

# Optimal Threshold Determination for the Maximum Product of Spacing Methodology with Ties for Extreme Events

Peter Murage<sup>1</sup>, Joseph Mung'atu<sup>2</sup>, Everlyne Odero<sup>3</sup>

<sup>1</sup>Department of Mathematics, Masinde Muliro University of Science and Technology, Kakamega, Western Province, Kenya

<sup>2</sup>Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

<sup>3</sup>Mainde Muliro University of Science and Technology, Kakamega Western Province, Kenya

Email: gitmurage@gmail.com, j.mungatu@fsc.jkuat.ac.ke, eodero@mmust.ac.ke

**How to cite this paper:** Murage, P., Mung'atu, J. and Odero, E. (2019) Optimal Threshold Determination for the Maximum Product of Spacing Methodology with Ties for Extreme Events. *Open Journal of Modelling and Simulation*, 7, 149-168.  
<https://doi.org/10.4236/ojmsi.2019.73008>

**Received:** April 23, 2019

**Accepted:** June 1, 2019

**Published:** June 4, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Extreme events are defined as values of the event below or above a certain value called threshold. A well chosen threshold helps to identify the extreme levels. Several methods have been used to determine threshold so as to analyze and model extreme events. One of the most successful methods is the maximum product of spacing (MPS). However, there is a problem encountered while modeling data through this method in that the method breaks down when there is a tie in the exceedances. This study offers a solution to model data even if it contains ties. To do so, an optimal threshold that gives more optimal parameters for extreme events, was determined. The study achieved its main objective by deriving a method that improved MPS method for determining an optimal threshold for extreme values in a data set containing ties, estimated the Generalized Pareto Distribution (GPD) parameters for the optimal threshold derived and compared these GPD parameters with GPD parameters determined through the standard MPS model. The study improved maximum product of spacing method and used Generalized Pareto Distribution (GPD) and Peak over threshold (POT) methods as the basis of identifying extreme values. This study will help the statisticians in different sectors of our economy to model extreme events involving ties. To statisticians, the structure of the extreme levels which exist in the tails of the ordinary distributions is very important in analyzing, predicting and forecasting the likelihood of an occurrence of the extreme event.

## Keywords

Extreme Value Theory (EVT), Maximum Product of Spacing MPS, Generalized Pareto Distribution (GPD), Peak Over Threshold (POT)

## 1. Introduction

Certain values in the tails of any distribution, represent extreme events and they are pointers to eventuality. The values in the tails are rare, few, but can have great impact on the conclusion arrived at by the analysts. Different sectors of our life experience extreme events and here we mention just but a few. According to [1] and [2], extremely low production in agriculture results to famine if the agriculture depends on rainfall. This means that the amount of rain experienced in that region was too low that crops dried up [3] or very high rainfall that it destroyed all crops that had been planted. [4] studying extreme rainfall in mountainous region and [5] studying extreme rainfall in west Africa did observe that, how low or high the amount of rainfall depends on the threshold attached to the rainfall in that region. In insurance industries [6], while discussing tools in finance and insurance, noted that extreme high claims by the customers can be very dangerous for the company while extreme low claims by the customers can be very beneficial for the company's profit. This means that there is a critical level that the insurance company would wish it is not surpassed and if it is, according to [7], it must be prepared for this eventuality. Very high emissions of the waste product from the manufacturing industries is detrimental to the environment and ozone layer. However, countries must continue to industrialize or expand their industries for economic prosperity. A certain level of emissions must not be exceeded otherwise the environment and ozone layer would be destroyed. The critical value for which if exceeded, an eventuality occurs is called a threshold. The events beyond this threshold are called extreme events and they happen to be at the tails of the distribution. Extreme value theory (EVT) is a tool which attempts to provide us with the best possible estimate of the tail area of the distribution. In [8] work on the importance of tail dependence in Bivariate frequency analysis, there are two principal kinds of model for extreme values. The oldest group of models is the block maxima models; these are models for the largest observations collected from large samples of identically distributed observations. According to [9] and [10], the block maxima/minima methods are fitted with the generalized extreme value (GEV) distribution. A more modern group of models is the peaks-over-threshold (POT) models; these are models for all large observations which exceed a high threshold. The POT models are generally considered to be the most useful for practical applications, due to a number of reasons. First, by taking all exceedances over a suitably high threshold into account, they use the data more efficiently and second, they are easily extended to situations where one wants to study how the extreme levels of a variable  $Y$  depend on some other variable  $X$  for instance,  $Y$  may be the level of tropospheric ozone on a particular day and  $X$  a vector of meteorological variables for that day. This kind of problem is almost impossible to handle through the annual maxima/minima method. POT methods are used where the exceedances are modeled to understand the behavior of the data in the tails. Many methods of determining an optimal threshold have been developed. The most common one is the graphical method proposed by [11]. This method

is however subjective and requires experts to determine the threshold. The most successful method is the Maximum Product of Spacing (MPS). Maximum product of spacing (MPS) method or maximum spacing estimation (MSE) method was proposed by Cheng [12] and Ramnaby [13] as an alternative method to maximum Likelihood Estimate method (MLE). A threshold approach for peaks over threshold using MPS was carried out by [14] and noted that the selection of a threshold is an important and challenging problem [15]. While studying traditional estimation methods and MPS in Generalized Inverted Exponential Distribution found out that MPS outperformed MLE and least square (LSE) methods on the basis of K-S distance and Akaike Information Criterion (AIC). This method however encounters a problem whenever the exceedances have a tie. This study intended to offer a solution to this problem.

## 2. Methodology

### 2.1. Improved MPS Methodology

The MPS allows efficient estimators in non regular cases where MLE may not exist. This is especially relevant to the GEV distribution in which the MLE does not exist when  $\varepsilon < -1$ . According to [12] Maximum spacing estimators are sensitive to closely spaced observations, and especially ties. In cases of ties, some scholars have suggested that one value of each tie is taken [16] and [17]. Let  $x_1, x_2, \dots, x_n$  be a random sample of independent observations from a continuous distribution  $F_{\theta_0}$  belonging to  $F_{\theta}, \theta \in \Theta$ . Applying the probability transform  $F_{\theta}(\cdot)$  to the order statistics  $x_{1,n} \leq x_{2,n} \leq \dots \leq x_{n,n}$  yields  $0 \equiv F_{\theta}(x_{0,n}) \leq F_{\theta}(x_{1,n}) \leq \dots \leq F_{\theta}(x_{n+1,n}) \equiv 1$ . We define the spacings as the gaps between the values of the distribution function at adjacent ordered points

$$D_i(\theta) = F_{\theta}(x_i) - F_{\theta}(x_{i-1}) \quad (1)$$

for  $i = 1, 2, \dots, n+1$ . The maximum spacing estimator of  $\theta_0$  was defined as value that maximizes the logarithm of the geometric mean of sample spacings [12].

