

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

6,900

Open access books available

186,000

International authors and editors

200M

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



On Modelling Extreme Damages from Natural Disasters in Kenya

Carolyn Ogotu and Antony Rono

Abstract

We seek to develop a distribution to model the extreme damages resulting from Natural Disasters in Kenya. The distribution is based on the Compound Extreme Value Distribution, which takes into account both the distributions of the frequency of occurrence and magnitude of the events. Threshold modelling is employed, where the extreme damages are identified as the points that lie above a sufficiently high threshold. The distribution of the number of the exceedance is found to be Negative Binomial, while that of the severity is approximated by a Generalised Pareto Distribution. Maximum likelihood estimation is used to estimate the parameters, and the log-likelihood is maximised using numerical methods. Probability weighted moments estimation is used to determine the starting values for the iterations. Prediction study is then carried out to investigate the performance of the proposed distribution in predicting future events.

Keywords: natural disasters, extreme events, compound distributions

1. Introduction

A natural disaster is a sudden and adverse event caused by natural forces that significantly disrupts the workings of society [1]. Such events result to loss of life, injury or other negative health impacts, loss of livelihoods, environmental damage, social and economic disruption. These natural forces causing natural disasters are called natural hazards. A natural hazards can be formally defined as a natural process or phenomenon that may cause loss of livelihoods and services, negative health impacts such as loss of life and injury, social and economic disruption or environmental damage [2]. Therefore, a natural hazard becomes a disaster once it occurs. The Centre for Research on the Epidemiology of Disasters Database (EM-DAT) generally classifies natural disaster according to the hazards that causes them. We have geophysical (earthquakes, volcanoes landslides and tsunamis), hydrological (floods and avalanches), climatological (droughts and wildfires), meteorological (storms and cyclones) or biological (epidemics and pests/insects/plagues). Hydrological, meteorological and climatological hazards are generally termed as weather-related hazards. Natural disasters are considered extreme events and thus we consider the method of Extreme Value Theory for analysis of our data.

1.1 Extreme value modelling

Extreme value theorem (EVT) is used to develop stochastic models aimed at solving real-world problems relating to extremal and unusual events for instance

stock market crashes, natural catastrophes, major insurance claims e.t.c. EVT models the tails of loss severity distributions as it only considers extreme and rare events. Extreme value modelling has been used widely in hydrology, insurance, finance and environmental application where extreme events are of interest. Due to the scope of this research, we will limit the literature to those relating to natural disasters. The relevant theories and statistical methods behind EVT are presented by [3].

[4, 5] propose the Generalised Extreme Value (GEV) distribution which is used to model the maxima of a set of independent and identically distributed (iid) events. The resulting model, referred to as the Block Maxima (BM) model, involves dividing the observations into non-overlapping blocks of equal sizes and restrict the attention to maximum observations in each period. The observations created are then assumed to follow the GEV [6]. The block sizes are vital in the EVT model as too small block sizes can lead to poor asymptotic approximation, while too big sizes will result to very little observations [7]. [8] use BM method to model hydrological floods and droughts. They find that using block sizes of one year or less to model drought leads to bias. Floods, on the other hand, is modelled using block sizes of less than one year. The same method is used by [9] to study seasonal variation of extreme precipitation across the UK. They use a one month block, having found little improvement of a longer-block. They were able to identify regions with a high risk of extreme precipitation in a given season. A major drawback of the Block Maxima is its inefficiency in terms of data usage. Dividing the dataset into blocks is wasteful of the data since not all the available data is used. To overcome this challenge, the data above some sufficiently high threshold is modelled in what is commonly referred to as excess over threshold (EOT). Several studies have been conducted to compare the performance of BM and EOT, all based on simulated data. [10] states that, in the case where the shape parameter is equal to zero, EOT estimates for a high quantile is better only if the number of exceedance is larger than 1.65 times the number of blocks. [11] find that EOT is preferable when the shape parameter is greater than zero and when the number of exceedance is greater than the number of blocks. However, when the shape parameter is less than zero, BM is more efficient. [12] carries out a vast simulation study to compare the two methods based on their accuracy as measured by mean squared errors, on the basis of time series with various characteristics and in estimating probabilities of exceedance. He finds that the EOT estimates for a sample that has an average of two or more exceedance per year are more accurate than the corresponding BM estimates. In addition, he finds that EOT should be used when the data is less than 50 years. When there is 200 years of data, both approaches have similar accuracies. In terms of the return value estimates, he finds that the EOT estimates are significantly more accurate than those obtained by BM. [6] find that the BM method at an optimal level gives lower asymptotic minimal square error than the EOT method. They conclude that BM is more efficient than EOT in practical situations.

Despite having mixed results, we can deduce that EOT is, on average, a preferred method to BM. Two features are however dominant from all these studies: first, EOT is more efficient than BM if the number of exceedance is greater than the number of blocks and secondly, the two methods have comparable performances for large sample sizes.

The biggest challenge in using EOT is selecting the appropriate threshold. The threshold selection process is always a trade-off between the bias and variance [13]. Moreover, the choice of the threshold significantly affects the tail of the distribution and hence the GPD parameters. The traditional extreme value modelling approach uses fixed threshold approach, where the threshold is chosen and fixed before the model is fitted. The threshold is usually chosen based on various diagnostics plots

that is used to evaluate the model fit. The traditional EVT approach only considers the extreme data above the threshold, and discards the rest that are below. [7] states that the motivation for ignoring extreme data is that extreme and non-extreme data are caused by different processes and the latter is rarely of concern. Furthermore, there is concern that the data below the threshold may compromise the examination of the tail fit. However, there has been increasing interests in extreme value mixture models, which uses all the data for parameter estimation in the inference. This class of models has the potential to overcome some of the difficulties encountered in using the traditional EVT, in regards to threshold selection and uncertainties associated with it. The next section will highlight the development of extreme value mixture models.

1.2 Extreme value mixture models

Extreme value mixture models typically comprise of two components: a bulk model which describes all the non-extreme data below the threshold, and a tail model which is the traditional extreme value model to model the extreme data above the threshold. The advantage of these models is that they consider all the data without wasting information and provides an automated approach both for estimation of the threshold and quantification of the resulting uncertainties from the threshold choice [7].

[14] propose a dynamic weighted mixture model combining a light-tailed distribution for the bulk model, with GPD for the tail model. They use a Weibull distribution as the bulk distribution. Rather than explicitly defining the threshold, they adopt a Cauchy distribution function as the mixing function to make the transition between the bulk and tail distributions. They use maximum likelihood estimation for parameter estimation, and numeric iteration to calculate empirical quantiles. They then test the model using Danish fire loss dataset, and find the parameter estimates to be comparable to those reported in the literature regarding EOT inference. The quantiles are comparable as well. The model is also compared to an existing robust thresholding approach, and it is found to be better for very small percentiles in small datasets.

[15] develop an extreme value mixture model comprising of a truncated gamma to fit the bulk distributions and GPD for the tail distributions. They note that any other parametric distribution can also be used. They treat the threshold as an unknown parameter, and use Bayesian statistics for inference about the unknown parameters. Inference on the Posterior distribution is done using Markov chain Monte Carlo methods, and simulation studies is carried out to analyse the performance of the model. They find that the parameter estimates are very close to the true value for large samples. However, for small samples, they encounter problems with convergence, and consequently parameter estimation. They also apply the model to Nasdaq 100 dataset, an index in the New York Stock Exchange, and compare its performance with the traditional extreme value model. The resulting GPD estimates is close to that obtained using the traditional approach.

1.3 Compound extreme value distribution

Unlike Extreme Value Theorem which is concerned with only the tails of loss severity distributions, Compound Extreme Value Distribution (CEVD) models can simultaneously model the frequency and severity of extreme events. The idea was proposed by Tebfu and Fengshi in 1980 to model typhoon in South China [16]. They only consider the events above the threshold, as is the case with the traditional EVT, and assume that the number of exceedance (frequency) is Poisson Distributed.

The exceedance (severity) are, on the other hand, assumed to follow a Gumbel Distribution, which is a special case of the General Extreme Distribution. Hence, the model is usually referred to as Poisson-Gumbel. Tebfu and Fengshi [7] also used CEVD to model hurricane characteristics along the Atlantic coasts and the Gulf of Mexico. They assume that the number of exceedance follows a Poisson Distribution, and the exceedance is Weibull Distribution.

Initially, CEVD was mostly used in hydrology, to model wave heights and the resulting extreme events. The model has been successfully used to predict design wave height. For instance, Hurricane Katrina of 2005 corresponded to 60 years return period as predicted by the Poisson-Weibull model [17]. As a result, there has been several extensions to these class of models including the Bivariate Compound Extreme Value Distribution (BCEVD) model. [18] and Multivariate Compound Extreme Value Distribution (MCEVD) model [17]. In addition, the model has been adopted in a wider range of areas, including finance, insurance, disasters and catastrophic modelling.

