IAENG International Journal of Applied Mathematics

Using Simple Fixed-Point Iterations to Estimate Generalized Pareto Distribution Parameters

Sri Purwani and Riza Andrian Ibrahim

***Abstract*—Estimating generalised Pareto distribution (GPD) parameters is a fundamental step in modelling the extreme-value distribution of random variables. It is generally done with the maximum likelihood method, but there are generally difficulties in estimating GPD parameters using this method as there is no closed-form solution for the first derivative of the GPD log-likelihood function. This makes the solution difficult to determine analytically. However, numerical methods can be used as an alternative. Therefore, this study estimates the solution numerically using a simple fixed-point iteration method that is intuitive for both practitioners and professionals. We obtained three fixed-point iterations when estimating GPD parameters that met the unbiased estimator and convergence criteria. The iterations allow practitioners and professionals to directly and efficiently estimate GPD parameters when modelling extreme-value distributions of random variables.**

***Index Terms*—generalized Pareto distribution; parameter estimation; maximum likelihood method; fixed-point iteration; unbiasedness criteria; convergence criteria**

I. Introduction

EXTREME value theory (EVT) is a unique approach to determining the probability of the extreme value of a random variable on the left or right tail of the probability function [1]. This approach is widely used in various fields, such as climatology, finance, and insurance. In general, studies that apply EVT are more likely to analyse the probability of extreme values of random variables in the right tail of the probability function since this is more interesting [2]. This includes studies on maximum rainfall risk [3], maximum temperature risk [4], [5], [6], maximum sea wave height [7], [8], maximum disaster loss risk [9], [10], and maximum earthquake magnitude risk [11], [12], [13].

There are two methods for measuring the extreme-value probability of random variables using the EVT approach: 1) the block maxima (BM) method; and 2) the peaks-over threshold (POT) method. The latter is the most modern method of the two, and it facilitates the design of the cumulative-distribution function (CDF) model for extreme values in which the 'father' CDF of a random variable is unknown [2]. The advantage of this method lies in its ability to define extreme values based on values that exceed a particular threshold [14]. This definition often makes the size of extreme values in the POT method more significant than those in the BM method, indicating that the POT method is more representative of the population and that the probability results will be closer to the actual probability.

One of the fundamental steps in modelling the right-tailed probability function of a random variable using the POT method entails estimating generalized Pareto distribution (GPD) parameters. GPD is a two-parameter probability distribution that is crucial in modelling. Many methods can be used to estimate GPD parameters, but the most popular is the maximum likelihood (ML) method [15]. This method is intuitive as the estimated parameter sought is a two-dimensional vector that maximizes the GPD likelihood function. If the value of the GPD likelihood function reaches the maximum value, the sample that represents the population will very likely be observed. However, there are obstacles to determining the vector solution due to the first derivative of the GPD log-likelihood function (i.e., the natural logarithmic form of the GPD likelihood function) not having a closed form. Hence, it is difficult to determine the solution for the local-maximum value using ordinary analytical methods.

The difficulty in determining this solution can be overcome by numerical methods. Identifying a solution that maximizes the GPD log-likelihood function is similar to finding the root of its first derivative when it has a value equal to zero. Therefore, these numerical methods are categorized as root-finding methods. Several numerical root-finding methods, such as Newton, secant, and fixed-point iteration, can be used, of which fixed-point iteration is the simplest basic method [16]. This method determines the root as the fixed point of a function [17], denoting that only slight manipulation of the function's form is required and that further differentiation is not needed, as is the case in the Newton or secant methods [18].

Thus, this study's research objective is to develop novel fixed-point iteration equations to estimate GPD parameters via the ML method. The literature review in Section II demonstrates that fixed-point iteration equations have not been studied in terms of GPD parameter estimation. Moreover, we also review the equations for accuracy and goodness of fit using case studies on the disaster loss data from the United States between 1977 and 2021. This study will allow professionals and practitioners to estimate GPD parameters in a way that the design of the extreme

Manuscript received 13 February 2023; revised 26 November 2023.

This work was supported in part by the Rector and the Directorate of Research and Community Service (DRPM) Universitas Padjadjaran.

Sri Purwani is a lecturer in the Mathematics Department, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jl. Raya Bandung-Sumedang km 21 Jatinangor, Sumedang 45363, Indonesia (corresponding author to provide e-mail: sri.purwani@unpad.ac.id).

Riza Andrian Ibrahim is a PhD candidate in the Mathematics Department, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jl. Raya Bandung-Sumedang km 21 Jatinangor, Sumedang 45363, Indonesia (e-mail: riza17005@mail.unpad.ac.id).

distribution of a random variable can be carried out directly and easily using the POT method.

II. Literature Review

Several studies were carried out on estimating GPD parameters between 1987 and 2000. For example, Hosking and Wallis [19] estimated GPD parameters using the method of moments (MOM) and the probability-weighted moment (PMW) approaches. They also compared these methods using Monte Carlo simulations. Moharram et al. [20] used the least squares (LS) method to estimate GPD parameters, after which they compared the estimation results with MOM and PWM methods based on root mean square error. The comparison showed that the estimated results deduced from the LS method were generally smaller than those deduced using the MOM and PWM methods. Furthermore, Singh and Guo [21] implemented the principle of maximum entropy (POME) method to estimate GPD parameters and found that the estimated results were better than those for the MOM and PMW methods in several data ranges. In addition, Lin and Wang [22] estimated GPD parameters using the ML method on censored data.