$$\hat{\theta} = \arg \max_{\theta \in \Theta} S_n(\theta) \quad (2)$$

where

$$\begin{aligned} S_n(\theta) &= \ln^{n+1} \sqrt{(D_1(\theta) \cdot D_2(\theta) \cdots D_{n+1}(\theta))} \\ &= \frac{1}{n+1} \sum_{i=1}^{n+1} \ln D_i(\theta) \end{aligned}$$

This maximum spacing estimator is sensitive to the ties. That is, for any

$$x_{i+m} = x_{i+m-1} = \dots = x_i$$

Then  $D_{i+m}(\theta) = D_{i+m-1}(\theta) = \dots = D_i(\theta)$ . This therefore collapses the method. The modified MPS method proposed here is to use grouped data frequency table. Let  $x_1, x_2, \dots, x_n$  occur  $f_1, f_2, \dots, f_n$  times respectively. The geometric mean is given by

$$G = (x_1^{f_1} \cdot x_2^{f_2} \cdots x_n^{f_n})^{\frac{1}{N}} = \left[ \prod_{i=1}^n x_i^{f_i} \right]^{\frac{1}{N}}$$

This implies that

$$\ln G = \frac{1}{N} \sum_{i=1}^n f_i \ln x_i \tag{3}$$

This leads to the modified MPS method as

$$S_n(\theta) = \ln^{n+1} \sqrt[n+1]{(D_1^{f_1}(\theta) \cdot D_2^{f_2}(\theta) \cdots D_{n+1}^{f_{n+1}}(\theta))} = \frac{1}{n+1} \sum_{i=1}^{n+1} f_i \ln D_i(\theta) \tag{4}$$

In case  $f_1 = f_2 = \cdots = f_{n+1} = 1$ , then we go back to the standard MPS. The Spacings are such that  $\sum_{i=1}^n D_i(\theta) = 1$ . Under MPS, the  $D_i(\theta)$ 's are defined as:

$$\begin{aligned} D_1(\theta) &= F(x_{1:n}, \theta) \\ D_i(\theta) &= F(x_{i:n}, \theta) - F(x_{i-1:n}, \theta) \\ D_{n+1}(\theta) &= 1 - F(x_{n:n}, \theta) \end{aligned}$$

Therefore, Equation (4) can be partitioned as:

$$S_n(x_i; \theta, \varepsilon, \sigma) = \frac{1}{n+1} \left\{ f_1 \ln D_1(\theta) + \sum_{i=2}^n f_i \ln D_i(\theta) + f_{n+1} \ln D_{n+1}(\theta) \right\} \tag{5}$$

## 2.2. Estimation of Generalized Pareto Distribution Using the Modified MPS Method

To estimate the parameters, we substitute the GPD

$$G(x; \varepsilon, \sigma) = \begin{cases} 1 - \left[ 1 + \varepsilon \left( \frac{x-u}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}}, & \varepsilon \neq 0 \\ 1 - \exp \left[ - \left( \frac{x-u}{\sigma} \right) \right], & \varepsilon = 0 \end{cases} \tag{6}$$

into the MPS method. This lead to two cases of estimating the GPD parameters.

### 2.2.1. Case 1: When $\varepsilon \neq 0$

In this case:

$$\begin{aligned} D_1 &= 1 - \left[ 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \tag{7} \\ D_i &= \left( 1 - \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \right) - \left( 1 - \left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \right) \end{aligned}$$

which leads to

$$D_i = \left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} - \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \tag{8}$$

and

$$D_{n+1} = 1 - \left( 1 - \left[ 1 + \varepsilon \left( \frac{x_n - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \right) \tag{9}$$

implying that

$$D_{n+1} = \left[ 1 + \varepsilon \left( \frac{x_n - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}}$$

Therefore, Equation (5) now becomes:

$$S_n(x_i; \theta, \varepsilon, \sigma) = \frac{1}{n+1} \left\{ f_1 \ln \left( 1 - \left[ 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} \right) + \sum_{i=2}^n f_i \ln \left( \left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} - \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} \right) + f_{n+1} \ln \left[ 1 + \varepsilon \left( \frac{x_n - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} \right\} \quad (10)$$

The estimation of the parameters involves taking partial derivatives of Equation (10) with respect to each of the parameters and setting the result to zero. For the estimation of  $\varepsilon$ , the first term on the R.H.S is worked out as:

Let

$$K_1 = \ln \left( 1 - \left[ 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} \right) \quad (11)$$

implying that

$$\frac{\partial K_1}{\partial \varepsilon} = \frac{(x_1 - \theta)}{\sigma \varepsilon \left( 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right)} - \frac{1}{\varepsilon^2} \ln \left( 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right) \quad (12)$$

Working out the second term of Equation (10);

Let

$$K_2 = \ln \left\{ \left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} - \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} \right\} \quad (13)$$

Therefore;

$$\begin{aligned} \frac{\partial K_2}{\partial \varepsilon} = & \left( \frac{1}{\left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}} - \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right]^{\frac{1}{\varepsilon}}} \right) \\ & \times \left\{ \left( \frac{1}{\varepsilon^2} \ln \left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right] - \frac{1}{\varepsilon} \times \frac{1}{1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right)} \times \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right) \right. \\ & \left. - \left( \frac{1}{\varepsilon^2} \ln \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right] - \frac{1}{\varepsilon} \times \frac{1}{1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right)} \times \left( \frac{x_i - \theta}{\sigma} \right) \right) \right\} \quad (14) \end{aligned}$$

And the last term of Equation (10);

Let

$$K_3 = \ln \left[ 1 + \varepsilon \left( \frac{x_n - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \tag{15}$$

Therefore,

$$\frac{\partial K_3}{\partial \varepsilon} = \left( \frac{1}{\left[ 1 + \varepsilon \left( \frac{x_n - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}}} \right) \left\{ \frac{1}{\varepsilon^2} \ln \left[ 1 + \varepsilon \left( \frac{x_n - \theta}{\sigma} \right) \right] - \frac{1}{\varepsilon} \left( \frac{x_n - \theta}{\sigma} \right) \right\} \tag{16}$$

Similarly, we parameter to estimate is  $\sigma$ , from the:

$$\frac{\partial K_1}{\partial \sigma} = \frac{1}{\left[ 1 - \left[ 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \right]} \left[ \left[ 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right]^{-\frac{(1+\varepsilon)}{\varepsilon}} \right] \times \left( \frac{x_1 - \theta}{\sigma^2} \right) \tag{17}$$

$$\frac{\partial K_2}{\partial \sigma} = \frac{\partial K_2}{\partial w} \times \frac{\partial w}{\partial \sigma}$$

$$\begin{aligned} \frac{\partial K_2}{\partial \sigma} &= \frac{1}{\left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} - \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}}} \\ &\times \left\{ \left( \frac{x_i - \theta}{\sigma^2} \right) \left[ 1 + \varepsilon \left( \frac{x_i - \theta}{\sigma} \right) \right]^{-\frac{(1+\varepsilon)}{\varepsilon}} \right. \\ &\left. - \left( \frac{x_{i-1} - \theta}{\sigma^2} \right) \left[ 1 + \varepsilon \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right]^{-\frac{(1+\varepsilon)}{\varepsilon}} \right\} \end{aligned} \tag{18}$$

and

$$\frac{\partial K_3}{\partial \sigma} = - \left( \frac{x_n - \theta}{\sigma^2} \right) \left[ 1 + \varepsilon \left( \frac{x_n - \theta}{\sigma} \right) \right]^{-1} \tag{19}$$