[19] investigate the global historical occurrences of tsunamis. They compare the distribution of the number of annual tsunami events using a Poisson distribution and a Negative binomial distribution. The latter provides a consistent fit to tsunami events whose height is greater than one. They also also investigate the interval distribution using gamma and exponential distributions. The former is found to be a better fit, suggesting that the number of tsunami events is not a Poisson process.

[20] study tsunami events in the USA. They assume that the occurrence frequency of tsunami in each year follow a Poisson distribution. They then identify the distribution of tsunami heights by fitting six distributions: Gumbel, Log normal, Weibull, maximum entropy, and GPD. They select GPD, which had the best fit. They use MLE for parameter estimation, and investigate the fit of the Poisson Compound Extreme Value Distribution using goodness-of-fit statistics. The results is consistent with [19], that the Poisson-Generalised Pareto Distribution is appropriate for disaster modelling.

2. Classical extreme value theory

The core of the classical Extreme value theorem is the study of the stochastic behaviour of some maximum (or minimum) of a sequence of random variables. Define

$$M_n = \max \{Y_1, \dots, Y_n\} \quad (1)$$

where $\{Y_1, \dots, Y_n\}$ is a sequence of independent random variables with a common distribution function F . M_n represents the maxima (minima) of observed process over n blocks or time units. If F is known, the distribution of M_n can be derived as follows:

$$\begin{aligned} P\{M_n \leq y\} &= P\{Y_1 \leq y, \dots, Y_n \leq y\} \\ &= P\{Y_1 \leq y\} \times \dots \times P\{Y_n \leq y\} \\ &= \{F(y)\}^n \end{aligned} \quad (2)$$

However, F is usually unknown in practice, and will have to be estimated from the data. This poses a problem since a small error in the estimation of F can lead to large disparities for $F^n\{y\}$. An alternative approach is to model $F^n\{y\}$ through asymptotic theory of M_n , where we study the behaviour of $F^n\{y\}$ as n tends towards

infinity. Since $F(y) < 1$ for $y < y_{sup}$, where y_{sup} is the upper end-point of F , we have $F^n\{y\} \rightarrow 0$ as $n \rightarrow \infty$. We can remove the degeneracy problem by allowing some linear re-normalisation of M_n . Consider a linear re-normalisation:

$$\hat{M}_n = \frac{M_n - d_n}{c_n} \quad (3)$$

where $\{c_n\}$ and $\{d_n\}$ are sequences of constants with $c_n > 0$. Under a suitable choice of c_n and d_n , the distribution of M_n can be stabilised and which leads to extremal types theorem:

Theorem 1.1 [Extremal Types Theorem] For a non-degenerate distribution function, G , if there exists sequences of constants $\{c_n > 0\}$ and $\{d_n\}$, as $n \rightarrow \infty$, such that

$$P\left(\frac{M_n - d_n}{c_n} \leq y\right) \rightarrow G(y) \quad (4)$$

then G belongs to one of the following families:

Gumbel:

$$G(y) = \exp \left\{ -\exp \left[-\left(\frac{y-d}{c} \right) \right] \right\}, \quad -\infty < y < \infty \quad (5)$$

Frechet:

$$G(y) = \begin{cases} 0 & y \leq d \\ \exp \left\{ -\left(\frac{y-d}{c} \right)^{-\alpha} \right\} & y > d \end{cases} \quad (6)$$

Weibull:

$$G(y) = \begin{cases} \exp \left\{ -\left[-\left(\frac{y-d}{c} \right) \right]^\alpha \right\} & y < d \\ 1 & y \geq d \end{cases} \quad (7)$$

for $c > 0$ and $d \in \mathbb{R}$.

The proof of this theory is presented in [21]. The three classes of distributions are called extreme value distributions, with type I (Gumbel), type II (Frechet) and type III (Weibull) respectively. The extremal types theorem suggests that regardless of the population distribution of M_n , if a non-degenerate limit can be obtained by linear re-normalisation, then the limit distribution will be one of the three extreme value distributions.

In modelling an unknown population distribution, we choose one of the three types of limiting distributions and then estimate the model parameters. This approach is, however, deemed to be inefficient as uncertainty associated with the choice is not considered in the subsequent inference [22]. A better approach is to combine the three models into one single family, with the distributions being special cases of the universal distribution.

2.1 The generalised extreme value distribution

Von Misses [4] and Jenkinson [5] combined the three types of extreme value distributions leading to the generalised extreme value distribution (GEV).

Theorem 1.2 If there exists sequences of constants c_n and d_n , such that

$$P\left(\frac{M_n - d_n}{c_n} \leq y\right) \xrightarrow{n \rightarrow \infty} G(y) \quad (8)$$

where G is a non-degenerate distribution function, then G is a member of the GEV family:

$$G(y) = \exp \left\{ - \left[1 + \zeta \left(\frac{y - \nu}{\sigma} \right) \right]^{-1/\zeta} \right\} \quad (9)$$

defined on y such that $1 + \zeta \left(\frac{y - \nu}{\sigma} \right) > 0$ and with parameters: scale $\sigma > 0$, location $\nu \in \mathbb{R}$ and scale $\zeta \in \mathbb{R}$.

The shape parameter determines the tail behaviour under the same values of location and scale, and thus indicates the type of extreme value distribution:

- If $\zeta = 0$, the GEV becomes a Gumbel distribution and the tail decays exponentially.
- If $\zeta < 0$, the GEV becomes a negative Weibull distribution with the upper endpoint being finite.
- If $\zeta > 0$, the GEV simplifies to a Frechet distribution with a heavy tail and which decays polynomially.

3. Threshold models and the generalised Pareto distribution

Modelling using block maxima is inefficient as it is wasteful of the data, considering that the complete dataset is available. An alternative approach is to model all the data above some high threshold, in what is commonly referred to as threshold modelling or Excess over Threshold (EOT) modelling.

Given a set of independent and identically distributed random variables Z_1, \dots, Z_n , having a common distribution function, F , we are interested in estimating the conditional excess distribution function, F_v , of random variable X above a high threshold v :

$$F_u(y) = P(Z - v \leq y | Z > v), \quad 0 \leq y \leq z_+ - v \quad (10)$$

where $y = z - v$ are the exceedances and z_+ is the right endpoint of F . We can express $F_v(y)$ in terms of z as:

$$F_v(y) = \frac{F(v + y) - F(v)}{1 - F(v)} = \frac{F(v + z - v) - F(v)}{1 - F(v)} = \frac{F(z) - F(v)}{1 - F(v)} \quad (11)$$

Piklands [23] posed that if the underlying distribution $F(z)$ is in the maximum domain of attraction of extreme value distribution, then the conditional excess distribution function $F_v(z)$ for a large v , can be approximated by:

$$F_v(z) \approx F_{(\zeta, \nu, \sigma)}(z), \quad v \rightarrow \infty \quad (12)$$

where $F_{(\zeta, \nu, \sigma)}(z)$ is the Generalised Pareto Distribution (GPD).

3.1 Generalised Pareto distribution (GPD)

Definition 1 The random variable Z has a Generalised Pareto Distribution if the cumulative distribution function of Z is given by:

$$F_{(\zeta, \nu, \sigma)}(z) = \begin{cases} 1 - \left(1 + \frac{\zeta(z - \nu)}{\sigma}\right)^{-1/\zeta} & \text{for } \zeta \neq 0 \\ 1 - \exp\left(-\frac{z - \nu}{\sigma}\right) & \text{for } \zeta = 0 \end{cases} \quad (13)$$

for $z > \nu$ and $\left(1 + \frac{\zeta(z - \nu)}{\sigma}\right) > 0$ and parameters: location $\nu > 0$, scale $\sigma > 0$, shape $\zeta \in \mathbb{R}$.

Remark 1 (Special Cases) Under specific conditions, the GPD simplifies to other continuous distributions:

- When $\zeta = 0$, the GDP simplifies to the exponential distribution
- When $\zeta > 0$, the GDP becomes an ordinary Pareto distribution, and
- when $\zeta < 0$, the GDP simplifies to a Pareto type II distribution.

The mean of the GPD is:

$$E(Z) = \nu + \frac{\sigma}{1 - \zeta}, \text{ for } \zeta < 1 \quad (14)$$

The Variance of the GPD is:

$$\text{Var}(Z) = \nu + \frac{\sigma^2}{(1 - \zeta)^2(1 - 2\zeta)}, \text{ for } \zeta < \frac{1}{2} \quad (15)$$

In general, the r -th moment of the GPD only exists if $\zeta < \frac{1}{r}$.

The shape parameter, ζ , determines the tail distribution of the GPD as indicated in remark 1. When $\zeta = 0$, there exists a decreasing exponential tail, when $\zeta > 0$, there is a heavy tail and when $\zeta < 0$, the tail is short, with finite upper end point $\nu - \frac{\sigma}{\zeta}$.