Several studies on estimating GPD parameters were also carried out from 2000 to 2022, with de Zea Bermudez and Turkman [23] first estimating GPD parameters using the Bayesian method and showing that this method was suitable for small data sizes in simulations. Juarez and Schucany [24] applied the minimum density power divergence estimator method to estimate GPD parameters, which demonstrated that the coefficients for the robustness and the efficiency of the estimated GPD parameters can be controlled. Zhang [25] estimated GPD parameters using likelihood-moment estimation (LME) and showed the high asymptotic efficiency that resulted from the parameter estimates. Zhang and Stephens [26] used the ML and quasi-Bayesian methods to estimate the GPD parameters. Following this, de Zea Bermudez and Kotz [27] estimated GPD parameters by combining robust and Bayesian techniques. Moreover, Mackay et al. [7] estimated GPD parameters using LME, which has a low sensitivity for selecting a threshold value and is suitable for data sizes lower than 500. Song and Song [28] estimated GPD parameters using the nonlinear least square method. Wang and Chen [29] introduced a hybrid estimation method that minimized the statistical value of the Anderson-Darling (AD) test and maximized the GPD log-likelihood function. In addition, El-Sagheer et al. [30] assessed GPD parameters on progressive and censored failure data using the Bayesian method, the performance of which was determined by Monte Carlo simulations. Finally, Form and Ratnasingam [31] estimated GPD parameters using the elemental percentile method.

None of these studies estimated GPD parameters using the ML method in the form of fixed-point iterations. Thus, this study makes a novel contribution to the literature.

III. Estimating GPD Parameters Using the ML Method

The GPD was first introduced by Pickands [32] and Balkema and de Haan [33]. Smith [34], Davison [35], Smith [36], and van Montfort and Witter [3] then contributed to the development of this method through their research. For example, X is a random variable with a fat-tailed CDF, μ is the threshold value between the extreme and non-extreme values of X, and $Y = X - \mu$ with $X > \mu$ denotes an extreme-excess random variable. Thus, if the value of μ is large, then the CDF of Y can be approximated by the CDF of the GPD, which is expressed as follows:

$$F_Y(y) = \begin{cases} 1-\left(1+\xi \dfrac{y}{\sigma}\right)^{-\frac{1}{\xi}} & ;\xi \neq 0 \\ 1-e^{-\frac{y}{\sigma}} & ;\xi = 0 \end{cases}. \tag{1}$$

In this equation, $\sigma > 0$ and $\xi \in \mathbb{R}$ represent the scale and shape parameters, respectively. Moreover, the probability density function (PDF) of Y, approximated by the PDF of the GPD, is stated as follows:

$$f_Y(y) = \begin{cases} \dfrac{1}{\sigma}\left(1+\xi \dfrac{y}{\sigma}\right)^{-\left(\frac{1}{\xi}+1\right)} & ;\xi \neq 0 \\ \dfrac{1}{\sigma} e^{-\frac{y}{\sigma}} & ;\xi = 0 \end{cases}. \tag{2}$$

The domain of Y is $[0,\infty)$ when $\xi \geq 0$ and is $\left[0,-\dfrac{\sigma}{\xi}\right]$ when $\xi < 0$ [37]. Based on its shape parameters, Salvadori et al. [2] stated that there are three types of GPD (see Table I). The shape parameter values in this study are assumed to be non-zero to limit the scope of the study.

The most popular method for estimating GPD parameters is the ML method. Herein, X_j with $j = 1, 2, \ldots, n$ are independent and identically distributed random variables for which the PDF tails are fat. Moreover $Y_i = X_i - u$, where $X_i > u$ $X_i > u$, $i = 1, 2, ..., m$, represents the extreme-excess random variables that are independent and identically distributed in the GPD. Thus, the likelihood function of Y_i is expressed as follows:

$$L(\boldsymbol{\theta}|\mathbf{y}) = \prod_{i=1}^{m} f_Y(y_i \mid \boldsymbol{\theta}) = \prod_{i=1}^{m} \frac{1}{\sigma}\left(1+\xi \frac{y_i}{\sigma}\right)^{-\left(1+\frac{1}{\xi}\right)}, \tag{3}$$

TABLE I
Three Types of GPD

Types of GPD	Shape Parameter Values	Sidelight
Type I	$\xi = 0$	GPD is equivalent to exponential distribution
Type II	$\xi > 0$	GPD is equivalent to Pareto distribution
Type III	$\xi < 0$	GPD is equivalent to beta distribution

where $\boldsymbol{\theta} = \begin{bmatrix} \sigma \\ \xi \end{bmatrix} \in \boldsymbol{\Theta}$ represents a parameter vector of GPD in space $\boldsymbol{\Theta}$ and $\mathbf{y} = \left[y_1, \ldots, y_m\right]^T$. The GPD log-likelihood function of Y_i is stated as follows:

$$\ln\left[L\left(\boldsymbol{\theta}|\mathbf{y}\right)\right] = -m\ln\left(\sigma\right) - \left(1 + \frac{1}{\xi}\right)\sum_{i=1}^{m}\ln\left[\left(1 + \xi\frac{y_i}{\sigma}\right)\right]. \quad (4)$$

The first derivative of (4) with respect to each element of $\boldsymbol{\theta}$ is expressed as follows:

$$\nabla\ln\left[L\left(\boldsymbol{\theta}\,|\,\mathbf{y}\right)\right] = \begin{bmatrix} \dfrac{\partial\ln\left[L\left(\boldsymbol{\theta}\,|\,\mathbf{y}\right)\right]}{\partial\sigma} \\ \dfrac{\partial\ln\left[L\left(\boldsymbol{\theta}|\mathbf{y}\right)\right]}{\partial\xi} \end{bmatrix}$$

$$= \begin{bmatrix} -\dfrac{m}{\sigma} + \dfrac{1+\xi}{\sigma}\displaystyle\sum_{i=1}^{m}\dfrac{y_i}{\sigma + \xi y_i} \\ \dfrac{1}{\xi^2}\displaystyle\sum_{i=1}^{m}\ln\left(1 + \xi\dfrac{y_i}{\sigma}\right) - \dfrac{1+\xi}{\xi}\displaystyle\sum_{i=1}^{m}\dfrac{y_i}{\sigma + \xi y_i} \end{bmatrix}. \quad (5)$$

The root of (5), which is denoted by $\boldsymbol{\alpha} = \begin{bmatrix} \hat{\sigma} \\ \hat{\xi} \end{bmatrix}$, fulfils the following equation:

$$\begin{bmatrix} -\dfrac{m}{\hat{\sigma}} + \dfrac{1+\hat{\xi}}{\hat{\sigma}}\displaystyle\sum_{i=1}^{m}\dfrac{y_i}{\hat{\sigma} + \hat{\xi} y_i} \\ \dfrac{1}{\hat{\xi}^2}\displaystyle\sum_{i=1}^{m}\ln\left(1 + \hat{\xi}\dfrac{y_i}{\hat{\sigma}}\right) - \dfrac{1+\hat{\xi}}{\hat{\xi}}\displaystyle\sum_{i=1}^{m}\dfrac{y_i}{\hat{\sigma} + \hat{\xi} y_i} \end{bmatrix} = 0. \quad (6)$$

Thus, the solution for (6) is difficult to estimate analytically since it does not have a closed form, but a numerical method is a fitting alternative for solving this problem. Therefore, this study uses the fixed-point iteration method to do so.