Finally, we estimate  $\theta$  as follows:

$$\begin{aligned} \frac{\partial K_1}{\partial \theta} &= \frac{1}{\left[ 1 - \left[ 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \right]} \\ &\times \frac{1}{\varepsilon} \left( 1 - \left[ 1 + \varepsilon \left( \frac{x_1 - \theta}{\sigma} \right) \right]^{-\frac{(1+\varepsilon)}{\varepsilon}} \right) \times \left( \frac{x_1 - \theta}{\sigma} \right) \\ \frac{\partial K_2}{\partial \theta} &= \frac{\partial K_2}{\partial h} \times \left( \frac{\partial A}{\partial \theta} - \frac{\partial M}{\partial \theta} \right) \end{aligned} \tag{20}$$

$$\frac{\partial K_2}{\partial \theta} = \frac{1}{\left[1 + \varepsilon \left(\frac{x_{i-1} - \theta}{\sigma}\right)\right]^{\frac{1}{\varepsilon}} - \left[1 + \varepsilon \left(\frac{x_i - \theta}{\sigma}\right)\right]^{\frac{1}{\varepsilon}}} \times \left\{ \left(\frac{x_{i-1} - \theta}{\sigma}\right) \left[1 + \varepsilon \left(\frac{x_{i-1} - \theta}{\sigma}\right)\right]^{\frac{(1+\varepsilon)}{\varepsilon}} - \left(\frac{x_i - \theta}{\sigma}\right) \left[1 + \varepsilon \left(\frac{x_i - \theta}{\sigma}\right)\right]^{\frac{(1+\varepsilon)}{\varepsilon}} \right\} \quad (21)$$

and

$$\frac{\partial K_3}{\partial \theta} = \left(\frac{x_n - \theta}{\sigma}\right) \left[1 + \varepsilon \left(\frac{x_n - \theta}{\sigma}\right)\right]^{-1} \quad (22)$$

Therefore, after differentiating partially Equation (10) with respect to the parameters, we get the normal Equations (23), (24) and (25);

$$S'_\varepsilon = \frac{1}{n+1} \left\{ f_1 \frac{\partial K_1}{\partial \varepsilon} + \sum_{i=2}^n f_i \frac{\partial K_2}{\partial \varepsilon} + f_{n+1} \frac{\partial K_3}{\partial \varepsilon} \right\} = 0 \quad (23)$$

where the terms  $\frac{\partial K_1}{\partial \varepsilon}$ ,  $\frac{\partial K_2}{\partial \varepsilon}$  and  $\frac{\partial K_3}{\partial \varepsilon}$  are from Equations (12), (14) and (16) respectively.

$$S'_\sigma = \frac{1}{n+1} \left\{ f_1 \frac{\partial K_1}{\partial \sigma} + \sum_{i=2}^n f_i \frac{\partial K_2}{\partial \sigma} + f_{n+1} \frac{\partial K_3}{\partial \sigma} \right\} = 0 \quad (24)$$

where the terms  $\frac{\partial K_1}{\partial \sigma}$ ,  $\frac{\partial K_2}{\partial \sigma}$  and  $\frac{\partial K_3}{\partial \sigma}$  are from Equations (17), (18) and (19) respectively.

$$S'_\theta = \frac{1}{n+1} \left\{ f_1 \frac{\partial K_1}{\partial \theta} + \sum_{i=2}^n f_i \frac{\partial K_2}{\partial \theta} + f_{n+1} \frac{\partial K_3}{\partial \theta} \right\} = 0 \quad (25)$$

where the terms  $\frac{\partial K_1}{\partial \theta}$ ,  $\frac{\partial K_2}{\partial \theta}$  and  $\frac{\partial K_3}{\partial \theta}$  are from Equations (20), (21) and (22) respectively. The parameters were obtained from Equations (23), (24) and (25) using numerical analysis procedures for optimization.

### 2.2.2. Case 2: When $\varepsilon = 0$

Parameters under this case are estimated here. When  $\varepsilon = 0$ , the spacings become;

$$D_1 = 1 - \exp\left[-\left(\frac{x_1 - \theta}{\sigma}\right)\right] \quad (26)$$

$$D_i = \exp\left[-\left(\frac{x_{i-1} - \theta}{\sigma}\right)\right] - \exp\left[-\left(\frac{x_i - \theta}{\sigma}\right)\right] \quad (27)$$

$$D_{n+1} = \exp\left[-\left(\frac{x_n - \theta}{\sigma}\right)\right] \quad (28)$$

Therefore, Equation (5) can be written as;

$$S_n(x_i; \theta, \varepsilon, \sigma) = \frac{1}{n+1} \left\{ f_1 \ln \left[ 1 - \exp \left[ - \left( \frac{x_1 - \theta}{\sigma} \right) \right] \right] + \sum_{i=2}^n f_i \ln \left[ \exp \left[ - \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right] - \exp \left[ - \left( \frac{x_i - \theta}{\sigma} \right) \right] \right] - f_{n+1} \times \left( \frac{x_n - \theta}{\sigma} \right) \right\} \quad (29)$$

Let

$$K_1^* = \ln D_1 \quad (30)$$

Therefore,

$$\frac{\partial K_1^*}{\partial \sigma} = - \frac{x_1 - \theta}{\sigma^2 \left( 1 - \exp \left[ - \left( \frac{x_1 - \theta}{\sigma} \right) \right] \right)} \exp \left[ - \left( \frac{x_1 - \theta}{\sigma} \right) \right] \quad (31)$$

Similarly, let  $P = \exp \left[ - \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right] - \exp \left[ - \left( \frac{x_i - \theta}{\sigma} \right) \right]$  and

$$K_2^* = \ln P \quad (32)$$

Therefore,

$$\frac{\partial K_2^*}{\partial \sigma} = \frac{\left( (x_i - \theta) \exp \left[ - \left( \frac{x_i - \theta}{\sigma} \right) \right] - (x_{i-1} - \theta) \exp \left[ - \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right] \right)}{\sigma^2 \left( \exp \left[ - \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right] - \exp \left[ - \left( \frac{x_i - \theta}{\sigma} \right) \right] \right)} \quad (33)$$

Let

$$K_3^* = \frac{(x_n - \theta)}{\sigma} \quad (34)$$

implying that;

$$\frac{\partial K_3^*}{\partial \sigma} = \frac{1}{\sigma^2} (x_n - \theta) \quad (35)$$

The equations for estimating  $\theta$  can also be derived from,

$$\frac{\partial K_1^*}{\partial \theta} = - \frac{\exp \left[ - \left( \frac{x_1 - \theta}{\sigma} \right) \right]}{\sigma \left( 1 - \exp \left[ - \left( \frac{x_1 - \theta}{\sigma} \right) \right] \right)} \quad (36)$$