3.2 Relationship between GEV and GPD

Theorem 1.3 Let Z_1, \dots, Z_n be a sequence of independent and identically distributed random variables with a common cumulative distribution function F , and let $M_n = \max\{Z_1, \dots, Z_n\}$ satisfying the conditions to be approximated by GEV, i.e., for large n :

$$P\{M_N \leq z\} = G(z), \text{ where } G(z) = \exp\left\{-\left[1 + \zeta\left(\frac{z - \nu}{\sigma}\right)\right]^{-1/\zeta}\right\} \quad (16)$$

Then, for a sufficiently high threshold ν , the conditional distribution function of $(Z - \nu)$, conditioned to $Z > \nu$ is approximately given by

$$H(y) = 1 - \left(1 + \frac{\zeta y}{\delta}\right)^{-1/\zeta} \quad (17)$$

where $y = z - v > 0$, $\left(1 + \frac{\zeta y}{\delta}\right) > 0$ and $\delta = \sigma + \zeta(v - \nu)$.

Proof. Denote the distribution function of the random variable Z by F. By theorem 1.2, for large n,

$$\begin{aligned} G(z) = F^n(z) &\approx \exp \left\{ - \left[1 + \zeta \left(\frac{z - \nu}{\sigma} \right) \right]^{-1/\zeta} \right\} \\ n \log F(z) &\approx \left\{ - \left[1 + \zeta \left(\frac{z - \nu}{\sigma} \right) \right]^{-1/\zeta} \right\} \end{aligned} \quad (18)$$

For large values of z, Taylor series expansion implies:

$$\log F(z) \approx - \{1 - F(z)\} \quad (19)$$

Hence,

$$\begin{aligned} n\{1 - F(z)\} &\approx \left[1 + \zeta \left(\frac{z - \nu}{\sigma} \right) \right]^{-1/\zeta} \\ \{1 - F(z)\} &\approx \frac{1}{n} \left[1 + \zeta \left(\frac{z - \nu}{\sigma} \right) \right]^{-1/\zeta} \end{aligned} \quad (20)$$

So, for $z = v + y$

$$\{1 - F(v + y)\} \approx \frac{1}{n} \left[1 + \zeta \left(\frac{v + y - \nu}{\sigma} \right) \right]^{-1/\zeta} \quad (21)$$

Thus,

$$\begin{aligned} P\{Z > v + y | Z > v\} &= \frac{1 - P(Z < v + y)}{1 - P(Z < v)} \\ &\approx \frac{\left[1 + \zeta(v + y - \nu)/\sigma \right]^{-1/\zeta}}{\left[1 + \zeta(v - \nu)/\sigma \right]^{-1/\zeta}} \\ &= \frac{\left[1 + \zeta(v + y - \nu)/\sigma \right]^{-1/\zeta}}{\left[1 + \zeta(v - \nu)/\sigma \right]^{-1/\zeta}} \\ &= \left[\frac{1 + \zeta(v - \nu)/\sigma + \zeta y/\sigma}{1 + \zeta(v - \nu)/\sigma} \right]^{-1/\zeta} \\ &= \left[\frac{1 + \zeta(v - \nu)}{1 + \zeta(v - \nu)} + \frac{\zeta y/\sigma}{1 + \zeta(v - \nu)/\sigma} \right]^{-1/\zeta} \\ &= \left[1 + \frac{\zeta y/\sigma}{\sigma + \zeta(v - \nu)} \right]^{-1/\zeta} \\ &= \left[1 + \frac{\zeta y}{\sigma + \zeta(v - \nu)} \right]^{-1/\zeta} \\ &= \left[1 + \frac{\zeta y}{\delta} \right]^{-1/\zeta} \end{aligned} \quad (22)$$

Therefore,

$$P\{Z \leq v + y | Z > v\} = 1 - \left[1 + \frac{\zeta y}{\delta} \right]^{-1/\zeta} \quad (23)$$

where, $y > 0$, $\left(1 + \frac{\zeta y}{\delta} \right) > 0$ and $\delta = \sigma + \zeta(v - \nu)$.

Eq. (23) is the generalised Pareto family of distributions.

Theorem 1.3 implies that we can use GPD as an approximation to the distribution of maxima using EOT as alternative to GEV in block maxima. We can observe how the parameters of the GPD are uniquely determined by those of the GEV. In particular, the shape parameter is equal in both cases. When the block sizes change in the GEV, the parameters ν and σ change, while the shape parameter remains constant.

Eq. (23) clearly indicates the dependence of the scale parameters σ on the threshold. The threshold choice is an important part of threshold modelling. As with the case with the block sizes in GEV, the choice of threshold is a trade-off between the bias and variance [13]. If the threshold is too low, we violate the asymptotic arguments underlying the GPD, whereas, if the threshold is too high, we will generate few exceedances to estimate the parameters, resulting to a large variance. We now describe the process of selecting the appropriate threshold, to be used in modelling extreme events using GPD.

3.3 Threshold selection

In modelling using EOT, the threshold is usually chosen before the model is fitted. We will present three tools that will help in identifying the threshold:

3.3.1 Mean residual life plot

This method is based on the mean of the GPD: $E(Z) = \frac{\sigma}{1-\zeta}$, when $\zeta < 1$. If the excesses $z - v$ are approximated by a GPD for a high threshold v , the mean excess, for $\zeta < 1$ is:

$$e(v) = E(Z - v | Z > v) = \frac{\sigma_v}{1 - \zeta} \quad (24)$$

For any higher threshold $r > v$:

$$e(r) = E(Z - r | Z > r) = \frac{\sigma_r}{1 - \zeta} = \frac{\sigma_v + \zeta(r - v)}{1 - \zeta} \quad (25)$$

Therefore, the mean excess function $e(r)$ is a linear function of v , once a suitable high threshold has been reached. The sample mean residual life plot is drawn using:

$$\left(v, \frac{1}{n_v} \sum_{i=1}^{n_v} (z_{(i)} - v) \right) \quad (26)$$

where $z_{(i)}$ is the observation, i , above the threshold, v , and n_v is the total number of observations above v . For a high v , all the exceedances, $r > v$, in the mean residual life plot changes linearly with v .

Using this result, the procedure for estimating the threshold is as follows:

- Draw a mean residual life plot using (26)
- Choose a threshold above which the plot is approximately linear. We use confidence interval to determine whether the plot looks linear.

3.3.2 Parameter stability plot

Assuming that the exceedances, $(z - v)$, over a threshold v follow a GPD (ζ, σ_v) , the exceedances will still follow a GPD for any higher threshold $r > v$, with the same ζ , but with scale parameter of:

$$\sigma_r = \sigma_v + \zeta(r - v) \quad (27)$$

(This follows from theory 1.3).

Let us re-parametrise the scale parameter σ_r :

$$\sigma^* = \sigma_r - \zeta_r \quad (28)$$

such that:

$$\sigma_r = \sigma_v + \zeta r - \zeta v \quad (29)$$

$$\sigma_r - \zeta_r = \sigma_v - \zeta v$$

$$\sigma^* = \sigma_v - \zeta v$$

The parameter σ^* now only depends on a sufficiently high threshold v . The parameter stability plot involves plotting the GPD parameter estimates for a range of values v . The threshold is chosen to be the point where the shape and the modified scale parameters become stable (that is, the parameter estimates is constant above the threshold at which the GPD becomes valid). We use confidence interval to select this point.

3.3.3 Gertensgarbe and Werner plot

The test was proposed by Gertensgarbe and Werner (1989) and is used to select a threshold by detecting the starting point of the extreme region. The idea behind the test is that the behaviour of a series of differences that correspond to the extreme observations is different from the one corresponding to the non-extreme observations. So, given a series of differences $\Delta_r = z_{(r)} - z_{(r+1)}, i = 2, 3, \dots, n$, of the order statistics, $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(n)}$, the starting point of the extreme region, and hence the threshold, will be the point at which the series of differences exhibit a significant statistical change.

To detect this point, we apply a sequential version of the Mann-Kendall test to check the null hypothesis that there is no change in the series of differences. Define a series:

$$U_r = \frac{U_r^* - \frac{r(r-1)}{4}}{\sqrt{\frac{i(r-1)(i+5)}{72}}} \quad (30)$$

where $U_r^* = \sum_{j=1}^r n_j$ and n_j is the number of values in the the series of differences $\Delta_1, \dots, \Delta_j$ less than Δ_j . We also define another series, U_p , by applying the same procedure to the series of differences from the end to start, $\Delta_n, \dots, \Delta_1$,

rather than from the start to the end. The starting point of change in the series of differences, and hence the threshold, is the the intersection point of the series U_r and U_p .

4. Data analysis

We use data for all the natural disasters that has been recorded in Kenya in the period 1964 – 2018. The data is obtained from CRED database, which is the currently the most comprehensive database for natural events. The data was also cross-referenced with that from other sources including UN-agencies and NDMU. The impact of natural disasters is quantified in terms of the total number of people affected on an annual basis, which we deemed to be more reliable than the the total damage in monetary terms. The total number of people affected includes those who were injured, died, left homeless or affected in any other way by natural disasters. A total of 112 events have been recorded over that period of time with a resulting damage of approximately 62 million people affected.

