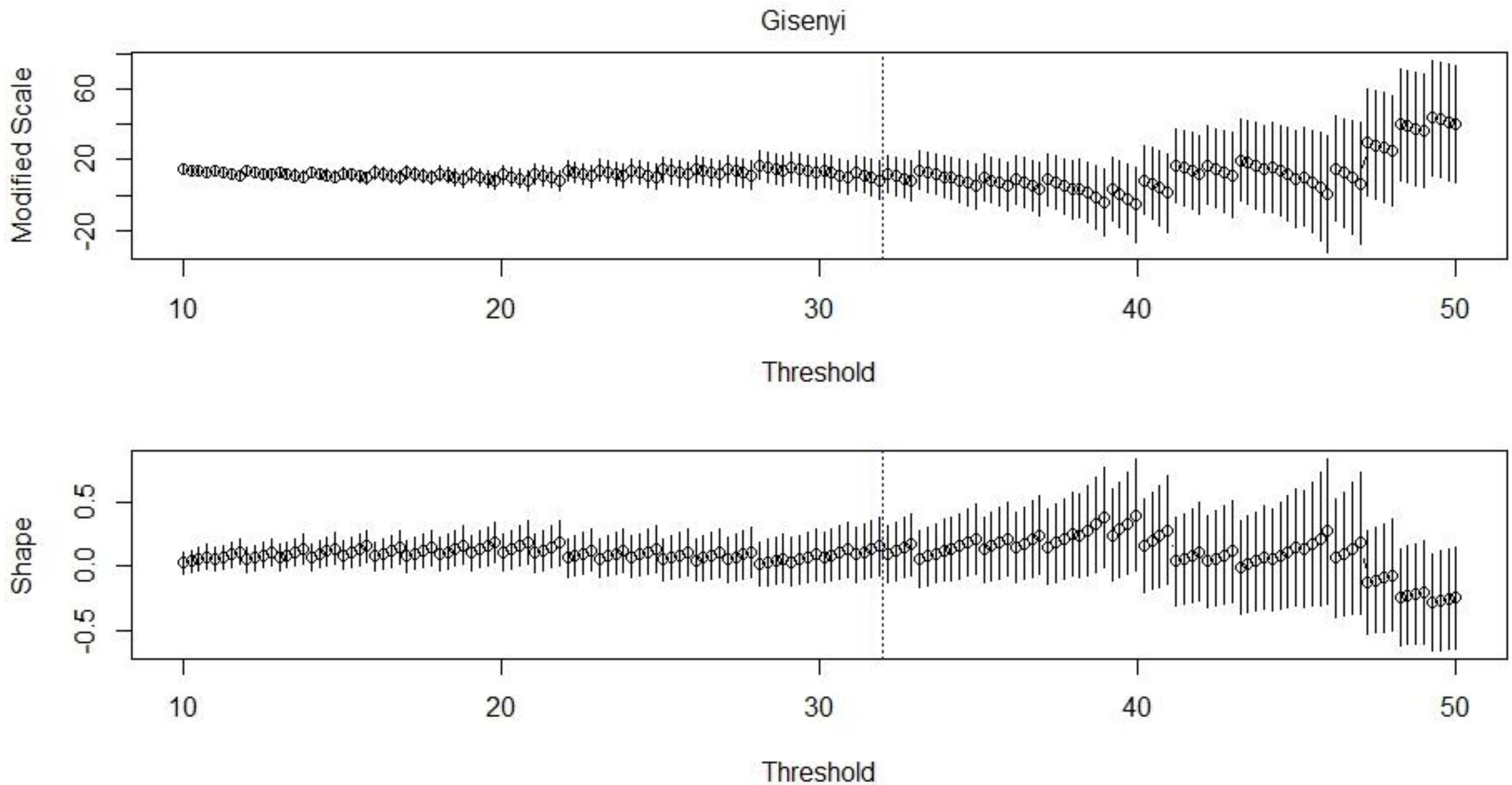
The aim of this work was to study the analysis of extreme rainfall over northwestern of Rwanda using the peak over threshold method. As the consider emphasized on extreme daily rainfall and found that the return periods increase as the return levels keep increase, this indicate that N-W of Rwanda when we will be having rainfall greater than the above threshold obtained then it will be taken as extreme, therefore, decision makers should take some measure to develop the early alarming systems to warn the public.

To keep on encourage he people living in the rural areas to follow the rules and regulations advised by the authorities.

However, future studies can model and predict extreme rainfall in Rwanda with respect to specific regions in the country using both generalized Pareto distribution and generalized extreme value distribution. Also, modelling both extreme rainfall and temperature in Rwanda is a possible research direction.