Next, let  $K_2^*$  and  $P$  be defined as Equation (32), so that

$$\frac{\partial K_2^*}{\partial P} = \frac{1}{P} = \frac{1}{\exp \left[ - \left( \frac{x_{i-1} - \theta}{\sigma} \right) \right] - \exp \left[ - \left( \frac{x_i - \theta}{\sigma} \right) \right]}$$

Simplifying to;



$$\frac{\partial K_2^*}{\partial \theta} = \frac{1}{\sigma} \quad (37)$$

Lastly, let  $K_3^*$  be defined as Equation (15), implying that;

$$\frac{\partial K_3^*}{\partial \theta} = -\frac{1}{\sigma} \quad (38)$$

Therefore, after differentiating partially Equation (29) with respect to the parameters, we get the normal Equations (39) and (40);

$$S'_{\sigma^*} = \frac{1}{n+1} \left\{ f_1 \frac{\partial K_1^*}{\partial \sigma} + \sum_{i=2}^n f_i \frac{\partial K_2^*}{\partial \sigma} + f_{n+1} \frac{\partial K_3^*}{\partial \sigma} \right\} = 0 \quad (39)$$

where the terms  $\frac{\partial K_1^*}{\partial \sigma}$ ,  $\frac{\partial K_2^*}{\partial \sigma}$  and  $\frac{\partial K_3^*}{\partial \sigma}$  are from Equations (31), (33) and (35) respectively.

$$S'_{\theta^*} = \frac{1}{n+1} \left\{ f_1 \frac{\partial K_1^*}{\partial \theta} + \sum_{i=2}^n f_i \frac{\partial K_2^*}{\partial \theta} + f_{n+1} \frac{\partial K_3^*}{\partial \theta} \right\} = 0 \quad (40)$$

where the terms  $\frac{\partial K_1^*}{\partial \theta}$ ,  $\frac{\partial K_2^*}{\partial \theta}$  and  $\frac{\partial K_3^*}{\partial \theta}$  are from Equations (36), (37) and (38) respectively. The parameters were obtained from Equations (39) and (40) using numerical analysis procedures for optimization.

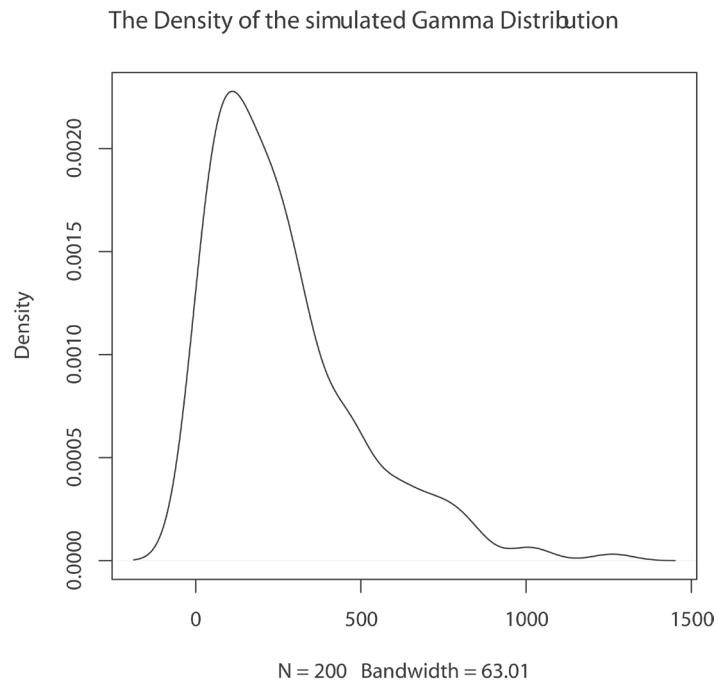
### 3. Results and Discussion

#### 3.1. Simulation Study

A simulation was performed to compare the standard MPS methodology with the improved MPS methodology. We simulated data from a gamma distribution with the parameters shape = 2.6, scale = 1:1000. Repetitions were later introduced in the order of 0, 20, 40 and 60. The repeated values gave rise to situations of ties. Gamma distribution is known to have fairly heavy tails. To determine our threshold, we simulated a set of data constituting 300 values. 100 values did not have a repetition while 100 values had each a repetition making them to have a frequency of 2 each. This set of data was used in the improved MPS model. After the simulation, this set of data was reorganized in such a way that the 300 values had a frequency of 1 each regardless of whether it was repeated or not. This set of data was used in the standard MPS model. The normal equations derived above were used as the model for the improved MPS methodology. For the three parameter models 23, 24 and 25 were used while for two-parameter model 39 and 40 were used.

Suitable values for  $k$  and  $\theta$  for which the gamma distribution would produce long tail were selected. In this case, a simulation using  $k = 2.6$  and  $\theta = 1:1000$  was made and the density of the simulation is shown in **Figure 1**.

The x-axis of **Figure 1** represents the number of values generated for the specified parameters of gamma distribution. The y-axis represents the density of those values. **Figure 1** indicates that majority of the values are concentrated



**Figure 1.** The density of the Gamma distributed initial values.

between 0 and 450. The values with big magnitude concentrate to the right of the distribution. The simulated values are Gamma distributed with a long right tail.

When the two parameter standard and improved models were used, the following results **Table 1** from the simulation were obtained.

The threshold (location parameter) from the improved MPS model was higher than that obtained through standard MPS model **Table 1**. However the scale parameter of the improved MPS model was lower than that obtained through the standard MPS model. Three parameter improved MPS model and the standard MPS model were used to determine the threshold of the same data set. The results are shown in **Table 2**.

The threshold (location parameter) determined from the improved MPS model was high compared to the threshold determined through the standard MPS model **Table 2**. The shape parameter determined through the improved MPS model was also high compared to the one obtained through the standard MPS model. However, the scale parameter obtained through the improved MPS model was lower than that obtained through the standard MPS model. To assess the performance of the parameters in the standard and improved models, the GPD parameters determined in **Table 1** and **Table 2** were backtested in the simulated data. **Table 3** contains a summary of the values obtained.