IV. Main Results

A. Designing Possible Fixed-Point Iterations in GPD Parameter Estimation using the ML Method

Fixed-point iteration is a method used to numerically determine the root of a function. This method can also be used to determine the maximum solution of a function, which is the root of its first derivative. Assuming that the function for which the maximum solution is sought is (4), using the fixed-point iteration method, (6) is first modified to form the equation $\boldsymbol{\theta} = g\left(\boldsymbol{\theta}|\mathbf{y}\right)$. This transformation results in the following fixed-point iteration equation:

$$\boldsymbol{\theta}_t = g\left(\boldsymbol{\theta}_{t-1}|\mathbf{y}\right) = \begin{bmatrix} g_1\left(\boldsymbol{\theta}_{t-1}|\mathbf{y}\right) \\ g_2\left(\boldsymbol{\theta}_{t-1}|\mathbf{y}\right) \end{bmatrix}, \quad (7)$$

where $\boldsymbol{\theta}_t = \begin{bmatrix} \sigma_t \\ \xi_t \end{bmatrix}$ and $\boldsymbol{\theta}_0$ is given. The error value of each iteration, denoted by ε_t, is determined using the following equation:

$$\varepsilon_t = \left\|\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1}\right\|. \quad (8)$$

If ε_t is less than ε, where ε is the tolerance error, the iteration is stopped. To start the iteration for the GPD parameter estimation, (7) must first be designed. It is important that the equation satisfies the unbiased estimator and convergence requirements. Firstly, the equations for the possible iterations of σ_t and ξ_t were designed separately. These possible iterations are presented in Table II, which demonstrates that there were eight possible schemes of combination for σ_t and ξ_t. Secondly, each scheme's unbiased estimators and convergence were examined.

B. Determining the Unbiased Estimators of the GPD Parameter Estimation

The unbiasedness of $\boldsymbol{\alpha}$ as the estimator of $\boldsymbol{\theta}$ was investigated by determining each element's unbiased estimator. $\boldsymbol{\alpha}$ is an unbiased estimator of $\boldsymbol{\theta}$ if

$$Bias\left(\hat{\sigma},\sigma\right) = E\left(\hat{\sigma}\right) - \sigma = 0 \text{ and } Bias\left(\hat{\xi},\xi\right) = E\left(\hat{\xi}\right) - \xi = 0, \quad (9)$$

where $Bias\left(\hat{\sigma},\sigma\right)$ is the unbiased estimator of both $\hat{\sigma}$ and

TABLE II
THE POSSIBLE ITERATIONS OF σ_t AND ξ_t

No.	$\boldsymbol{\sigma_t}$	$\boldsymbol{\xi_t}$
1.	$\dfrac{\sigma_{t-1}\left(\xi_{t-1}+1\right)}{m}\displaystyle\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}$	$\dfrac{\sum_{i=1}^{m}\ln\left(1+\xi_{t-1}\dfrac{y_i}{\sigma_{t-1}}\right)}{\xi_{t-1}\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}}-1$
2.	$\dfrac{m\sigma_{t-1}}{\left(\xi_{t-1}+1\right)\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}}$	$\sqrt{\dfrac{\xi_{t-1}\sum_{i=1}^{m}\ln\left(1+\xi_{t-1}\dfrac{y_i}{\sigma_{t-1}}\right)}{\left(1+\xi_{t-1}\right)\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}}}$
3.		$-\sqrt{\dfrac{\xi_{t-1}\sum_{i=1}^{m}\ln\left(1+\xi_{t-1}\dfrac{y_i}{\sigma_{t-1}}\right)}{\left(1+\xi_{t-1}\right)\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}}}$
4.		$\dfrac{\sum_{i=1}^{m}\ln\left(1+\xi_{t-1}\dfrac{y_i}{\sigma_{t-1}}\right)}{\left(1+\xi_{t-1}\right)\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y}}$

σ as well as where $Bias\left(\hat{\xi},\xi\right)$ is the unbiased estimator of $\hat{\xi}$. If $\sigma_{p,t}$ with $p=1,2$ is the iteration σ_t of type p in Table II, and if $\xi_{q,t}$ with $q=1,2,3,4$ represents the iteration ξ_t of type q in Table II, the unbiased estimator of $\sigma_{1,t}$ is determined first. Thus, the unbiased estimator of $\sigma_{1,t}$ must be zero.

$$Bias(\hat{\sigma},\sigma)=E(\hat{\sigma})-\sigma=E\left[\frac{\sigma(\xi+1)}{m}\sum_{i=1}^{m}\frac{Y_i}{\sigma+\xi Y_i}\right]-\sigma$$

$$=\frac{\sigma(\xi+1)}{m}\times m\times E\left(\frac{Y}{\sigma+\xi Y}\right)-\sigma=\sigma(\xi+1)\times E\left(\frac{Y}{\sigma+\xi Y}\right)-\sigma. \quad (10)$$

In this equation, $E\left(\frac{Y}{\sigma+\xi Y}\right)$ must be determined.

$$E\left(\frac{Y}{\sigma+\xi Y}\right)=\int_0^\infty\left(\frac{y}{\sigma+\xi y}\right)\frac{1}{\sigma}\left(1+\frac{\xi}{\sigma}y\right)^{-\left(\frac{1}{\xi}+1\right)}dy=\frac{1}{\xi+1}. \quad (11)$$