Addendum 1: List of figures




Gisenyi
Modified Scale
-20
20
60
10
20
30
40
50
Threshold
Shape
-0.5
0.0
0.5
10
20
30
40
50
Threshold

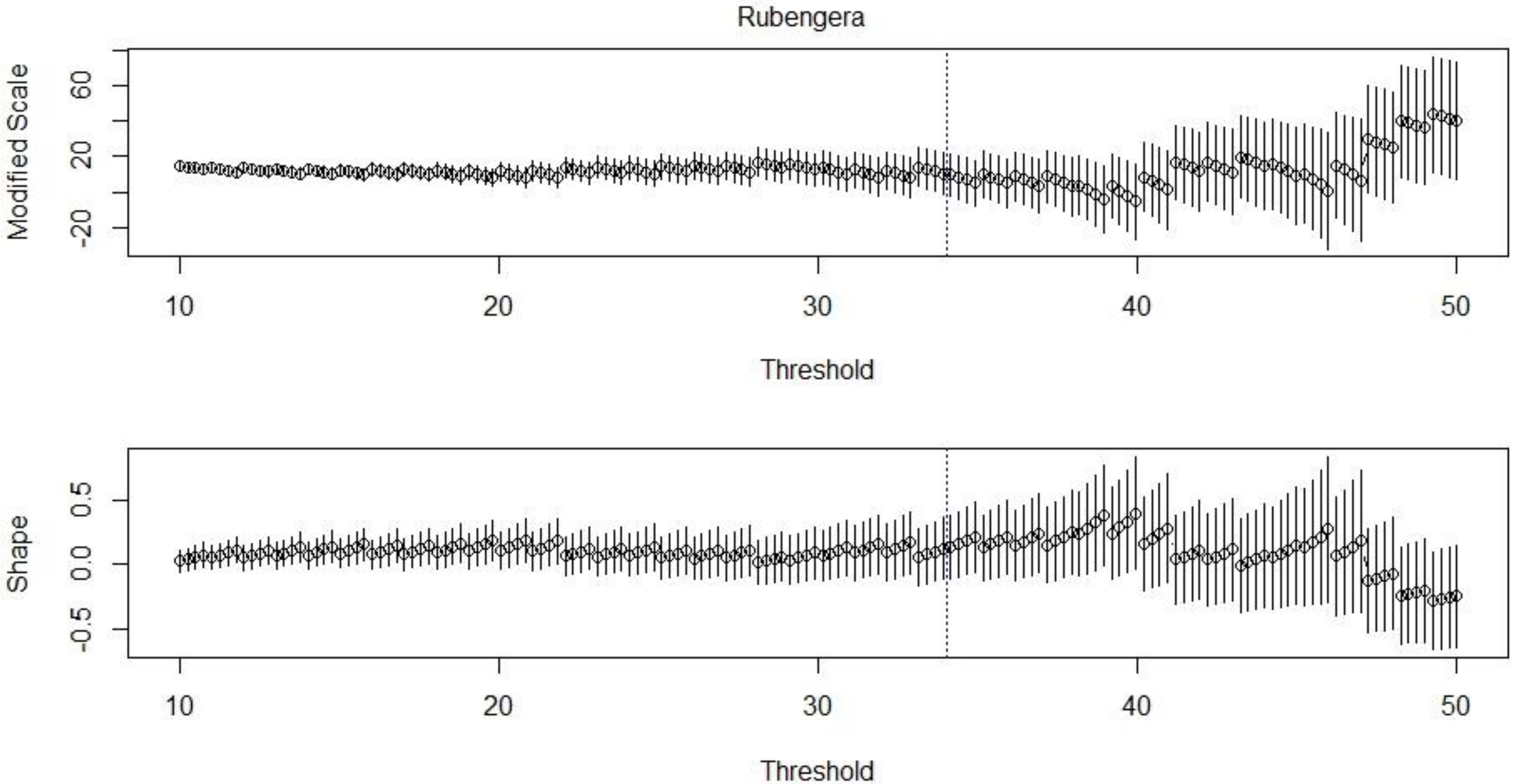
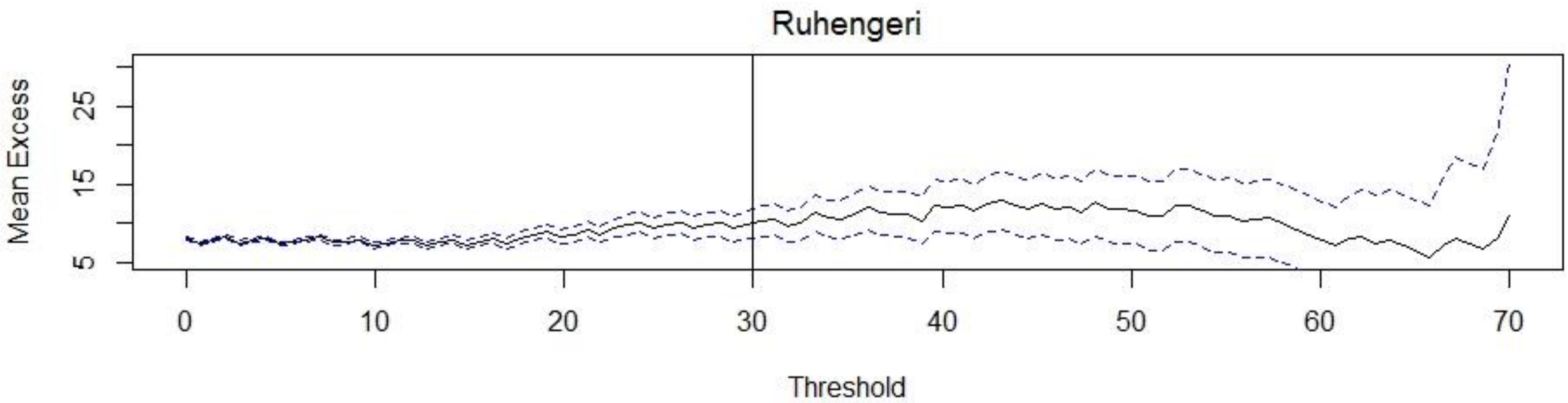
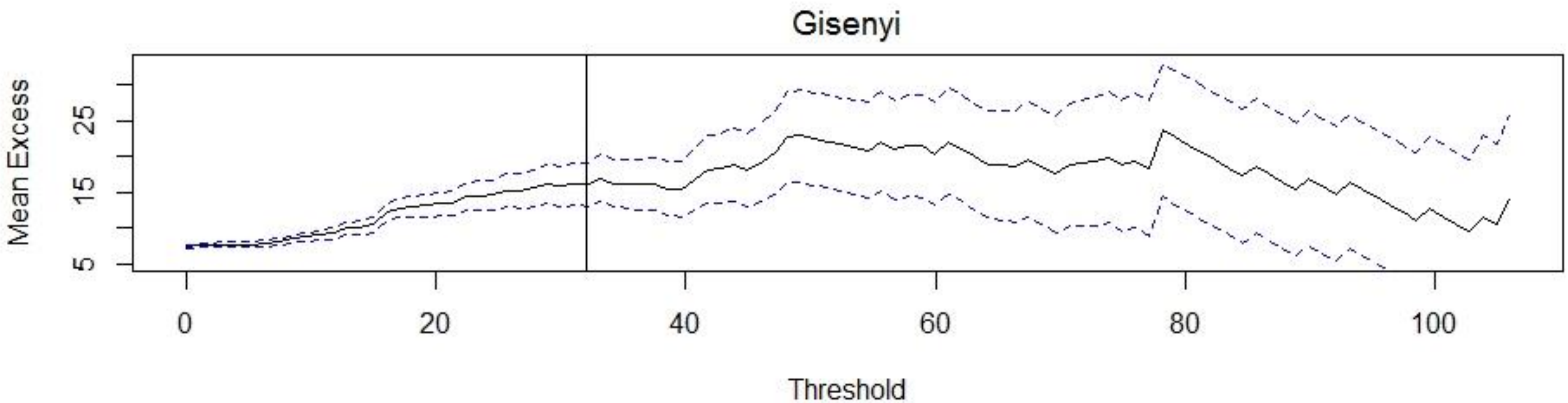