From **Table 1**, a threshold of 725.5767 was determined for the standard MPS model from a data that contained ties. When this threshold was backtested on the same data, It produced 18 data values as excesses **Table 3**. This was a proportion of 0.06. A threshold of 736.476 was determined for the improved MPS model from the same data that contained ties **Table 1**. When this threshold was

**Table 1.** Pars estimates of two parameter model.

	Location	Scale
Improved	736.476	13.72969
Standard	725.5767	16.31062

**Table 2.** Pars estimates of three parameter model.

	Location	Scale	Shape
Improved	738.1303	9.483573	-0.84884
Standard	726.3707	13.33941	-5.49648

**Table 3.** Model selection.

MLE and MPLE Estims	2parStd	2parImpovd	3parStd	3parImpvd
Threshold	725.5767	736.476	726.3707	738.1303
No. above	18	15	18	15
Proportion above	0.06	0.05	0.06	0.05
Scale estimate	146.3	163.9	98.69	130
Scale std. err	34.49	42.31	31.66	40.8
Shape estimate	0.3728	0.1903	0.3879	0.2146
Shape std. err	0.2808	0.2573	0.2855	0.2644
Asympt Var-Cov-Scale	1189	1790	1002	1664
Asympt Var-Cov-Shape	0.07887	0.06623	0.08151	0.06993
Deviance	214.8238	183.1161	214.5401	182.7789
Penalized Deviance	215.4867	182.9699	215.796	183.3916
AIC	216.8238	185.1161	216.5401	184.7789
Penalized AIC	217.4867	184.9699	217.796	185.3916

backtested, it produced excesses of 15 data values which was a proportion of 0.05 **Table 3**. The threshold from improved model was higher than that of standard MPS model hence fewer data values in the excesses. The deviance and AIC criterion were used to compare the threshold that would make a generalized pareto distribution (GPD) model more fitting **Table 3**. The deviance of the two parameter standard MPS model was 214.8238 while that of the two parameter improved MPS model was 183.1161. The deviance of the improved model was lower than that of the standard MPS model. The AIC value of the standard MPS model was 216.8238 while that of the improved MPS model was 185.1161. The improved model had a lower AIC value. Treating the GPD model of the standard MPS model and that of improved MPS model as competing models, then the improved MPS model produced a more optimal threshold than that of standard MPS model. A threshold value of 726.3707 was determined for the three parameter standard model while that for the improved MPS model was

738.1303 **Table 2**. These two thresholds were back tested in the appropriate data sets **Table 3**. 18 excesses were realized on backtesting 726.3707 while 15 excesses were realized on back testing the value of 738.1303. The deviance of the standard model was 214.5401 while that of the improved model was 182.7789. The deviance of the improved MPS model was lower than that of the standard model. The AIC value of the improved model is lower than that of the standard MPS model. Whenever two models are competing, the model with the lower AIC is a better model. Thus the threshold obtained from improved MPS model makes the GPD to perform better than that obtained from a standard MPS.

### 3.2. Effects of Number of Repetitions on Threshold

In this section, the effect of number of repetitions on threshold was investigated. The gamma distribution with parameters  $k = 2.6$  and  $\theta = 1:1000$  was used. However, to change the number of repetitions,  $y_1$  and  $y_2$  were made to take different values. To create a sample with 20 repetitions, variables  $y_1 = rep(1, 240)$  and  $y_2 = rep(1:2, 20)$  were used. To create a sample with 40 repetitions, variables  $y_1 = rep(1, 180)$  and  $y_2 = rep(1:2, 40)$  were used, while a sample of 60 repetitions was created by using the variables  $y_1 = rep(1, 120)$  and  $y_2 = rep(1:2, 60)$ . For 0 repetitions, the variables used were  $y_1 = rep(1, 50)$  and  $y_2 = rep(1, 250)$

The repetitions cause ties and therefore, these samples contain grouped ties. The improved MPS model used this raw sample. For this sample to be used with standard MPS model, the data had to be ungrouped so as each value had a frequency of one. The density of the distribution in the four cases of repetitions were plotted as shown in **Figure 2**.

The x-axis of the plots in **Figure 2** represents the number of the values specified in the gamma function while the y-axis represents the density of those values for different repetitions. The figure shows that there were some values in the tail and that the distribution was skewed towards right. This distribution was similar to Gamma density which was used in the simulation of this data set. The data indicate that several values are in the tails of the distribution meaning that the density's contained extreme values. This justified the use of GPD to analyze the simulated data.

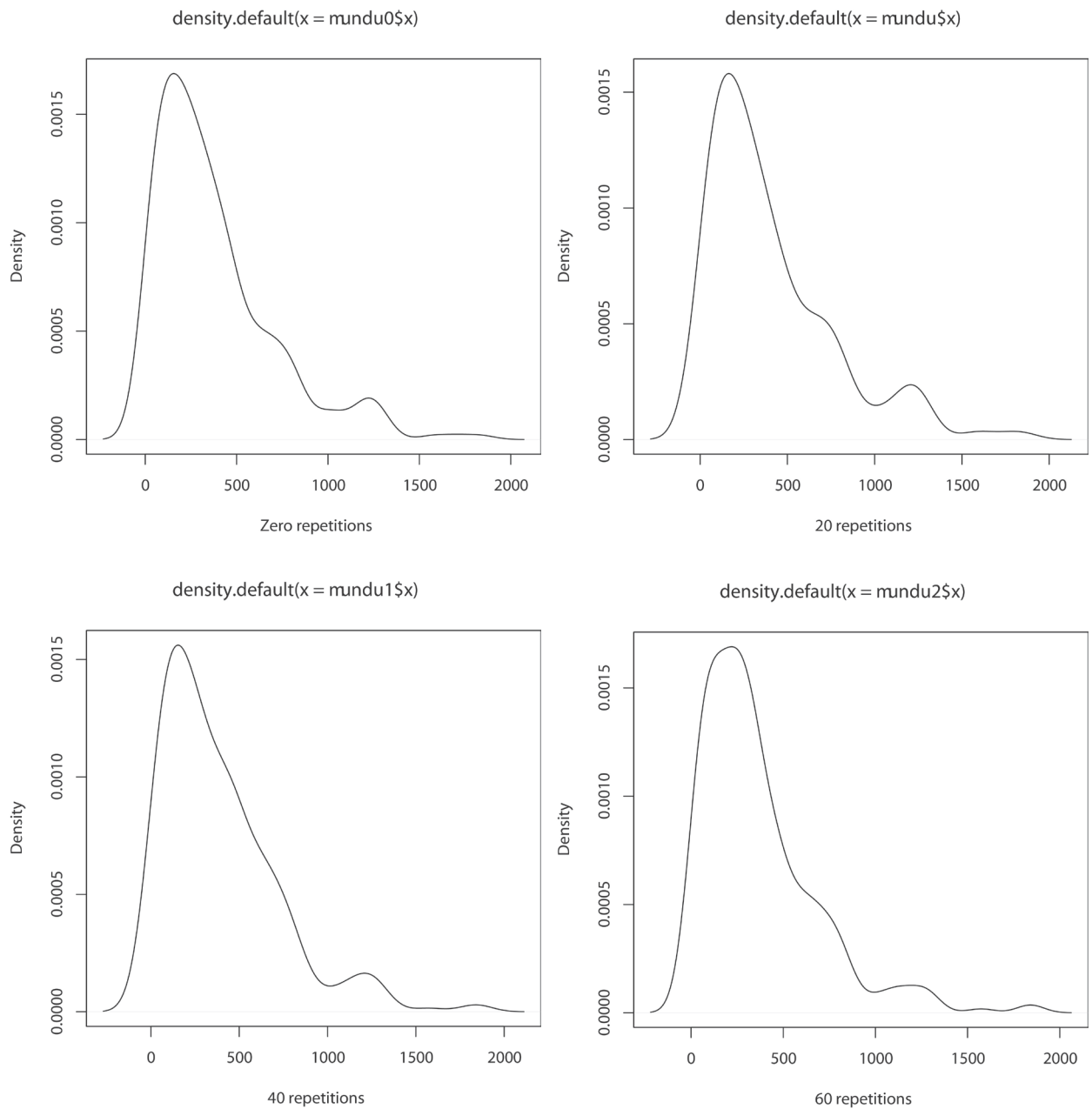
According to [16], some values of each tie have to be dropped to leave only one value of each tie for the standard MPS to work. Therefore, samples of size 300 with ungrouped ties of 0, 20, 40 and 60 repetitions were each subjected to the two parameter standard MPS model. The results obtained were as in **Table 4**. The threshold (location) value increased as ties increased. However, the scale parameter increased from 5.993496 to 9.99465 then decreased as the number of ties increased. The samples with grouped ties were subjected to the two parameter improved MPS model and the results obtained are as in **Table 5**. It was observed that, the location parameter(threshold value) increased as the number of ties increased. The scale parameter increased from 4.097801 to