Therefore,

$$Bias(\hat{\sigma},\sigma)=\sigma(\xi+1)\times E\left(\frac{Y}{\sigma+\xi Y}\right)-\sigma=\sigma(\xi+1)\times\frac{1}{\xi+1}-\sigma=0. \quad (12)$$

If $Bias(\hat{\sigma},\sigma)=0$, $\hat{\sigma}$ is the unbiased estimator of σ, the unbiased estimator of $\xi_{1,t}$ must be reviewed. Herein, this unbiased estimator of $\xi_{1,t}$ must be zero.

$$Bias\left(\hat{\xi},\xi\right)=E\left(\hat{\xi}\right)-\xi=E\left[\frac{\sum_{i=1}^{m}\ln\left(1+\xi\frac{Y_i}{\sigma}\right)}{\xi\sum_{i=1}^{m}\frac{Y_i}{\sigma+\xi Y_i}}-1\right]-\xi$$

$$=\frac{mE\left[\ln\left(1+\xi\frac{Y}{\sigma}\right)\right]}{\xi mE\left(\frac{Y}{\sigma+\xi Y}\right)}-1-\xi=\frac{E\left[\ln\left(1+\xi\frac{Y}{\sigma}\right)\right]}{\xi E\left(\frac{Y}{\sigma+\xi Y}\right)}-1-\xi. \quad (13)$$

Following this, $E\left[\ln\left(1+\xi\frac{Y}{\sigma}\right)\right]$. must be determined:

$$E\left[\ln\left(1+\xi\frac{Y}{\sigma}\right)\right]=\int_0^\infty\ln\left(1+\xi\frac{y}{\sigma}\right)\frac{1}{\sigma}\left(1+\frac{\xi}{\sigma}y\right)^{-\left(\frac{1}{\xi}+1\right)}dy=\xi. \quad (14)$$

Therefore,

$$Bias\left(\hat{\xi},\xi\right)=\frac{E\left[\ln\left(1+\xi\frac{Y}{\sigma}\right)\right]}{\xi E\left(\frac{Y}{\sigma+\xi Y}\right)}-1-\xi=\frac{\xi}{\xi\frac{1}{\xi+1}}-1-\xi=0. \quad (15)$$

As $Bias\left(\hat{\xi},\xi\right)=0$, $\hat{\xi}$ is the unbiased estimator of ξ, the processes of determining the unbiased estimators of $\sigma_{2,t}$, $\xi_{2,t},\xi_{3,t}$, and $\xi_{4,t}$ are generally the same as those for $\sigma_{1,t}$ and $\xi_{1,t}$. The results of these checks demonstrate that only $\xi_{3,t}$ is biased since it results in $Bias\left(\hat{\xi},\xi\right)=-2\xi\neq 0$. Therefore, $\xi_{3,t}$ was omitted from the schemes, resulting in the iteration $\boldsymbol{\theta}_t=\begin{bmatrix}\sigma_{p,t}\\ \xi_{q^*,t}\end{bmatrix}$, in which $p=1,2$ and $q^*=1,2,4,$ are the unbiased estimators of $\boldsymbol{\theta}$.

C. Determining the Convergence of the GPD Parameters' Iterations

The convergence of the iteration equation $\boldsymbol{\theta}_t=g\left(\boldsymbol{\theta}_{t-1}\right)$ was determined using Theorem 1.

Theorem 1. Let $g(\boldsymbol{\theta}|\boldsymbol{y})$ and $g'(\boldsymbol{\theta}|\boldsymbol{y})$ be continuous for a two-dimensional closed set $\mathcal{D}\in\mathbb{R}^2$, and let

$$g(\boldsymbol{\theta}|\mathbf{y})\in\mathcal{D}\ \forall\boldsymbol{\theta}\in\mathcal{D}. \quad (16)$$

Moreover, if

$$\max_{\boldsymbol{\theta}\in\mathcal{D}}\left\|g'(\boldsymbol{\theta}|\mathbf{y})\right\|\leq\gamma<1, \quad (17)$$

then there is a unique solution $\boldsymbol{\alpha}$ for $\boldsymbol{\theta}=g(\boldsymbol{\theta}|\boldsymbol{y})$, and the iteration $\boldsymbol{\theta}_t$ will converge to $\boldsymbol{\alpha}$ for any initial estimators $\boldsymbol{\theta}_0\in\mathcal{D}$.

Proof. (16) shows that there is at least one solution $\boldsymbol{\alpha}$ for $\boldsymbol{\theta}=g(\boldsymbol{\theta}|\boldsymbol{y})$ in $\mathcal{D}$. Thus, the solution for $\boldsymbol{\alpha}$ is unique. Contrastingly, if there are two solutions for $\boldsymbol{\theta}=g(\boldsymbol{\theta}|\boldsymbol{y})$, namely $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, using the mean value theorem, it can be seen that

$$\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2=g(\boldsymbol{\alpha}_1|\mathbf{y})-g(\boldsymbol{\alpha}_2|\mathbf{y})=g'(\mathbf{c}|\mathbf{y})(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2), \quad (18)$$

where $\mathbf{c}\in\boldsymbol{\alpha}_1\times\boldsymbol{\alpha}_2$. If the absolute value of (18) is taken and (17) is then used, we obtain the following:

$$\begin{aligned}\left\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\right\|&=\left\|g(\boldsymbol{\alpha}_1|\mathbf{y})-g(\boldsymbol{\alpha}_2|\mathbf{y})\right\|=\left\|g'(\mathbf{c}|\mathbf{y})(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)\right\|\\&=\left\|g'(\mathbf{c}|\mathbf{y})\right\|\left\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\right\|\leq\gamma\left\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\right\|.\end{aligned} \quad (19)$$

Note that

$$\begin{aligned}&\left\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\right\|\leq\gamma\left\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\right\|,\\&(1-\gamma)\left\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\right\|\leq 0.\end{aligned} \quad (20)$$