Descriptive statistics for both the annual occurrence and the impacts are provided in **Table 1**. The minimum number of disasters and the resulting impact is zero, which corresponds to those years where no natural disasters occurred. A total of 22 years recorded no natural disaster events. The average annual number of natural disaster occurrences in Kenya is two, and about 1.3 million people are affected every year. The maximum number of natural disaster occurrences observed in any year is 9 and the worst disaster recorded in any year affected approximately 23, 331, 469 people. The mean is greater than the median for both variables, indicating that the data is right-skewed. We can also observe that the spread of the impacts is large as suggested by both the standard deviation and the inter-quantile range (about 3.5 million and 252718 respectively), as opposed to that of the occurrence.

We are interested in the distribution of the number of occurences and the corresponding impact of natural disasters in Kenya in the period of study. **Figure 1** shows the distribution of the number of natural disaster occurrences in the last 55 years. The number of natural disasters between 1964 and the late 1990s was fairly low, with no year experiencing more than two events. We can then observe a sharp

Statistic	Annual Occurrence	Annual Impact
N	55	55
Total	112	62,160,910
Mean	2	1,130,198
Std. Dev	2.26	3,494,302
Minimum	0	0
st Quantile	0	0
Median	1	15000
rd Quantile	4	252718
Maximum	9	23,331,469
Inter Quantile Range	4	252,718

Table 1.
Descriptive Statistics.

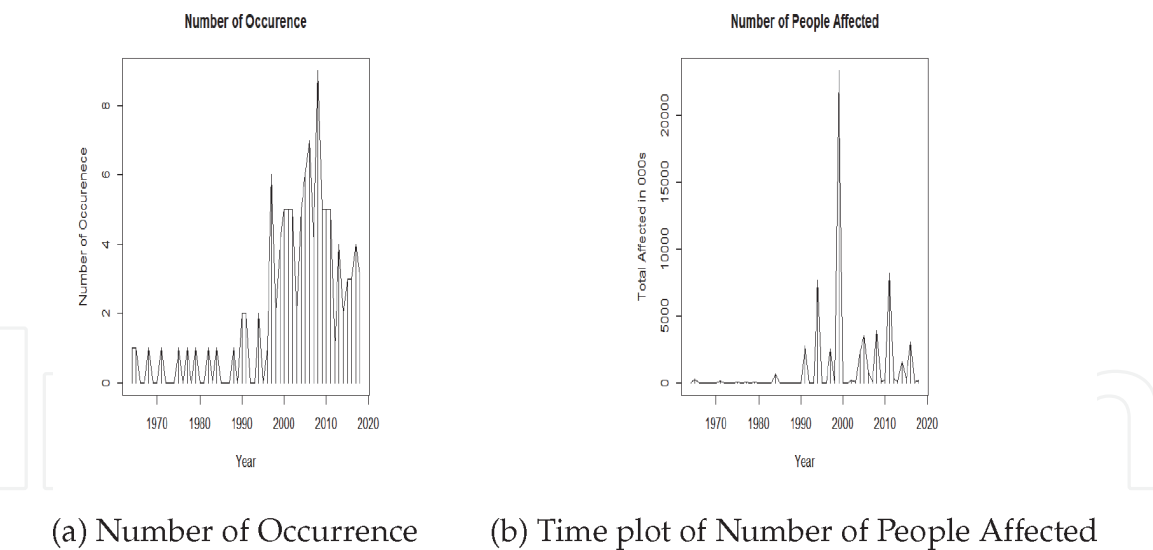


Figure 1.
Distribution of Natural Disasters in Kenya in the period: 1964–2018.

increase in the annual occurrences in the last 20 years. During this period, all the years have experienced at least two events.

A similar trend is observed for the impacts of the natural disasters, as shown in **Figure 1b**. Very few people were being affected in the years between 1964 and 1990. Then from 1990, the severity of the natural disasters increased and a lot more people were being affected.

4.1 Check for Independence assumption

We assumed that the annual occurrences are independent of each other as with is the case with annual impact. We also assumed that the occurrences and impact are independent of each other. We can observe that the scatter plots 2,3 and 4 indicate that there is no serious violation of these assumptions (**Figures 2-4**).

4.2 Detection of heavy-tailed behaviour

The statistical methods used in this project rely on the heavy-tailedness of the underlying distribution. We observed that the data is right skewed as the mean is

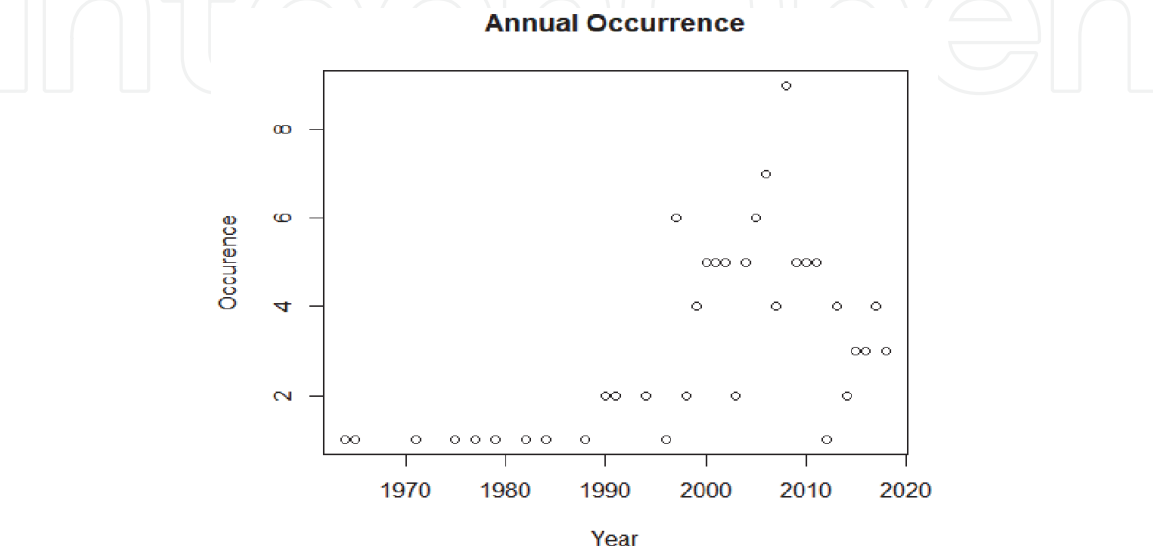


Figure 2.
Scatter plot of Annual Occurrence.

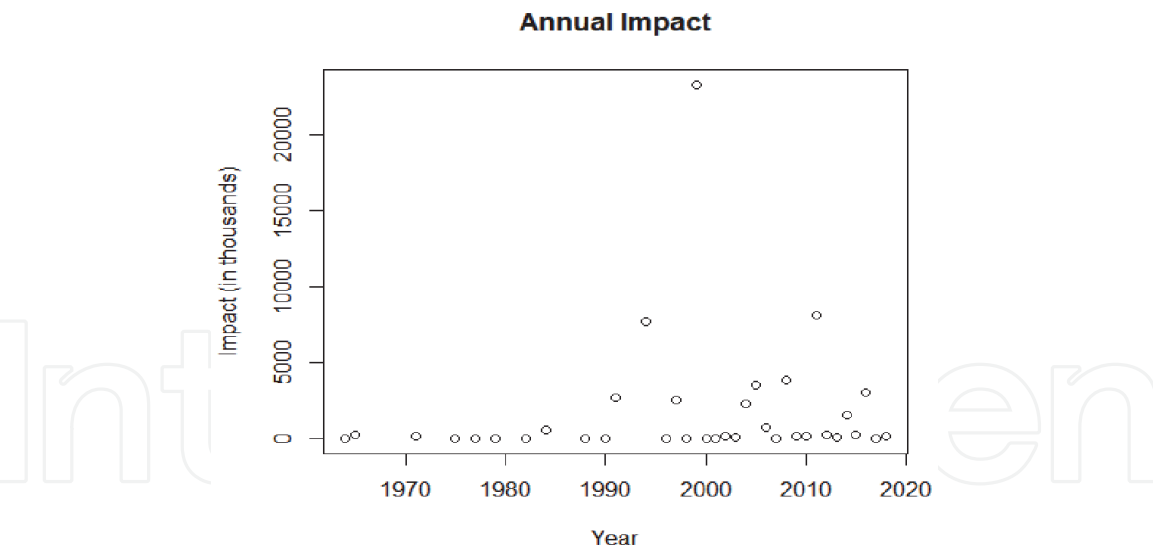


Figure 3.
Scatter plot of Annual Impact.