Figure 4:Threshold selection diagnostic with tcplot modified scale (up) and shape (down) for the stations of Gisenyi, Ruhengeri and Rubengera.


Ruhengeri
Mean Excess
5
15
25
0
10
20
30
40
50
60
70
Threshold


Gisenyi
Mean Excess
5
15
25
0
20
40
60
80
100
Threshold

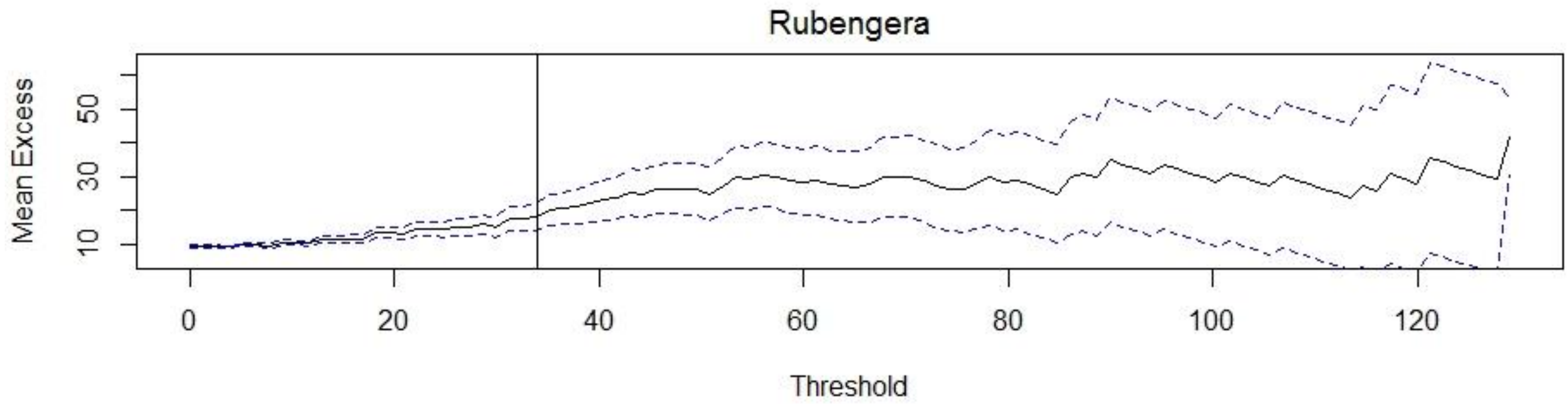


Figure 5:Threshold selection diagnostic with Mean excess Residual Plot for Gisenyi, Ruhengeri and Rubengera stations.

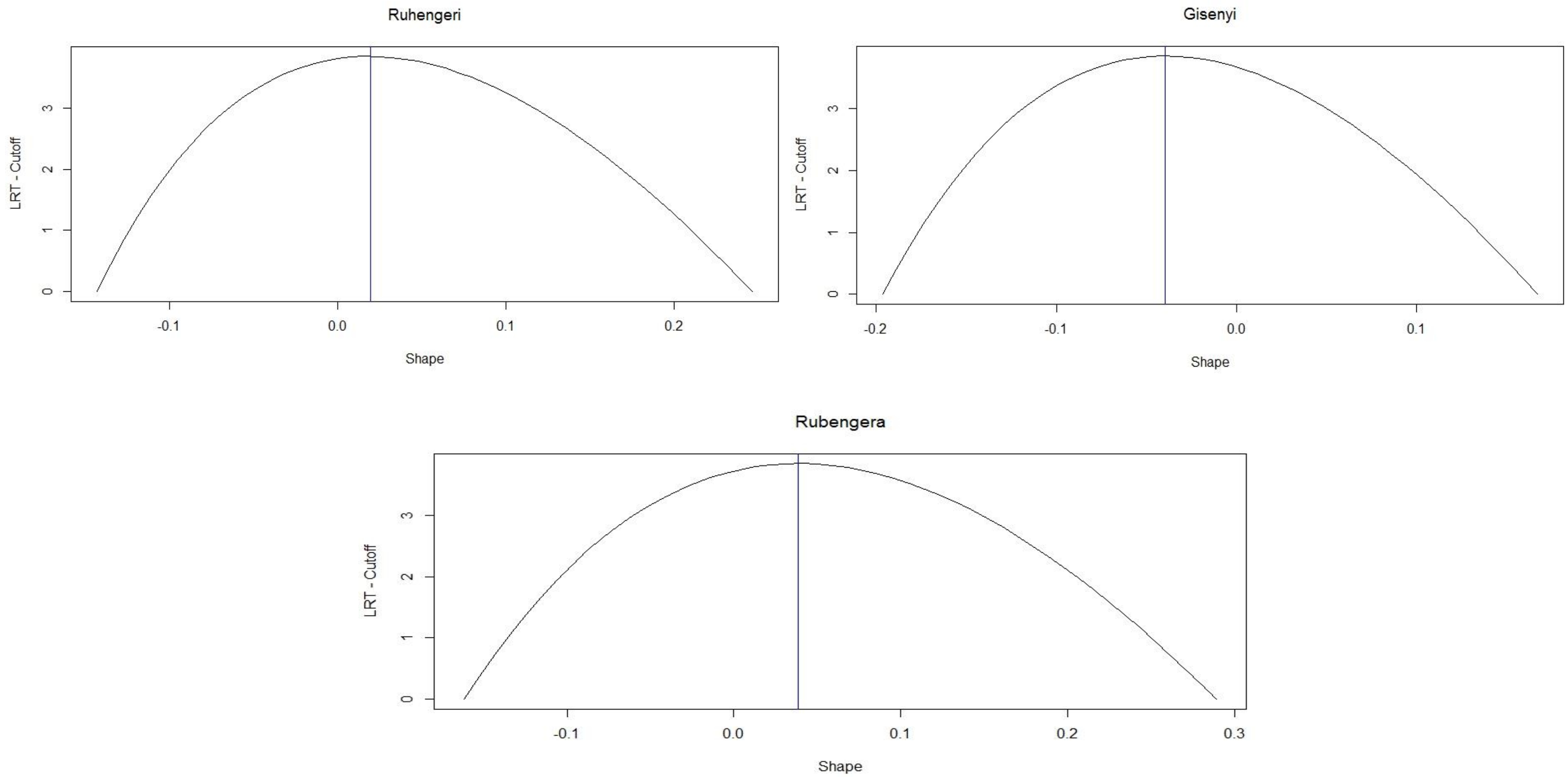


Figure 6:Profile likelihood GPD for Gisenyi, Ruhengeri and Rubengera stations.