Since $\gamma<1$, $\boldsymbol{\alpha}_1$ is the same as $\boldsymbol{\alpha}_2$, the solution for $\boldsymbol{\alpha}$ is unique. Following this, the iteration $\boldsymbol{\theta}_t$ will converge to $\boldsymbol{\alpha}$ for any initial estimators $\boldsymbol{\theta}_0\in\mathcal{D}$. Based on (16), it is inductively clear that $\boldsymbol{\theta}_0 \in \mathcal{D}, \boldsymbol{\theta}_t \in \mathcal{D}$. This is similar to (18), which demonstrates that

$$\begin{aligned}\left\|\boldsymbol{\alpha}-\boldsymbol{\theta}_{t+1}\right\| &= \left\|g(\boldsymbol{\alpha}|\mathbf{y})-g(\boldsymbol{\theta}_t|\mathbf{y})\right\| = \left\|g'(\mathbf{d}|\mathbf{y})(\boldsymbol{\alpha}-\boldsymbol{\theta}_t)\right\| \\ &= \left\|g'(\mathbf{d}|\mathbf{y})\right\|\left\|(\boldsymbol{\alpha}-\boldsymbol{\theta}_t)\right\| \le \gamma\left\|\boldsymbol{\alpha}-\boldsymbol{\theta}_t\right\|, \end{aligned} \tag{21}$$

where $\mathbf{d} \in \boldsymbol{\alpha} \times \boldsymbol{\theta}_t$. Thus, equation (21) can be inductively generalized as follows:

$$\left\|\boldsymbol{\alpha}-\boldsymbol{\theta}_t\right\| \le \gamma^t \left\|\boldsymbol{\alpha}-\boldsymbol{\theta}_0\right\|. \tag{22}$$

Since $\gamma < 1$, $\gamma^t \left\|\boldsymbol{\alpha}-\boldsymbol{\theta}_0\right\| \to 0$ as $t \to \infty$, this demonstrates that $\boldsymbol{\theta}_t \to \boldsymbol{\alpha}$ as $t \to \infty$.

In practice, the convergence of iteration $\boldsymbol{\theta}_t$ using Theorem 1 is rarely determined. This is because it is not easy to find the closed set $\mathcal{D} \in \mathbb{R}^2$ that satisfies (16). Therefore, Theorem 1 can be practically used by applying Corollary 1.

Corollary 1. If $g(\boldsymbol{\theta}|\boldsymbol{y})$ and $g'(\boldsymbol{\theta}|\boldsymbol{y})$ are continuous for a two-dimensional closed set $\mathcal{D} \in \mathbb{R}^2$, in which the solution $\boldsymbol{\alpha}$ of $\boldsymbol{\theta} = g(\boldsymbol{\theta}|\boldsymbol{y})$ is included, and if

$$\left\|g'(\boldsymbol{\alpha}|\mathbf{y})\right\| < 1, \tag{23}$$

there is a closed neighbourhood $\mathcal{K}$ for $\boldsymbol{\alpha}$, thereby proving the results of Theorem 1.

Proof. Using assumption (23), we can create a neighbourhood for $\mathcal{K}$ with a radius $\varepsilon > 0$ around $\boldsymbol{\alpha}$, $\mathcal{K} = \left\{\mathbf{k} \in \mathbb{R}^2 \middle| \boldsymbol{\alpha}-\mathbf{k} \le \varepsilon\right\}$, which satisfies the following conditions:

$$\max_{\mathbf{k} \in \mathcal{K}} \left\|g'(\mathbf{k}|\mathbf{y})\right\| \le \gamma < 1. \tag{24}$$

It must be shown that $g(\mathbf{k}|\mathbf{y}) \in \mathcal{K}\ \forall \mathbf{k} \in \mathcal{K}$. If any element $\mathbf{k} \in \mathcal{K}$ is selected, it must be noted that

$$\begin{aligned}\left\|\boldsymbol{\alpha}-g(\mathbf{k} \mid \mathbf{y})\right\| &= \left\|g(\boldsymbol{\alpha} \mid \mathbf{y})-g(\mathbf{k} \mid \mathbf{y})\right\| \\ &= \left\|g'(\mathbf{s}|\mathbf{y})\right\|\left\|\boldsymbol{\alpha}-\mathbf{k}\right\| \le \gamma\varepsilon \le \varepsilon, \end{aligned} \tag{25}$$

where $\mathbf{s} \in \boldsymbol{\alpha} \times \mathbf{k}$. Equation (25) indicates that $g(\mathbf{k}|\mathbf{y}) \in \mathcal{K}$. Finally, the iteration $\boldsymbol{\theta}_t$ converging to $\boldsymbol{\alpha}$ for any initial estimator $\boldsymbol{\theta}_0 \in \mathcal{K}$ can be justified in the same ways as the methods used in Theorem 1, except by replacing $\mathcal{D}$ with $\mathcal{K}$. □

In this study, the convergence for both $\sigma_{p,t}$ and $\xi_{q^*,t}$ was determined using Corollary 1. To facilitate this determination, we used the data on the disaster losses in the United States from 1977 to 2021 in billions of dollars, with a data size of 280 being obtained. This information was accessed at https://www.emdat.be (accessed 8 November 2022). The data were confirmed to have extreme values since the probability function tail was fat and sloped more to the right. The fat tail of the probability function can be seen in the data's Kurtosis (80.1725), which was >3, and the slope can be seen in the data skewness (8.0095), which was positive.

To obtain the vector $\mathbf{y}$, the threshold value μ was first determined. Herein, the threshold value was determined using the Kurtosis method. The threshold value selection algorithm is presented in Fig. 1.

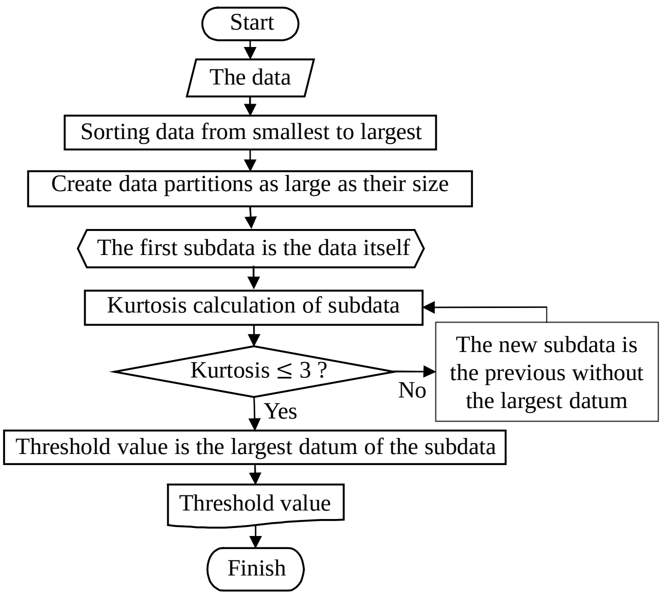

Fig. 1. The threshold value selection algorithm via the kurtosis method.