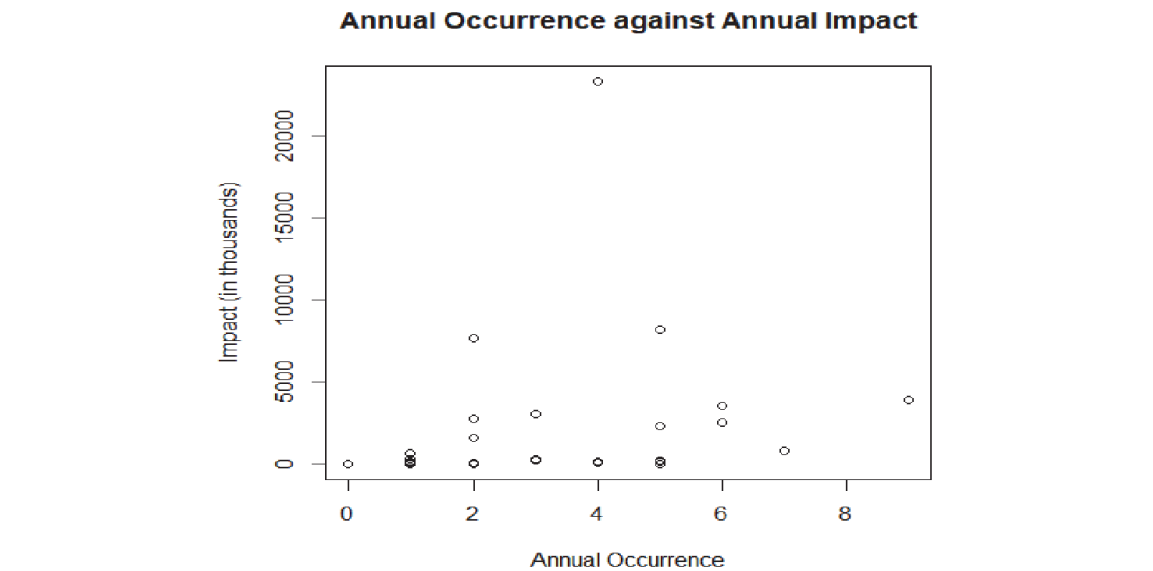


Figure 4.
Scatter plot of Annual Occurrence against Annual Impact.

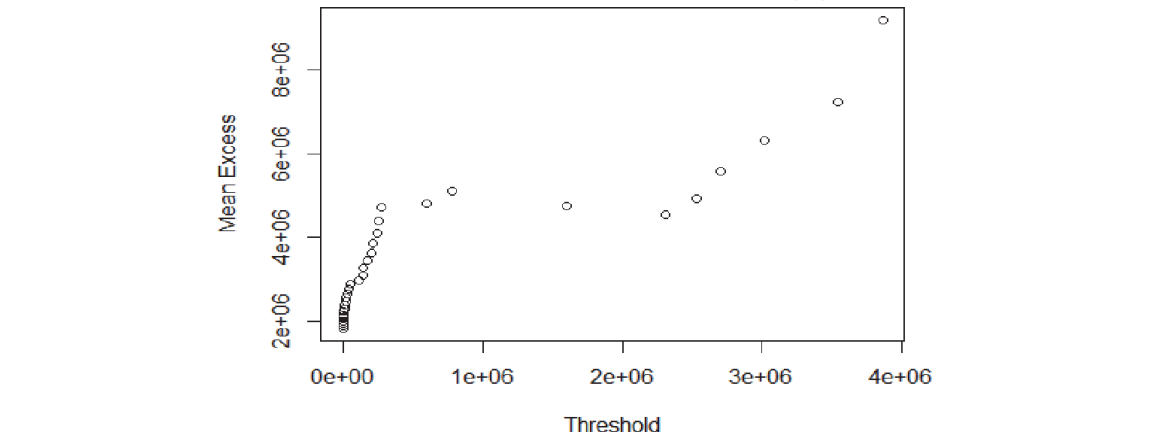


Figure 5.
Mean Excess Plot.

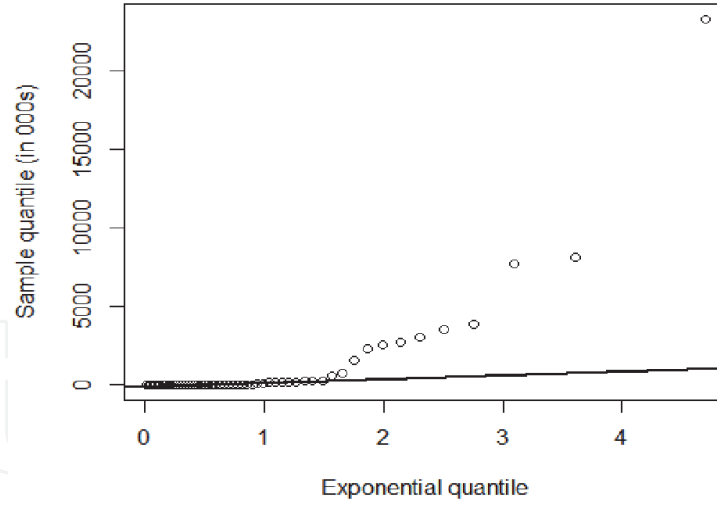


Figure 6.
Exponential Q-Q plot.

greater than the median (**Table 1**), indicating that the underlying distribution has a long tail. We will further investigate this by plotting the empirical mean excess function against different threshold values. We test the assumption using an empirical mean excess plot.

An upward trend in the mean excess plot indicates heavy-tailed behaviour, as explained in Section 4.2. From **Figure 5**, we can observe a general upwards trend in the graph, except for the area between one million and two million.

To get more conclusive results, we also plot an exponential Q-Q and observe the pattern of the points in relation to the straight line. Heavy-tailed behaviour will be indicated by a convex departure from the straight line as explained in Section 4.2. Shorter tailed-distribution will have a concave departure and if the data are a sample from an exponential distribution, the points should be approximately linear.

We can observe the convex behaviour of the exponential Q-Q plot in **Figure 6**. Thus, we can conclude that the underlying distribution is heavy-tailed.

5. Threshold selection

Having proved that the underlying distribution is heavy-tailed, we now select a threshold, above which the distribution of the impact is approximated by the GPD. We will use the techniques provided in Section 3.3 to estimate the appropriate threshold. First, we plot the mean excesses for each value of 200 different thresholds across the whole dataset, against their corresponding thresholds, with a significance level of 0.05. The goal is to find the lowest threshold such that the graph is linear with increasing thresholds, within uncertainty.

The mean residual life plot in **Figure 7** looks like a straight line right from the beginning except at the very right end. The plot suggests that the threshold should be between zero and around 250, 000.

The threshold selection can be carried out more carefully through a parameter stability plot. Recall that if the threshold is too low, the assumption underlying the GPD will be violated. On the other hand, if the threshold is too high, we will have too few observations to effectively fit the distribution. Given that natural disasters can be regarded, as extreme and rare events, we will limit the range of the threshold to within the upper and lower quantiles of the impact dataset: $0 \leq v \leq 250,000$, which is surprisingly in alignment to the values indicated by the mean excess plot. The resulting range of the number of exceedance is then 13 and 42. **Figure 8** shows

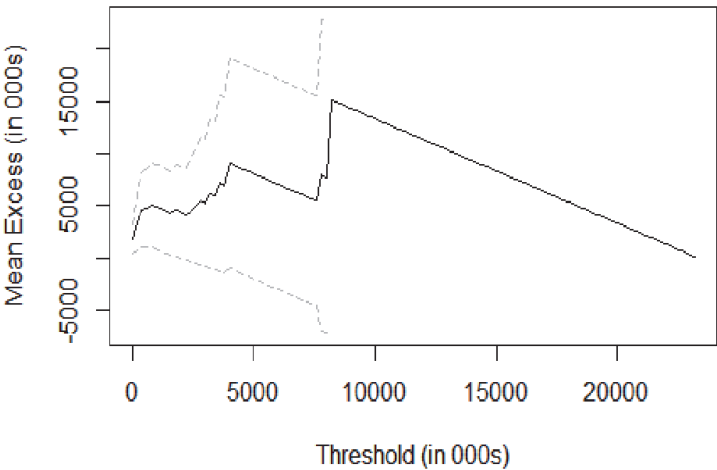


Figure 7.
Mean residual Life Plot.

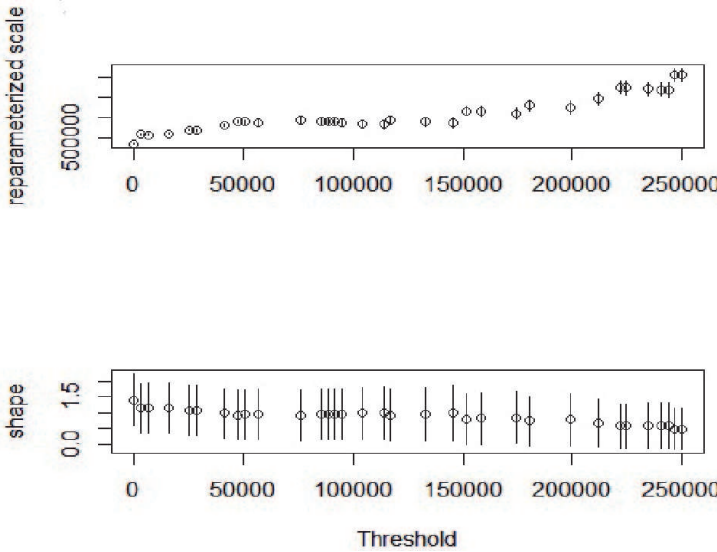


Figure 8.
Parameter Stability Plot.