**Table 4.** Two parameter std model estimates.

Repetitions	0	20	40	60
Location	1111.897	1116.442	1121.845	1124.372
Scale	5.993496	9.99465	3.056949	1.373637

**Table 5.** Two parameter improved model estimates.

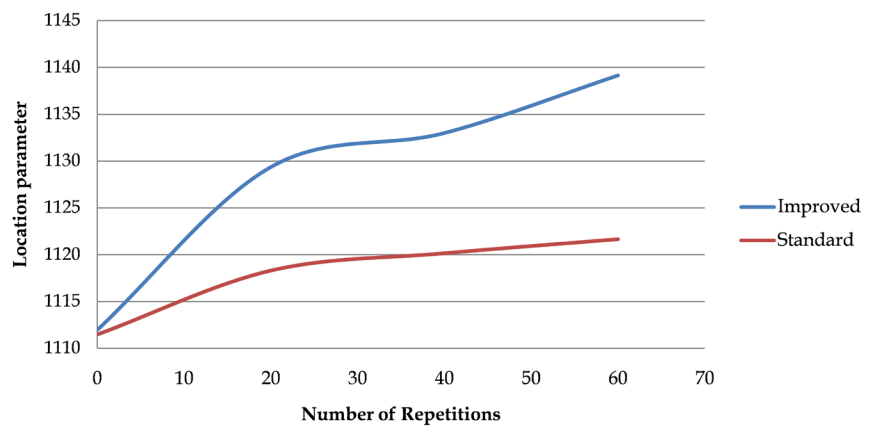
Repetitions	0	20	40	60
Location	1111.009	1127.579	1130.066	1140.35
Scale	4.097801	7.830554	5.119941	4.395648



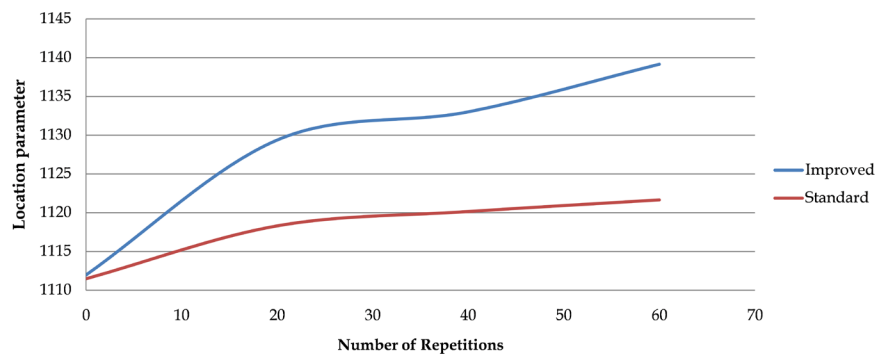
**Figure 2.** The densities for data with repetitions.

7.830554 then decreased as the number of ties increased. The samples with ungrouped ties were subjected to a three parameter standard MPS model and the results are as in **Table 6**. The location (threshold value) parameter increased from 1111.473 to 1121.675 as the number of ties increased. The scale decreased from 17.44149 to 4.950339. The shape also decreased from  $-0.05918$  to  $-6.59819$  as ties increased. The samples with grouped ties were subjected to the three parameter improved MPS model and the results are as in **Table 7**. The location (threshold value) parameter increased from 1111.954 to 1141.156 as the number of ties increased from 0 to 60. The scale parameter decreased from 15.4904 to 9.42994 as the number of ties increased from 0 to 60. The shape parameter increased from 0.05365 to 2.54892 at 40 repetitions then decreased to  $-4.74385$  at 60 repetitions.

Plots to compare the performance of the parameters obtained were made as indicated in **Figures 3-7**.



**Figure 3.** Behavior of location parameter in the two parameter model.



**Figure 4.** Behavior of location parameter in the three parameter model.

**Table 6.** Three parameter standard model estimates.

Repetitions	0	20	40	60
Location	1111.473	1118.298	1120.145	1121.675
Scale	17.44149	12.04152	4.716338	4.950339
Shape	$-0.05918$	$-0.04098$	$-1.90142$	$-6.59819$

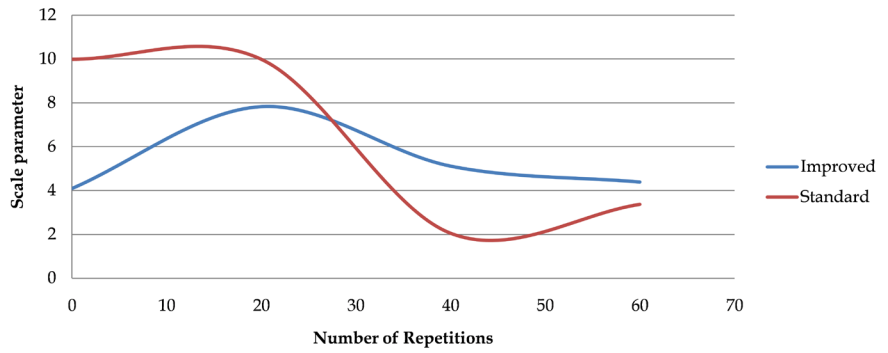


Figure 5. Behavior of scale parameter in the two parameter model.

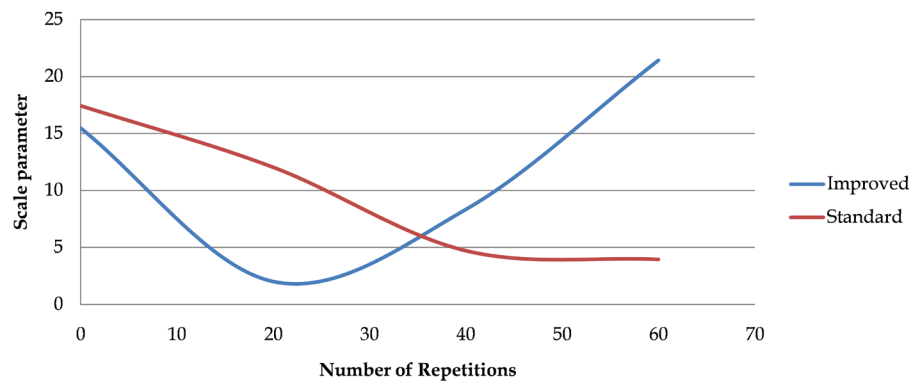


Figure 6. Behavior of scale parameter in the three parameter model.