The data were first partitioned into 280 sub-data sets. The first sub-data set was the data itself, and the second sub-data set was the former sub-data set without the most prominent datum. The threshold value was the most prominent datum of the largest sub-data set that had a Kurtosis of ≤3. The threshold value was determined using the Kurtosis method in Scilab v. 6.1.1. The threshold value obtained was 4.4881, and the most prominent datum was the 196th sub-data set, which was the largest sub-data set, with a Kurtosis of <3, namely 2.9399. The vector of $\mathbf{y}$ is presented as follows:

$$\mathbf{y} = \begin{bmatrix} 0.0118 \\ 0.1754 \\ \vdots \\ 204.2792 \end{bmatrix}. \tag{26}$$

The next stage comprised running each iteration pair. The algorithm for determining the convergence with the data is presented in Fig. 2. The initial vector estimator and the tolerance error were determined in advance. According to Hosking and Wallis [19], the initial vector can be determined using the following equation:

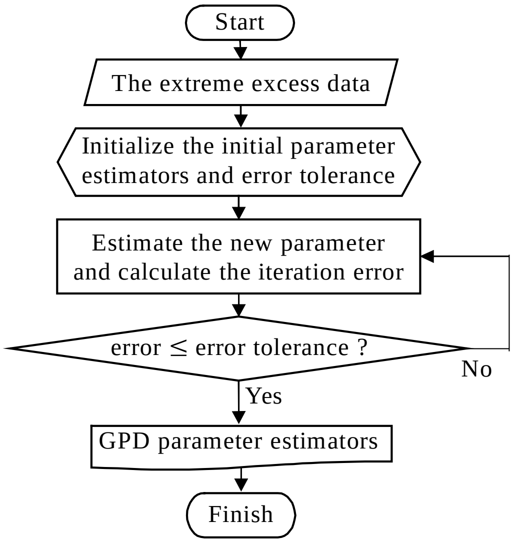

Fig. 2. Checking iteration convergence algorithm.

$$\boldsymbol{\theta}_0=\left[\frac{2\lambda\bar{y}}{\bar{y}-2\lambda},\frac{\bar{y}}{\bar{y}-2\lambda}-2\right]^T, \tag{27}$$

where $\bar{y}$ represents the average of $\mathbf{y}$, and

$$\lambda=\frac{1}{m}\sum_{i=1}^{m}\frac{m-i}{m-1}y_i. \tag{28}$$

Alternatively, the initial vector can also be determined using the following equation [15]:

$$\boldsymbol{\theta}_0=\left[\frac{\bar{y}}{2}\left(\frac{\bar{y}}{s^2}+1\right),\frac{1}{2}\left(\frac{\bar{y}}{s^2}-1\right)\right]^T, \tag{29}$$

where s^2 represents the variance of $\mathbf{y}$. This study determined the initial vector estimator using equation (27). Following this, the selected tolerance error was $\varepsilon=1\times10^{-6}$. The iteration pairs involved were the vector $\begin{bmatrix}\sigma_{p,t}\\ \xi_{q^*,t}\end{bmatrix}$ with $p=1,2$ and $q^*=1,2,4$. The iteration results for each $\begin{bmatrix}\sigma_{p,t}\\ \xi_{q^*,t}\end{bmatrix}$ are presented in Table III, which shows that only iterations $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ converge, whereas the other iterations diverge. The fixed point $\boldsymbol{\alpha}$ obtained was almost the same for each iteration. Therefore, each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ has a $\left\|g'(\boldsymbol{\alpha}|\mathbf{y})\right\|$ value <1. Thus, based on Corollary 1, there are closed neighbourhoods for $\boldsymbol{\alpha}$, so each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ converges to $\boldsymbol{\alpha}$.

TABLE III
THE ITERATION RESULTS OBTAINED

Iteration	Iteration Number	$\boldsymbol{\alpha}$	$\left\|g'(\boldsymbol{\alpha}\mid\mathbf{y})\right\|$
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{1,t}\end{bmatrix}$	18	$\begin{bmatrix}4.9712\\ 0.6866\end{bmatrix}$	0.7780
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{2,t}\end{bmatrix}$	41	$\begin{bmatrix}4.9793\\ 0.6839\end{bmatrix}$	0.9249
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{4,t}\end{bmatrix}$	27	$\begin{bmatrix}4.9743\\ 0.6856\end{bmatrix}$	0.8656
$\begin{bmatrix}\sigma_{2,t}\\ \xi_{1,t}\end{bmatrix}$	diverges	-	-
$\begin{bmatrix}\sigma_{2,t}\\ \xi_{2,t}\end{bmatrix}$	diverges	-	-
$\begin{bmatrix}\sigma_{2,t}\\ \xi_{4,t}\end{bmatrix}$	diverges	-	-

Following this, the diverging values of $\left\|g'(\boldsymbol{\alpha}|\mathbf{y})\right\|$ for each iteration $\begin{bmatrix}\sigma_{2,t}\\ \xi_{q^*,t}\end{bmatrix}$ were analysed (see Table III). Assume that the $\boldsymbol{\alpha}$ for each iteration $\begin{bmatrix}\sigma_{2,t}\\ \xi_{q^*,t}\end{bmatrix}$ is $\begin{bmatrix}4.9749\\ 0.6857\end{bmatrix}$, which is the mean of $\boldsymbol{\alpha}$ in iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$. The $\left\|g'(\boldsymbol{\alpha}|\mathbf{y})\right\|$ values for each iteration $\begin{bmatrix}\sigma_{2,t}\\ \xi_{q^*,t}\end{bmatrix}$ are presented in Table IV, which demonstrates that the $\left\|g'(\boldsymbol{\alpha}|\mathbf{y})\right\|$ value of each iteration $\begin{bmatrix}\sigma_{2,t}\\ \xi_{q^*,t}\end{bmatrix}$ had a value of >1. According to Corollary 1, this indicates that there is no closed neighbourhood for $\boldsymbol{\alpha}$, in which each iteration $\begin{bmatrix}\sigma_{2,t}\\ \xi_{q^*,t}\end{bmatrix}$ converges to $\boldsymbol{\alpha}$.