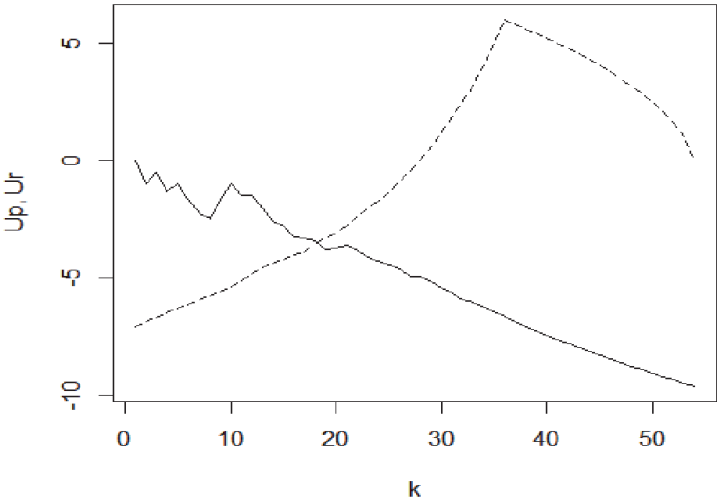


Figure 9.
Gertensgarbe Plot.

Statistic	Value
k_0	19
p-value	0.00014
threshold	150000
tail index	0.48507

Table 2.
Mann-Kendall Test Results.

the plot the MLE estimates for the parameters of the GPD against their 80 corresponding thresholds, within the range $0 \leq v \leq 250,000$, together with a 95% confidence intervals.

We can tell that the estimates of the shape parameter appear to be constant for the whole dataset (when the threshold, v , is zero). The re-parametrized scale parameter estimates on the other hand, seem to be stable beyond 50,000 but not so beyond 225,000.

The information provided by these plots can however, be rather approximative. The Gertensgarbe plot provides a more powerful procedure for threshold estimation [24].

The cross point of the Gertensgarbe graph (**Figure 9**) is at the observation numbered $k = 19$, which corresponds to a threshold of 150,000. The null hypothesis

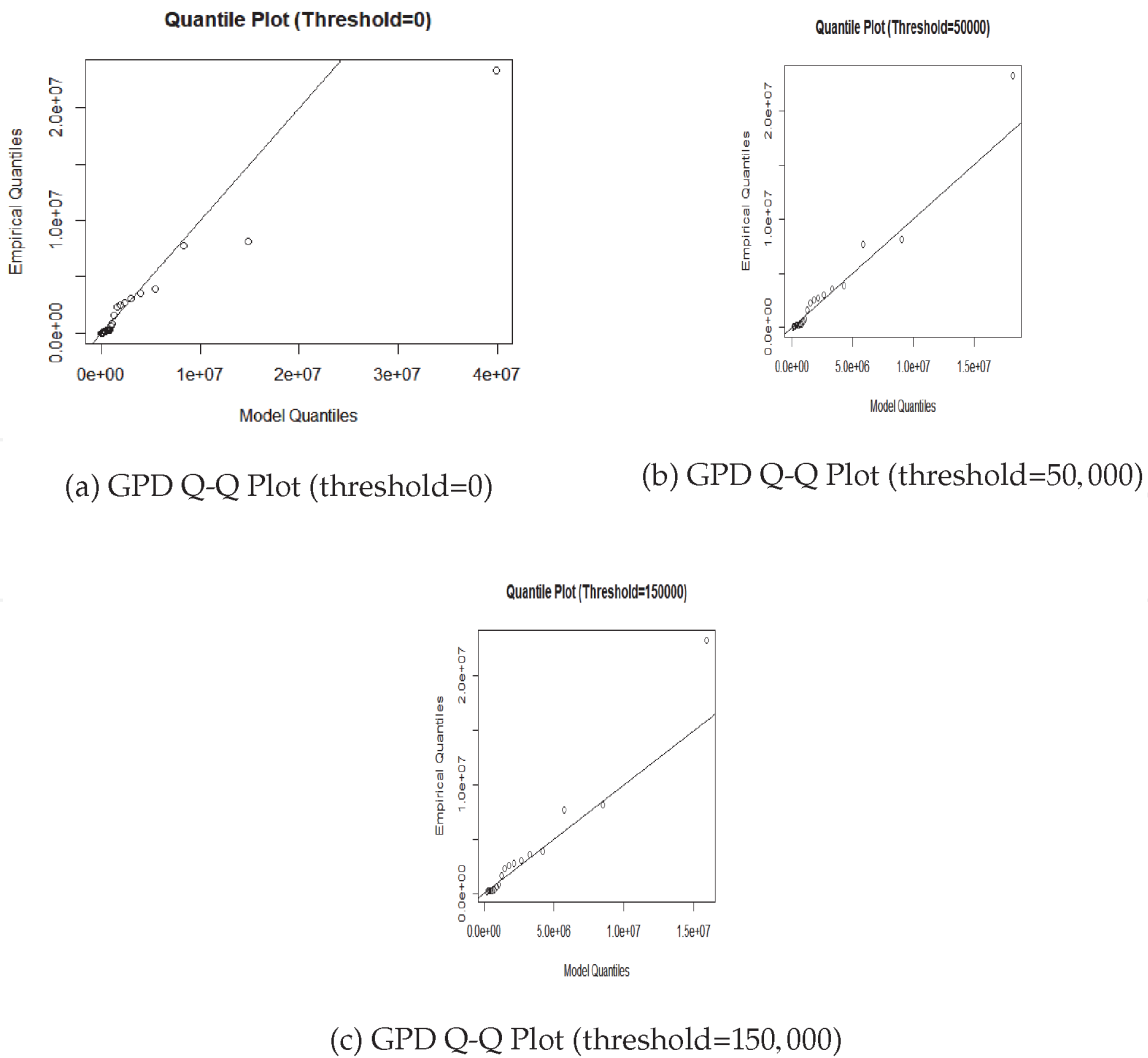


Figure 10.
Q-Q Plot of GPD under different thresholds.

that there is no change in the series of differences is rejected with a p-value less than 0.001 (see **Table 2**). Therefore, all the three techniques detect the threshold to be between $0 \leq v \leq 250,000$. The parameter stability plot estimates the threshold to be 50,000 and the Gertensgarbe plot 150,000. We will investigate the goodness-of-fit of the GPD to each of the three cases.

Figure 10 shows that GPD fits the data best when we set the threshold at 50,000 and 150,000. The differences between the fit in the two cases also appears to be very small. Since we want to have as many exceedances as possible, we choose the threshold to be 50,000.

6. Distribution of the exceedances

Given a threshold of 50,000, we are interested in the distribution of the excesses. The number of exceedances is 22, and **Figure 11** shows the plot of the kernel density estimates of the number and value of the exceedances.

The number of exceedances is assumed to be Negative-binomial-distributed. We carry out a chi-squared test using the MLE estimates to test its goodness of fit:

Table 3 shows the output of the chi-squared test carried out using “fitdistrplus” package in R. The p-value is greater than 0.01 hence, we fail to reject the null hypothesis that the data follows a Negative binomial distribution.

In section 3, we saw that the distribution of the exceedances can be approximated by a GPD. We will test whether this theorem is justified in our dataset. We use the “bootstrap goodness-of-fit test for the GPD” [25] provided in R package “gPdtest. This test investigates the goodness-of-fit of the GPD, for cases where the distribution is heavy-tailed (shape parameter $\zeta \geq 0$) and non-heavy tailed ($\zeta < 0$).