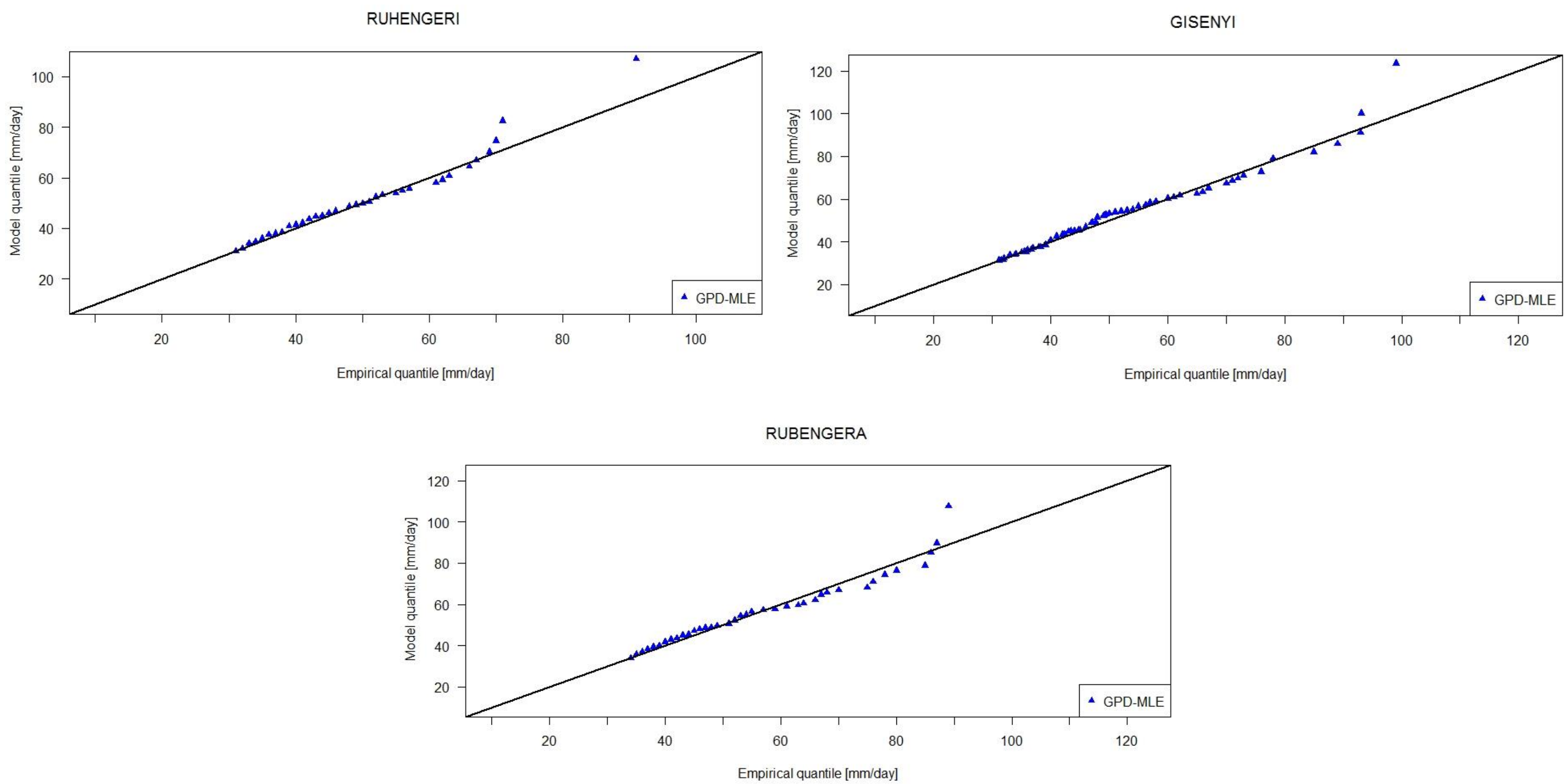


Figure 7:Q-Q plots for Gisenyi, Ruhengeri and Rubengera stations.