TABLE IV
$\left\|g'(\boldsymbol{\alpha}|\mathbf{y})\right\|$ VALUES FOR THE DIVERGE-ITERATIONS

Iteration	$\left\|g'(\boldsymbol{\alpha}\mid\mathbf{y})\right\|$
$\begin{bmatrix}\sigma_{2,t}\\ \xi_{1,t}\end{bmatrix}$	1.3450
$\begin{bmatrix}\sigma_{2,t}\\ \xi_{2,t}\end{bmatrix}$	1.3941
$\begin{bmatrix}\sigma_{2,t}\\ \xi_{4,t}\end{bmatrix}$	1.3748

D. Fixed-Point Iterations in GPD Parameter Estimation using the ML Method

These results indicate that three iterations meet the unbiased estimator and convergence conditions, namely

$\begin{bmatrix} \sigma_{1,t} \\ \xi_{q^*,t} \end{bmatrix}$ with $q^* = 1,2,4$. Expanding on this, the iteration $\begin{bmatrix} \sigma_{1,t} \\ \xi_{q^*,t} \end{bmatrix}$ for each q^* is expressed as follows:

$$\boldsymbol{\theta}_t = \begin{bmatrix} \sigma_{1,t} \\ \xi_{1,t} \end{bmatrix} = \begin{bmatrix} \dfrac{\sigma_{t-1}\left(\xi_{t-1}+1\right)}{m}\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i} \\ \dfrac{\sum_{i=1}^{m}\ln\left(1+\xi_{t-1}\dfrac{y_i}{\sigma_{t-1}}\right)}{\xi_{t-1}\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}}-1 \end{bmatrix}, \tag{30}$$

$$\boldsymbol{\theta}_t = \begin{bmatrix} \sigma_{1,t} \\ \xi_{2,t} \end{bmatrix} = \begin{bmatrix} \dfrac{\sigma_{t-1}\left(\xi_{t-1}+1\right)}{m}\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i} \\ \sqrt{\dfrac{\xi_{t-1}\sum_{i=1}^{m}\ln\left(1+\xi_{t-1}\dfrac{y_i}{\sigma_{t-1}}\right)}{\left(1+\xi_{t-1}\right)\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}}} \end{bmatrix}, \tag{31}$$

and

$$\boldsymbol{\theta}_t = \begin{bmatrix} \sigma_{1,t} \\ \xi_{4,t} \end{bmatrix} = \begin{bmatrix} \dfrac{\sigma_{t-1}\left(\xi_{t-1}+1\right)}{m}\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i} \\ \dfrac{\sum_{i=1}^{m}\ln\left(1+\xi_{t-1}\dfrac{y_i}{\sigma_{t-1}}\right)}{\left(1+\xi_{t-1}\right)\sum_{i=1}^{m}\dfrac{y_i}{\sigma_{t-1}+\xi_{t-1}y_i}} \end{bmatrix}. \tag{32}$$

Therefore, professionals and practitioners can use Equations (30), (31), and (32) to estimate GPD parameters using the ML method.

V. Discussion

A. The Unbiased Estimator Is Not Necessarily Convergent

The determination of the unbiased estimators of the GPD parameters in Section IV (B) shows that iteration $\begin{bmatrix} \sigma_{2,t} \\ \xi_{q^*,t} \end{bmatrix}$ with $q^* = 1,2,4$ is an unbiased estimate of $\boldsymbol{\theta}$, and it is justified by each unbiased estimator that is equal to zero. However, after determining the convergence of iteration $\begin{bmatrix} \sigma_{2,t} \\ \xi_{q^*,t} \end{bmatrix}$ in Section IV (C), only iteration $\begin{bmatrix} \sigma_{2,t} \\ \xi_{q^*,t} \end{bmatrix}$ was found to not meet the convergence criterion. Thus, an unbiased estimator of the GPD parameter is not necessarily convergent.

B. The Effect of the $\left\|\boldsymbol{g}'\left(\boldsymbol{\alpha}|\mathbf{y}\right)\right\|$ Value on Convergence Speed

As can be seen in Table III, the value of $\left\|\boldsymbol{g}'\left(\boldsymbol{\alpha}|\mathbf{y}\right)\right\|$ has a positive relationship with the number of iterations required to converge to the solution. The closer to zero the value of $\left\|\boldsymbol{g}'\left(\boldsymbol{\alpha}|\mathbf{y}\right)\right\|$ is, the fewer iterations it takes to converge, and vice versa.

C. The Estimated GPD Parameters' Accuracy and Goodness of Fit

The estimated GPD parameters' accuracy can be determined through various measures. Herein, the mean absolute percentage error (MAPE) was used. The accuracy of the GPD parameter estimation was considered 'good' if the MAPE value was between 10%–20%, and it was considered 'very good' if the MAPE values were <10% [41], [42]. The MAPE of the estimated GPD parameters was calculated using the following equation:

$$\text{MAPE} = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{F_Y^{-1}\left(F_e\left(y_i\right)\right)-y_i}{y_i}\right|\times 100\%, \tag{33}$$

where $F_e(\cdot)$ represented the empirical distribution function and $F_Y^{-1}(\cdot)$ represented the inverse function of (1). The MAPE of the estimated GPD parameters from each iteration $\begin{bmatrix} \sigma_{1,t} \\ \xi_{q^*,t} \end{bmatrix}$ is presented in Table V, which shows that this MAPE value was between 10%–20%. Therefore, the GPD parameter estimation for each iteration $\begin{bmatrix} \sigma_{1,t} \\ \xi_{q^*,t} \end{bmatrix}$ was classified as 'good'.