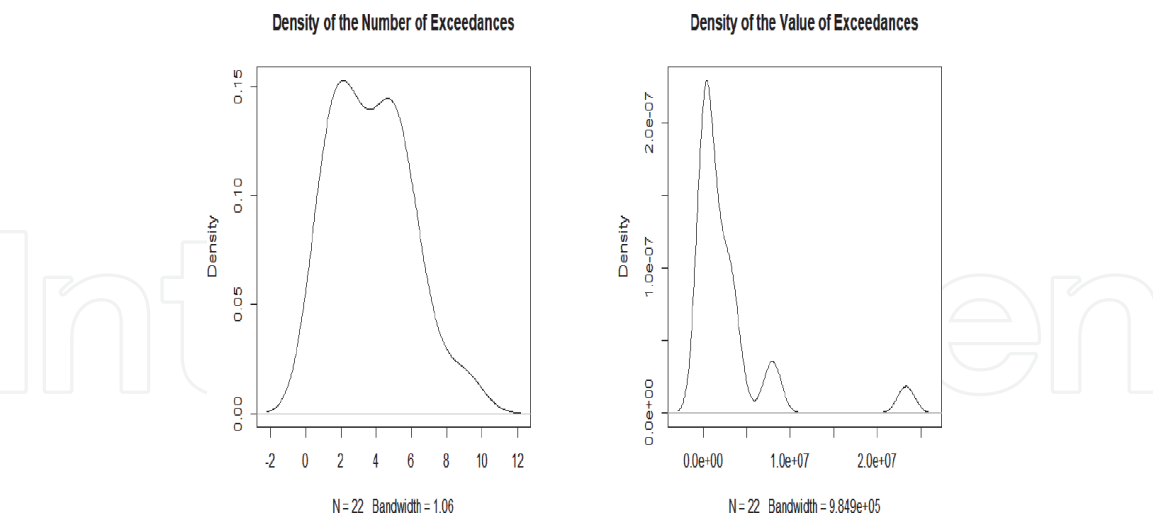


Figure 11.
Density of the Exceedances.

Name	Value
Ch-squared statistic	3.01
Degree of Freedom	2.00
Chi-squared p-value	0.22

Table 3.
Goodness of fit of Negative Binomial to the distribution of the number of Exceedances.

Item	P-value	R-statistic
H_0^- : X has a GPD with negative shape parameter	0.0961	0.9581
H_0^+ :X has a GPD with positive shape parameter	0.3874	0.9720

Table 4.
Goodness of fit test for GPD.

Table 4 shows that the p-value in both cases is greater than 0.01. Hence, we conclude that the exceedances do indeed follow a GPD.

7. Fitting the negative binomial-generalised Pareto compound extreme value distribution

7.1 Parameter estimation

We now fit the proposed distribution to the Kenyan data on natural disasters using MLE. We derived the log-likelihood function of the NB-GP CEVD in Section 7.1, and found that the function is not in closed form. Therefore, we will use iterative methods to estimate the parameters.

First, we have to select an appropriate initial value for the parameter estimates. These values represents the initial guess for the estimates. The parameters of the NB-GP CEVD are supported at $\alpha > 0, \sigma > 0, 0 < \theta < 1, 0 < a < 1$ and $\zeta \in \mathbb{R}$. We will study how the choice of the starting value affects the parameter estimates by choosing different starting points, within the support of the parameters.

From **Table 5**, we can observe how the starting values of the parameter estimates significantly affects the estimates. To get more reliable estimates, we use PWM estimates, as discussed in subsection X, to be the starting values. The fist five sample PWMs are found to be:

$\hat{p}_0 = 2817753, \hat{p}_1 = 2440020, \hat{p}_2 = 2185841, \hat{p}_3 = 1997587, \hat{p}_4 = 1851745 \tag{31}$

The resulting system of equations is then:

$$\begin{aligned} \frac{\sigma}{\zeta} \left\{ \alpha \left(\frac{a\theta}{1-\theta} \right)^\zeta B((1-\zeta), (\alpha+\zeta)) - 1.0 \right\} &= 2817753 \\ \frac{\sigma}{\zeta} \left\{ \alpha \left(\frac{a\theta}{1-\theta} \right)^\zeta B((1-\zeta), (2\alpha+\zeta)) - 0.5 \right\} &= 2440020 \end{aligned} \tag{32}$$

Parameter	α	θ	σ	ζ	a
Starting Values	4.5	0.1	1	0.5	0.4
Estimates	24.2281	0.9228	0.9757	0.5039	0.3900
Starting Values	0.5	0.1	1	0.5	0.4
Estimates	3.3001	0.9898	1.8597	0.7754	0.9016
Starting Values	0.1	0.1	0.1	0.1	0.1
Estimates	0.5632	0.9781	0.2440	0.2226	0.3991

Table 5.
MLE estimates of the parameters using different starting values.

$$\begin{aligned} \frac{\sigma}{\zeta} \left\{ \alpha \left(\frac{a\theta}{1-\theta} \right)^\zeta B((1-\zeta), (3\alpha+\zeta)) - 0.33 \right\} &= 2185841 \\ \frac{\sigma}{\zeta} \left\{ \alpha \left(\frac{a\theta}{1-\theta} \right)^\zeta B((1-\zeta), (4\alpha+\zeta)) - 0.25 \right\} &= 1997587 \\ \frac{\sigma}{\zeta} \left\{ \alpha \left(\frac{a\theta}{1-\theta} \right)^\zeta B((1-\zeta), (5\alpha+\zeta)) - 0.2 \right\} &= 1851745 \end{aligned} \tag{33}$$

We use “nleqslv” package in R to solve the resulting system of non-linear equations:

Using the values in **Table 6** as the initial values, the MLE estimates of the NB-GP CEVD distribution is [26]:

The likelihood function is also found to be -192.9873 .

7.2 Goodness-of-fit tests

Next, we test the goodness-of-fit of the NB-GP CEVD to the data. We create a sample of NB-GP CEVD with the parameters given in **Table 7**. Since this sample is the basis of comparison, we create a large sample of size 10,000 to closely approximate the distribution. Two-sample KS test and two-sampled AD test is then carried out to test the null hypothesis that the two samples have the same distribution.

As observed in **Table 8**, the p-values in both tests are greater than 0.01. Thus, we fail to reject the null hypothesis that the two samples come from the same population at 1% level of significance. Alternatively, for the KS-test, we have:

$$d(\alpha) \sqrt{\frac{n+m}{nm}} = \sqrt{-0.5 \ln(0.1)} \sqrt{\frac{10022}{220000}} = 0.3239 \tag{34}$$

Parameter	Estimate
$\hat{\alpha}$	9,000,000
$\hat{\theta}$	0.8509
$\hat{\sigma}$	0.2297
$\hat{\zeta}$	0.8378
\hat{a}	0.3836

Table 6.
Probability-weighted Moments Estimates.

Parameter	Estimate	Standard Error
$\hat{\alpha}$	9,000,000	2.9659
$\hat{\theta}$	0.0000178	—
$\hat{\sigma}$	1268.1920	0.0100
$\hat{\zeta}$	1.2336	0.0111
\hat{a}	0.4000	0.0341

Table 7.
Maximum Likelihood Estimates.

Test	statistic	p-value
Kolmogorov–Smirnov	0.3010	0.0375
Anderson-Darling	2.8140	0.0346

Table 8.
Goodness-of-fit Test results.

The test-statistic

$$D_{m,n} = 0.3010 < 0.3239 \tag{35}$$

so we fail to reject the null hypothesis. We conclude that the data of natural disasters in Kenya follow a NB-GP CEVD.

7.3 Prediction study

From the estimation results in **Table 7**, we can tell that the probability of damages exceeding 50,000 in any year is approximately 0.4. We now want to predict the expected maximum and minimum among those exceeding the threshold in any year with different certainties as explained in section X.

Using Eq. (X), the minimum damage that is expected to be exceeded among those exceeding the threshold with probability p, is predicted using:

$$\hat{x}_p = 50000 + \frac{\hat{\sigma}}{\hat{\zeta}} \left\{ \left[\frac{\hat{a}\hat{\theta}}{(1-\hat{\theta})((1-p)^{-1/\hat{\alpha}} - 1)} \right]^{\hat{\zeta}} - 1 \right\} \tag{36}$$

The corresponding minimum equation, x_p , among those exceeding 50000 that is expected not to be exceeded at any given year under different probabilities can be predicted using:

$$\hat{x}_p = 50000 + \left\{ \frac{\hat{\sigma}}{\hat{\zeta}} \left[\frac{\hat{a}\hat{\theta}}{(1-\hat{\theta})(p^{-1/\hat{\alpha}} - 1)} \right]^{\hat{\zeta}} - 1 \right\} \tag{37}$$

Using the parameter estimates in **Table 7**, the prediction results for $p = 0.9, 0.95, 0.99, 0.995, 0.999$ are given in **Table 9**.

The predicted maximum damage increases as the probabilities increase while the predicted minimum damage decreases as the probabilities increase. In other words, the range becomes larger at higher certainty levels. We can determine the accuracy of the prediction by comparing the prediction results to the statistics of the exceedances (**Table 10**). This can be done using two measures:

- 1. Absolute Error = Predicted Value - Actual Value
- 2. Relative Error = $\frac{\text{AbsoluteError}}{\text{ActualValue}}$

From **Table 11**, all the error measurements for the minimum values are positive, suggesting that the predicted values are higher than the actual value. This implies that the NB-GP CEVD tends to overestimate the minimum damage among the exceedances in any given year. We can also observe how the relative minimum error for the the minimum damage decreases as the probability increases.