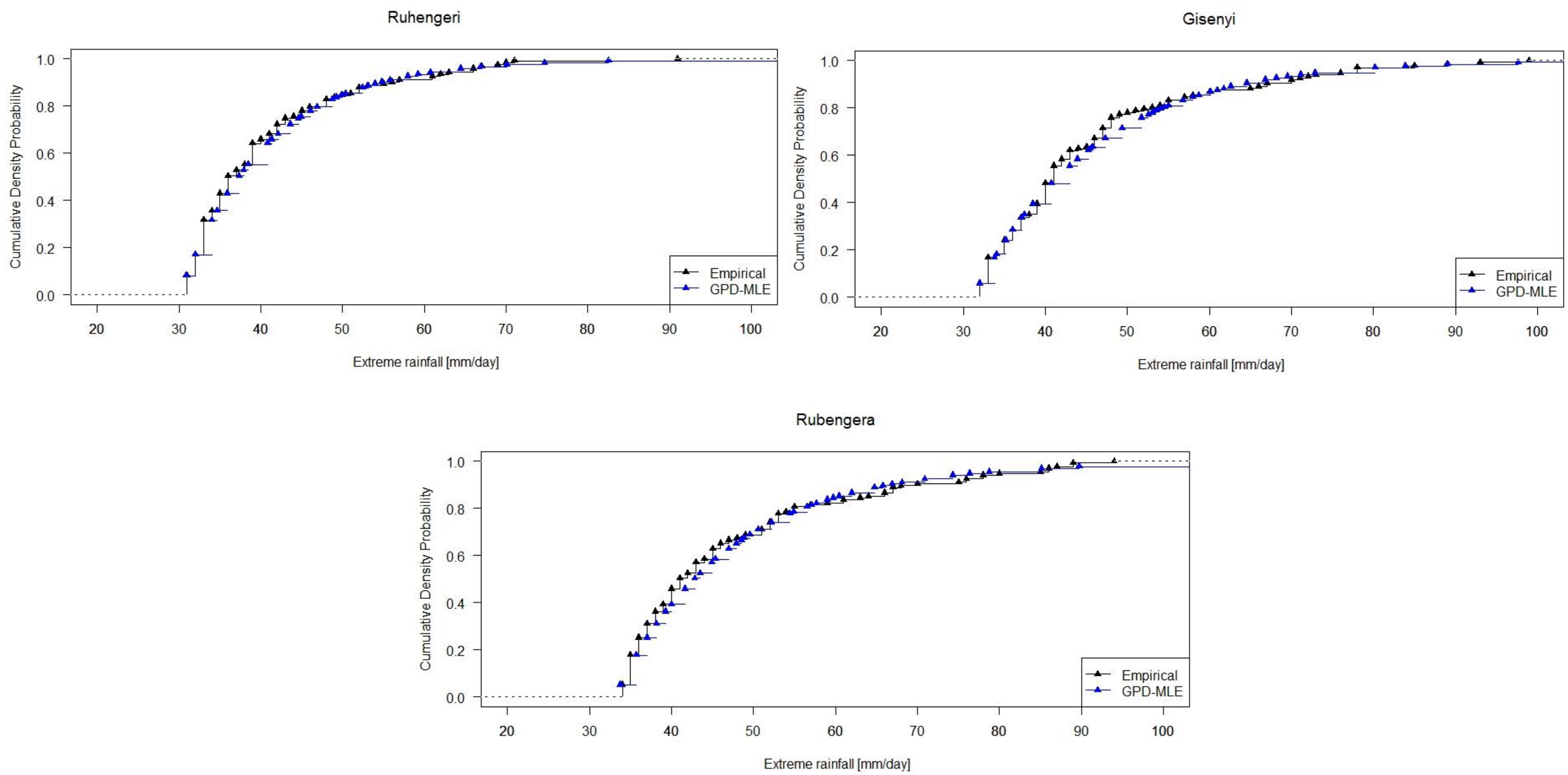


Figure 8: Cumulative distribution function for Gisenyi, Ruhengeri and Rubengera stations.