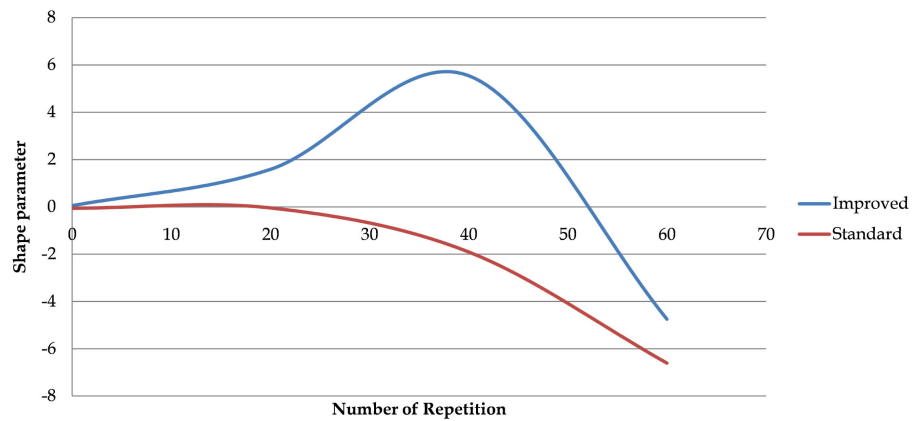


Figure 7. Behavior of shape parameter in the three parameter model.

Table 7. Three parameter improved model estimates.

Repetitions	0	20	40	60
Location	1111.954	1128.368	1133.003	1141.156
Scale	15.4904	7.009472	8.335507	9.42994
Shape	0.05365	1.586433	2.54892	-4.74385

### 3.3. Location Parameter

When there was no tie **Figure 3**, the two parameter standard MPS model and the

improved two parameter MPS model give the same threshold. The plot for the two parameter models indicate that the threshold improved as the number of repetitions increased. The increase rate of the improved MPS model was higher than that of the standard MPS model. In both models, there was a decrease of the threshold at around 32 repetitions after which the threshold continued to improve.

The trend observed in the two parameter model **Figure 3** was also observed in the performance of the three parameter models **Figure 4**. However, the drop at 32 repetitions in the three parameter models was not as big as in the two parameter models. From the two plots above, the improved MPS model performed better than the standard MPS models as ties increased.

### 3.4. Scale Parameter

A plot of scales was also made to compare the performance of the scale parameter as the number of repetitions increased for the two and three parameter MPS model **Figure 5**.

The two parameter model had different scale parameters when there were no ties. The scale of the standard MPS model increased faster than that of the improved MPS model as the repetitions increased up to the 20 repetitions after which the two models showed a downward trend **Figure 5**. However, the downward trend rate is higher in the standard MPS model than the improved MPS model. The two models had the same scale parameter at the 31 repetitions. Beyond 31 repetition, the improved model had a higher scale parameter than standard model despite the fact that both of them were still on the decline.

The three parameter MPS models **Figure 6** had different scale parameters at zero repetitions. The size of the parameters showed a downward trend but the improved MPS model had a higher rate of downward trend than the standard MPS model up to 20 repetitions after which the trend of the improved MPS model changed and showed some improvements. However the trend of the standard MPS model continued to drop until the 45 repetitions after which it started to improve. The two models had the same scale parameter at about 31 repetitions. The improved MPS model had a higher rate of improvement than the standard one. The trend of the scale parameter in the two parameter model and the three parameter model are similar.

### 3.5. Shape Parameter

The two parameter standard and improved model have location and scale parameters only. These models have a zero shape parameter. For the three parameter standard and improved models, the shape parameter performed as shown in **Figure 7**.

The two models had the same shape parameter when the data had no repetitions(ties) 7. The shape parameter of the improved MPS model improved



up to around 38 repetitions after which it showed a downward trend. The standard model remained steady between 0 and 20 repetitions after which it followed a downward trend. The drop rate in the shape parameter was higher on improved model than it was on the standard model. The shape parameter showed a consistent decrease beyond 20 repetitions for the standard model and 38 repetitions for the improved model. A general observation was that there was a change in the trend of all parameters in between 30 and 40 repetitions.

### 3.6. Back Testing Two Parameter Models

The determined threshold obtained in different samples were back tested in the sample data they were obtained from to assess their performance.

For two parameter models at 0 repetitions, both models, the standard and the improved MPS model **Table 8**, had the same number of observations over the threshold and both had the same deviance value and AIC criterion value. Therefore, at this level both models performance were the same. At 20 repetitions, the standard model had 23 observations over the threshold while the improved model had 21 observations over the threshold. However, the deviance value and AIC criterion values were lower in the case of improved model compared to those of the standard model. For 40 repetitions, the number above the threshold was 14 for both models but the deviance value and AIC criterion values of the improved model were lower compared to those criterion values for improved model were lower than those of standard model. From this analysis, the improved MPS model performed better than the standard MPS model.

**Table 8.** Two Parameter model with repetitions.

Ests	0Std	0Impd	20Std	20Impd	40Std	40Impd	60Std	60Impd
Thresh	1111.897	1111.009	1116.402	1127.579	1121.845	1130.066	1124.372	1140.35
No.abv	16	16	23	21	14	14	11	11
Pro.abv	0.0533	0.0533	0.0767	0.07	0.0467	0.0467	0.0367	0.0367
Sca.est	168.3	169.3	169.6	178.3	180.7	170.8	199.5	174.7
Scastd.er	51.09	51.37	42.41	46.64	59.37	56.54	76.02	67.84
Sha. Est.	0.5588	0.5505	0.3235	0.6702	0.5486	0.6172	0.5641	0.7014
Shastd.er	0.309	0.3065	0.1047	0.2861	0.3211	0.3412	0.3779	0.4265
AVarCo. Sc	2610	2639	1798	2175	3525	3197	5779	4602
AVarCo. Sh	0.09545	0.09395	0.1047	0.08186	0.1031	0.1164	0.1428	0.1819
Deviance	208.5296	208.6634	301.9033	277.4688	184.4465	183.3083	147.1072	145.1982
Pen. Devia	204.2043	204.3792	292.9033	269.3051	181.331	180.0228	145.7075	144.0077
AIC	210.5296	210.6634	303.9033	279.4688	186.4465	185.3038	149.1072	147.1982
Pena. AIC	206.2043	206.3792	292.5338	271.3051	183.331	182.0228	147.7075	146.0077