Probability	0.9	0.95	0.99	0.995	0.999
Predicted Maximum Value	2,788,420.00	6,689,195.00	49,35,2186.00	116,056,448.00	842,007,196.00
Predicted Minimum Value	110,633.00	93,581.66	75,257.41	71,093.15	64,934.09

Table 9.
Prediction Results for Extreme Damages in Kenya.

Statistic	Value
Minimum	60,340
Median	692,142
Mean	2,817,753
Maximum	23,331,469

Table 10.
Statistics of the Exceedances.

Probability	0.9	0.95	0.99	0.995	0.999
Absolute Error for the Minimum	50,293.000	33,241.663	14,917.407	10,753.149	4,594.092
Relative Error for the Minimum	0.833	0.551	0.247	0.178	0.076
Absolute Error for the Maximum	−20,543,049	−16,642,274	26,020,717	92,724,979	818,675,727
Relative Error for the Maximum	−0.880	−0.713	1.115	3.974	35.089

Table 11.
Prediction Accuracy for Extreme Damages.

The Relative error is highest at 0,9 level of certainty and lowest at 0.999 certainty level. So, accuracy of the predictions increases as the level of certainty increases. Generally, the distribution performs well at predicting the minimum values of those exceeding the threshold.

For the maximum values, the predicted values are lower than the actual values at 0.9 and 0.95 probabilities, and higher at $p \geq 0.99$. So, the proposed distribution underestimates the maximum values at probability levels less than 0.95, and over-estimates them at probability levels greater than 0.95. In addition, the relative maximum errors are smaller when the predicted values are lower than the actual ones, as compared to when they are greater. So, the distribution overestimates the maximum values by a greater margin than it tends to underestimate. IN general, the NB-GP CEVD performs poorly at predicting the maximum values of the damages exceeding the threshold.

In overall, the absolute errors are smaller for the minimum values than the maximum values. This implies that the NB-GPD CEVD performs better at predicting the minimum values among those exceeding the threshold, than it does for the maximum values.

8. Conclusions

The main goal for this research is to model the extreme damages in Kenya resulting from natural disasters by considering both occurrence and the magnitude

of the events. The magnitude of the natural disasters was quantified in terms of the total number of people affected. We use Excess over Threshold modelling, where the extreme damages are identified as the points beyond a sufficiently high threshold. The choice of the threshold significantly affects the parameter estimates, so we use three methods to identify the threshold: mean excess plot, parameter stability plot and Gertensgarbe and Werner plot. The threshold is found to be at 50,000.

In order to capture both the frequency and magnitude of the natural disasters, we use Compound Extreme Value Distribution that was proposed by Liu et.al in 1980. We identify the distribution of the number of exceedances to be Negative Binomial Distribution, while that for the magnitude of the exceedances is approximated by a GPD. We estimate the parameters of the CEVD using maximum likelihood estimation. The log-likelihood function is not in closed form, so we use PWM to determine the starting value for the iterations. We then carry out goodness-of-fit tests using a two-sample Kolmogorov-Smirnov test and a two-sample Anderson-Darling Test. We find that the NB-GP CEVD is a good fit for the data.

Finally, we carry out a prediction study for the NB-GP CEVD. The probability of exceeding the threshold in any year is found to be 0.4. We also predict, with different levels of certainty, the minimum and maximum values expected to be exceeded and not to be exceeded, respectively in the event of the occurrence of an extreme damage. We study the accuracy of these predictions by comparing the predicted values to the actual maximum and minimum. We find that the proposed distribution tends to overestimate the minimum damage at all the certainty levels. On the other hand, the distribution tends to underestimate the maximum damage at $p < 0.95$ certainty levels, and overestimate them at $p \geq 0.99$ levels. Generally, we find that the proposed distribution performs better at predicting the minimum damages among those exceeding the threshold, than it does for the corresponding maximum damages.

We conclude that the NB-GP CEVD is a good fit for the distribution of the extreme damages resulting from natural disaster in Kenya. It performs better at predicting the minimum value that is expected to be exceeded by extreme damages as compared to maximum damages expected not to be exceeded.

Acknowledgements

The authors would like to thank the School of Mathematics, University of Nairobi for providing a conducive environment to conduct this research.

Conflict of interest

The authors declare no conflict of interest.

IntechOpen

Author details

Carolyne Ogutu^{*,†} and Antony Rono[†]
University of Nairobi, Nairobi, Kenya

*Address all correspondence to: cogutu@uonbi.ac.ke

† These authors contributed equally.

IntechOpen

© 2021 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] E. Montz, A. Tobin and R. Hagelman, (2017). Natural hazards: explanation and integration
- [2] Zlateva and Velez (2016). *A Method for Risk Assessment from Natural Disasters Using an Actuarial Model*, Journal of Economic, Business and Management, Vol. 4.
- [3] P. Embrechts, C. Klüppelberg and T. Mikosch (2013), *Modelling extremal events: for insurance and finance*, vol 33, Springer.
- [4] R. von Mises, (1954). La distribution de la plus grande de n valeurs. *American Mathematical Society: Providence RI II*, 271–294.
- [5] A. F. Jenkinson (1955). The frequency distribution of the annual maximum (or minimum) value of meteorological events. *Quarterly Journal of the Royal Meteorological Society* 81 (348), 158–172.
- [6] A. Ferreira and H. De (2015), *On the block maxima method in extreme value theory: PWM estimators*, The Annals of Statistics, volume 43, number 1, pages 276–298
- [7] Y. Hu (2013), *Extreme Value Mixture Modelling with Simulation Study and Applications in Finance and Insurance*, University of Canterbury.
- [8] K. Engeland, H. Hisdal, and A. Frigessi (2004). Practical extreme value modelling of hydrological floods and droughts: a case study. *Extremes*, 7(1), 5–30.
- [9] H. Rust, D. Maraun and T. Osborn (2009), *Modelling seasonality in extreme precipitation*, The European Physical Journal Special Topics, Volume 174, number 1, page 99–111.
- [10] C. Cunnane, (1973). *A particular comparison of annual maxima and partial duration series methods of flood frequency prediction*, Journal of Hydrology, Volume 18, number 3–4, pages 257–271.
- [11] H. Madsen, C. Pearson and D. Rosbjerg (1997), *Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 2. Regional modeling*, Water Resources Research, volume 33, number 4, pages 759–769.
- [12] S. Caires (2009), *A comparative simulation study of the annual maxima and the peaks-over-threshold methods*, Deltares report 1200264–002 for Rijkswaterstaat, Waterdienst.
- [13] S. Coles, J. Bawa, L. Trenner, and P. Dorazio, (2001). *An introduction to statistical modeling of extreme values* (Vol. 208). Springer.
- [14] A. Frigessi, O. Haug and H. Rue (2002), *A dynamic mixture model for unsupervised tail estimation without threshold selection*, Extremes, volume 5, number 3, pages 219–235.
- [15] C. Behrens, H. Lopes and D. Gamerman (2004), *Bayesian analysis of extreme events with threshold estimation*, Statistical Modelling, volume 4, number 3, pages 227–244.
- [16] T. F. Liu, and F. S. Ma(1980). Prediction of extreme wave heights and wind velocities. *Journal of the Waterway, Port, Coastal and Ocean Division*, 106 (4), 469–479.
- [17] D. Liu, L. Wang and L. Pang (2006), *Theory of multivariate compound extreme value distribution and its application to extreme sea state prediction*, Chinese Science Bulletin, volume 51, number 23, pages 2926–2930.

[18] D. Liu, S. Wen and L. Wang (2002), *Poisson-Gumbel mixed compound distribution and its application*, Chinese Science Bulletin, volume 47, number 22, pages 1901–1906.

[19] E.L. Geist and T. Parsons (2011), Assessing historical rate changes in global tsunami occurrence, *Geophysical Journal International*, 187(1):497–509.

[20] S. Dong, J. and Zhai, S. Tao. (2017), Long-term statistics of extreme tsunami height at Crescent City, *Journal of Ocean University of China*, 16(3):437–46.

[21] M.R. Leadbetter, G. Lindgren, and H. Rootzén (1983), *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics [Internet]. Springer New York.

[22] E. Omey, F. Mallor, and E. Nualart (2009). *An introduction to statistical modelling of extreme values*. Application to calculate extreme wind speeds.

[23] J. Pickands III et al (1975), *Statistical inference using extreme order statistics*, The Annals of Statistics, volume 3, number 1, pages 119–131.

[24] Cebrián, A. C., Denuit, M., & Lambert, P. (2003). Generalized pareto fit to the society of actuaries' large claims database. *North American Actuarial Journal*, 7(3), 18–36.

[25] Villaseñor-Alva, J. A., & González-Estrada, E. (2009). A bootstrap goodness of fit test for the generalized pareto distribution. *Computational Statistics & Data Analysis*, 53(11), 3835–3841.

[26] J. Liu, D. Shi, and X. Wu (2008). Estimation of poisson-generalized pareto compound extreme value distribution by probability-weighted moments and empirical analysis. *Transactions of Tianjin University*, 14(1), 50–54.