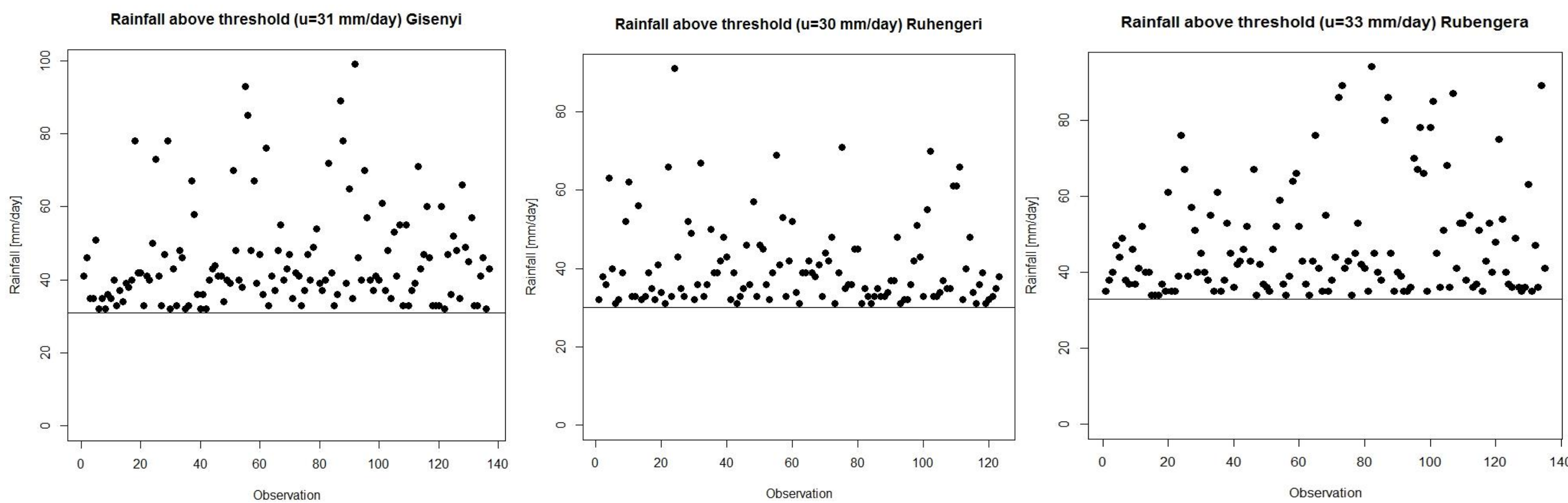


Figure 9: Rainfall above threshold at Gisenyi, Ruhengeri, and Rubengera stations for the period of 1981-2018.

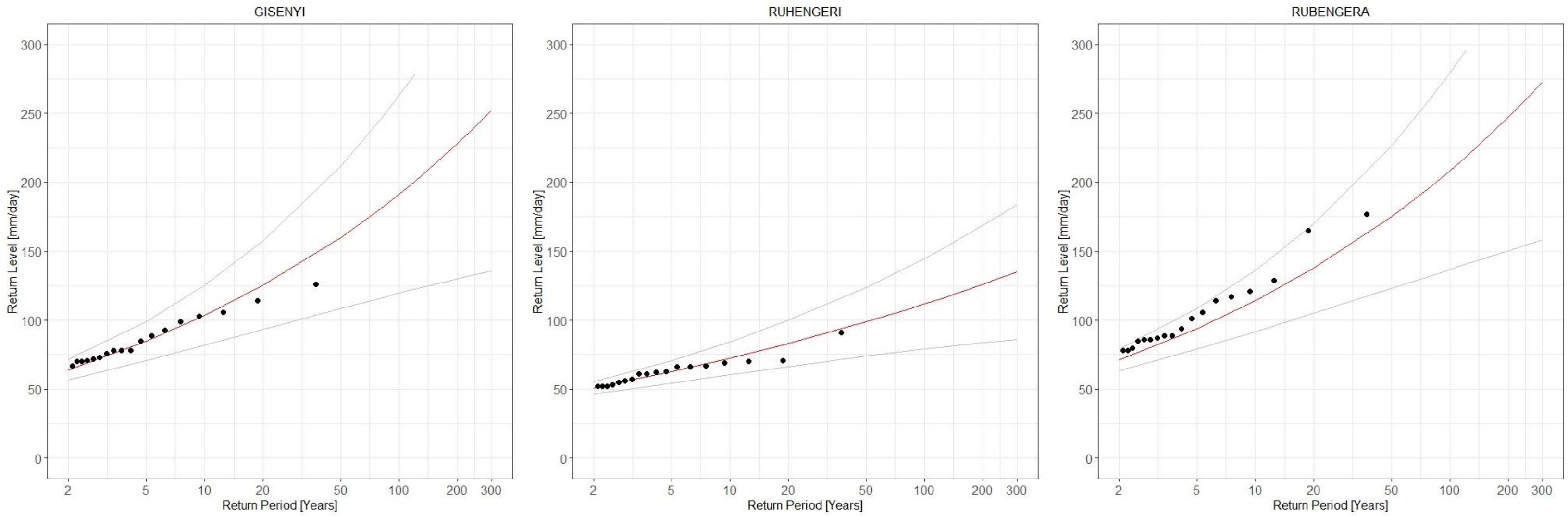


Figure 10: Return level of extreme rainfall for the stations of Gisenyi, Ruhengeri and Rubengera.