**Table 9.** Three parameter model with repetitions.

Ests	0repsStd	0repsImpd	20repsStd	20repsImpd	40repsStd	40repsImpd	60repsStd	60repsImpd
Thresh	1111.473	1111.954	1118.298	1128.3679	1120.145	1133.003	1121.647	1141.156
No.abv	16	16	21	21	14	14	11	11
Pro.abv	0.0533	0.0533	0.07	0.07	0.0467	0.0467	0.0367	0.0367
Sca.est	168.8	168.2	188.5	177.5	182.6	167.1	203	173.3
Scastd.err	51.23	51.07	49.17	46.42	59.96	55.54	77.09	67.39
Sha. Est.	0.5548	0.5594	0.5979	0.6765	0.5362	0.6467	0.5427	0.709
Shapstd.er	0.3078	0.3091	0.2705	0.2875	0.318	0.3488	0.3709	0.4295
AsVarSc	2624	2608	2418	2155	3595	3085	5944	4542
AsVarSh	0.09473	0.09555	0.07319	0.08265	0.1011	0.1217	0.1376	0.1845
Deviance	208.5939	208.5209	279.2341	277.3027	184.6582	182.8486	147.3715	145.0816
Pen. Devia	204.2876	204.1926	271.4901	269.1153	181.5982	179.5488	145.997	143.9223
AIC	210.5939	210.5209	281.231	279.3027	186.6582	184.8486	149.3715	147.0816
Pena. AIC	206.2876	206.1926	273.4911	271.1153	183.5982	181.5488	147.997	145.9223

### 3.7. Back Testing Three Parameter Model

To assess the performance of the threshold obtained when the GPD had the shape parameter, a back testing was done on the samples through the two models and the results were as in **Table 9**.

For 0 repetitions, the number of observations above the threshold were 16 in both models **Table 9**. The deviance values and the AIC values were the same for the three parameter model. Meaning that the two models performed equally the same. For 20 repetitions, the number of observations were 21 but the deviance value and AIC criterion value were lower in case of improved model compared to those of standard model. The deviance value and AIC criterion value for the improved model were also lower than those of standard model in the 40 and 60 repetitions although, the number of observations in the 40 repetitions were 14 while that of the 60 repetitions were 11. In this case therefore, the improved model performed better than the standard model. The presence of ties in the data set causes a difference in the performance of the two models.

## 4. Conclusion

This study helped to improve the MPS model by introducing the concept of  $f$  to both two-parameter and three-parameter model 23, 24, 25, 39 and 40. Through simulation, the improved MPS, both two-parameter and three-parameter models yielded a higher threshold as compared to the two standard MPS model **Table 1** and **Table 2**. The scale parameter of the improved MPS model was lower than that of the standard model. The shape parameter of the improved model performed well compared to the standard model. When the determined GPD parameters

were back tested into the same simulated data **Table 3**, the improved model performed better than the standard model. The deviance statistics and the AIC criterion for the improved model yielded lower values compared to the standard model. This study therefore managed to come up with an improved MPS model which was able to take care of the analysis of ties in the data set. The back testing for each repetition indicates that the improved MPS model performed better than the standard MPS model. When frequency  $f = 1$  of each data point, the improved MPS model reduces to the standard model. The improved MPS model helps to yield a more optimal threshold which in turn would help different sectors of the country's economy to be adequately prepared for any eventuality.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- [1] Butterfield, R. (2009) DFID Economic Impacts of Climate Change in Kenya, Rwanda and Burundi. ICPAC Kenya and SEI, 1-45.
- [2] Elizabeth, B., Claudia, R., Barack, O., Jawoo, K., Silvia, S. and Mario, H. (2011) Agricultural Management for Climate Change Adaptation. IFPRI, 1-40.
- [3] Katz, R. and Murphy, A. (1997) Economic Value of Weather and Climate Forecast. Cambridge University Press, London. <https://doi.org/10.1017/CBO9780511608278>
- [4] Prudhome, C. and Duncan, W. (1999) Mapping Extreme Rainfall in Mountaneous Region Using Geostatistical Techniques: A Case Study of Scotland. *International Journal of Climatology*, **19**, 1337-1356. [https://doi.org/10.1002/\(SICI\)1097-0088\(199910\)19:12<1337::AID-JOC421>3.3.CO;2-7](https://doi.org/10.1002/(SICI)1097-0088(199910)19:12<1337::AID-JOC421>3.3.CO;2-7)
- [5] Visschel, T., *et al.* (2012) Extreme Rainfall in West Africa: A Regional Modeling. *Water Resources Research*, **10**, 10-29.
- [6] Box, G.E.P. and Wilson, K.B. (1951) On the Experimental Attainment of Optimum Conditions. *Journal of the Royal Statistical Society Series B*, **13**, 1-45. <https://doi.org/10.1111/j.2517-6161.1951.tb00067.x>
- [7] Embrechts, P., McNeil, A. and Straumann, D. (2002) Correlation and Dependence in Risk Management: Properties and Pitfalls. In: Dempster, M.A.H., Ed., *Risk Management: Value at Risk and Beyond*, Cambridge University Press, Cambridge, 176-223. <https://doi.org/10.1017/CBO9780511615337.008>
- [8] Annie, P., David, H., Ann, C. and Stephene, P. (2007) Importance of Tail Dependence in Bivariate Frequency Analysis. *Journal of Hydrologic Engineering*, **12**, 394-403.
- [9] Balkema, A. and de Haan, L. (1974) Residual Life Time at Great Age. *The Annals of Probability*, **2**, 792-804. <https://doi.org/10.1214/aop/1176996548>
- [10] Pickands, J. (1975) Statistical Inference Using Extreme Order Statistics. *Annals of Statistics*, **3**, 119-131. <https://doi.org/10.1214/aos/1176343003>
- [11] Hill, B.M. (1975) A Simple General Approach to Inference about the Tail of a Distribution. *Annals of Statistics*, **13**, 331-341.
- [12] Cheng, R.C.H. and Amin, N.A.K. (1983) Estimating Parameters in Continuous Un-

- ivariate Distribution with a Shifted Origin. *Journal of the Royal Statistical Society*, **45**, 394-403. <https://doi.org/10.1111/j.2517-6161.1983.tb01268.x>
- [13] Ramnaby, B. (1984) The Maximum Spacing Method: An Estimation Method Related to the Maximum Likelihood Method. *Scandinavian Journal of Statistics*, **11**, 93-112.
- [14] Wong, S.T.W. and Wai, K.L. (2010) A Threshold Approach for Peaks over Threshold Modeling Using Maximum Product of Spacing. *Statistica Sinica*, **20**, 1257-1272.
- [15] Umash, S., Kumar, S. and Singh, R.K. (2014) A Comparative Study of Traditional Estimation Methods and Maximum Product of Spacing Method in Generalized Inverted Exponential Distribution. *Journal of Statistics Applications and Probability*, **3**, 153-169. <https://doi.org/10.12785/jsap/030206>
- [16] Cheng, R.C.H. and Stephen, M.A. (1989) A Goodness-of-Fit Test Using Moran's with Estimated Parameters. *Biometrika*, **76**, 386-392. <https://doi.org/10.1093/biomet/76.2.385>
- [17] Smith, R. (1985) *Statistics of Extreme Values*. International Statistical Institute, Amsterdam.