Addendum 2: List of tables

Table 1:Geographic coordinates of rainfall stations used in the study

Stations	**Years of study**	**Latitude in Decimal degrees**	**Longitude in Decimal degrees**	**Altitude in meter**
Gisenyi	1981-2018	-1.6761	29.2599	1554
Ruhengeri	1981-2018	-1.48	29.61	1878
Rubengera	1981-2018	-2.07	29.41	1700

Table 2:Shape parameters and scale parameters with their 95% CI intervals for the three stations

Station	Shape			Scale		
	Low CI	**Estimated**	**High CI**	**Low CI**	**Estimated**	**High CI**
Ruhengeri	-0.174	0.018	0.211	7.703	10.435	13.167
Gisenyi	-0.218	-0.039	0.139	8.409	11.345	14.282
Rubengera	-0.183	0.038	0.260	11.14	14.436	17.731

Table 3:Return level plot of extreme rainfall for the stations of Gisenyi, Ruhengeri and Rubengera

Return period (year)	**Return level (mm)**		
	GISENYI	RUHENGERI	RUBENGERA
20	90.22	75.41	97.33
50	102.13	85.14	112.55
80	104.13	87.84	112.55
100	108.15	91.26	120.57
150	110.98	93.85	124.43
200	116.10	98.58	131.52

REFERENCES

Bačová-Mitková, V., & Onderka, M. (2010). Analýza extrémnych hydrologických situácii na dunaji využitim metódy pot. *Journal of Hydrology and Hydromechanics*, *58*(2), 88–101. https://doi.org/10.2478/v10098-010-0009-x

Coles, S. (2015). An introduction to extreme values. In *The effects of brief mindfulness intervention on acute pain experience: An examination of individual difference* (Vol. 1).

Friederichs, P. (2007). *An Introduction to Extreme Value Theory Applications of EVT*.

GoR. (201 C.E.). *The National Risk Atlas of Rwanda The National Risk Atlas of Rwanda*.

Iyamuremye, E., Wanyonyi, S. W., & Mbete, D. A. (2019). Analysis and Modelling of Extreme Rainfall: A Case Study for Dodoma, Tanzania. *Asian Journal of Probability and Statistics*, *3*(2), 1–29. https://doi.org/10.9734/ajpas/2019/v3i230086

J.ean Baptiste, N., & Habiyaremye, G. (2014). High Risk Zones on Floods and Landslides Disasters in Rwanda. *Proceedings of the International Conference "InterCarto/InterGIS," 1*(20), 262. https://doi.org/10.24057/2414-9179-2014-1-20-262

Manatsa, D., Chingombe, W., & Matarira, C. H. (2008). The impact of the positive Indian Ocean dipole on Zimbabwe droughts Tropical climate is understood to be dominated by. *International Journal of Climatology*, *2029*(March 2008), 2011–2029. https://doi.org/10.1002/joc

Muhire, I., Ahmed, F., & Elbasit, A. (2015). Spatio-temporal variations of rainfall erosivity in Rwanda. *Journal of Soil Science and Environmental Management*, *6*(4), 72–83. https://doi.org/10.5897/JSSEM14

Mung'atu, U. O. J. K. (2018). Modelling Extreme Maximum Rainfall Using Generalized Extreme Value Distribution: Case Study Kigali City. *International Journal of Science and Research (IJSR)*, *7*(6), 121–126. https://doi.org/10.21275/ART20183033

Ngailo, T. J., Reuder, J., & Rutalebwa, E. (2016). Modelling of Extreme maximum Rainfall using Extreme Value Theory for Tanzania. *International Journal of Scientific and Innovative Mathematical Research*, *4*(3), 34–45. https://doi.org/10.20431/2347-3142.0403005

Ngarukiyimana, J. P., Fu, Y., Yang, Y., Ogwang, B. A., Ongoma, V., & Ntwali, D. (2018). Dominant atmospheric circulation patterns associated with abnormal rainfall events over Rwanda, East Africa. *International Journal of Climatology*, *38*(1), 187–202. https://doi.org/10.1002/joc.5169

Okoola, R. E. (1999). Midtropospheric circulation patterns associated with extreme dry and wet episodes over equatorial Eastern Africa during the Northern Hemisphere spring. *Journal of Applied Meteorology*, *38*(8), 1161–1169. https://doi.org/10.1175/1520-0450(1999)038<1161:MCPAWE>2.0.CO;2

Onwuegbuche, F. C., Kenyatta, A. B., Affognon, S. B., Enock, E. P., & Akinade, M. O. (2019). Application of Extreme Value Theory in Predicting Climate Change Induced Extreme Rainfall in Kenya. *International Journal of Statistics and Probability*, *8*(4), 85. https://doi.org/10.5539/ijsp.v8n4p85

Red Cross. (2013). *DREF final report Rwanda : Floods*.

REMA. (2013). *The assessment of economic impacts of the 2012 wet season flooding in Rwanda*. 1–103.

Sikhwari, T., Nethengwe, N., Sigauke, C., & Chikoore, H. (2022). Modelling of Extremely High Rainfall in Limpopo Province of South Africa. *Climate*, *10*(3), 1–17. https://doi.org/10.3390/cli10030033

USAID. (2019). *Rwanda Tropical Forests and Biodiversity Analysis*. *June*.