TABLE V
MAPE of Estimated GPD Parameters of the Converge-Iterations

Iteration	MAPE (%)
$\begin{bmatrix} \sigma_{1,t} \\ \xi_{1,t} \end{bmatrix}$	17.3639
$\begin{bmatrix} \sigma_{1,t} \\ \xi_{2,t} \end{bmatrix}$	17.4081
$\begin{bmatrix} \sigma_{1,t} \\ \xi_{4,t} \end{bmatrix}$	17.3805

Furthermore, the GPD parameters' goodness of fit can be determined visually or formally. Visually, determining goodness of fit can be conducted with probability-probability plots (P-P plots), where set $\mathcal{A} = \left\{\left(F_e\left(y_i\right), F_Y\left(y_i\right)\right), i = 1,2,\ldots,m\right\}$ is placed into a Cartesian diagram. The estimated GPD parameters are considered visually fit for describing the data if the data for set $\mathcal{A}$ are scattered around a line with a gradient of 1 [43].

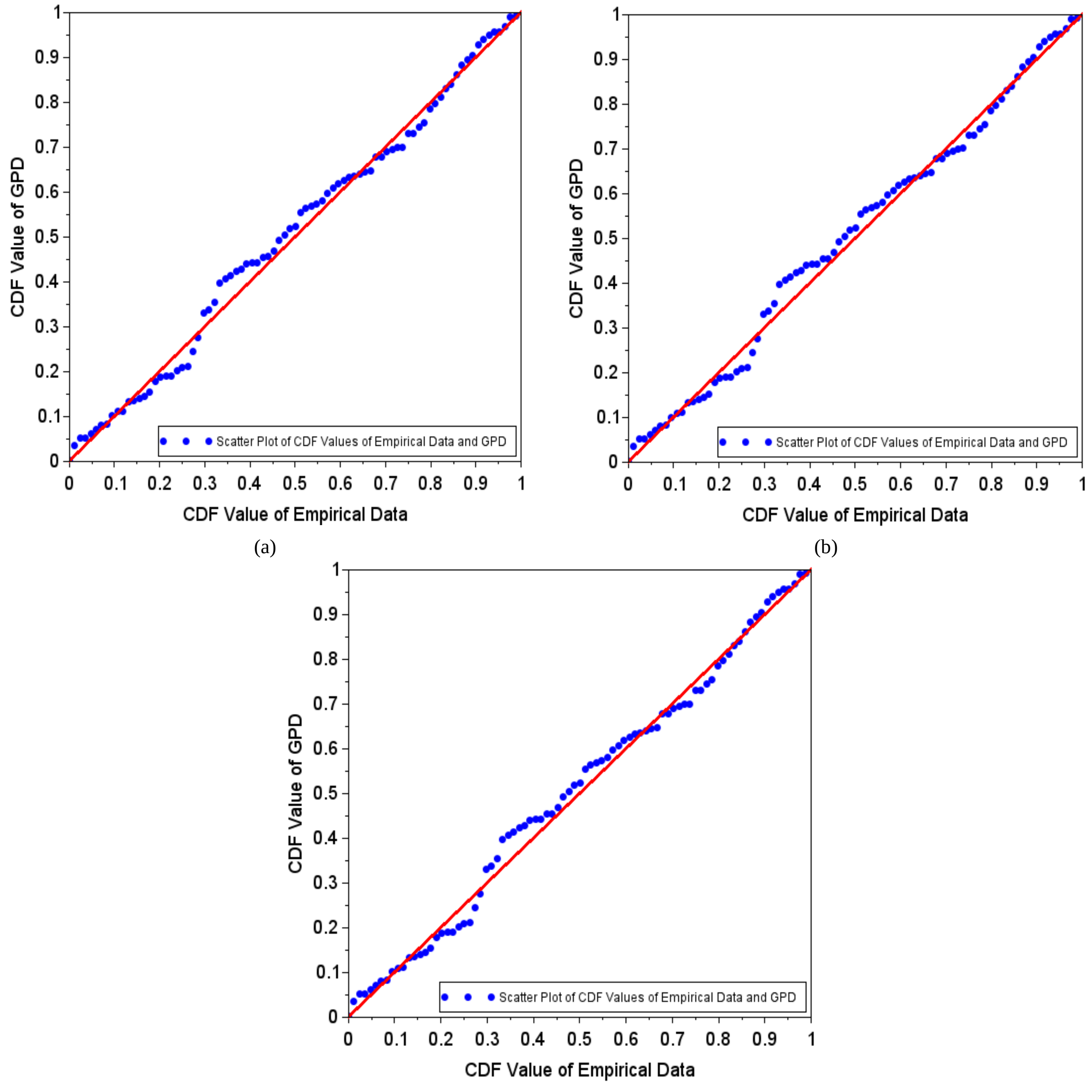

Fig. 3. P-P Plots for Estimated GPD Parameters of Iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{1,t}\end{bmatrix}$, $\begin{bmatrix}\sigma_{1,t}\\ \xi_{2,t}\end{bmatrix}$, and $\begin{bmatrix}\sigma_{1,t}\\ \xi_{4,t}\end{bmatrix}$ on (a), (b), and (c) respectively.

The P-P plots for each estimated GPD parameter of iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ are presented in Fig. 3, which demonstrates that the scatterplot of set $\mathcal{A}$ for each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ is centred around a red line with a gradient of one. Thus, the GPD parameters estimated by iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ are visually fit to describe the data distribution. Following this, the goodness of fit can be formally reviewed using the Kolmogorov-Smirnov (KS) and AD tests. The statistical value of the KS test is determined using the following equation [44], [45], [46], [47]:

$$\tau = \sup_{y}\left|F_e\left(y\right) - F_Y\left(y\right)\right|. \tag{34}$$

If the statistical value of the KS test is greater than the critical value $\left(\tau_\alpha\right)$, the estimated GPD parameter is not fit for describing the data distribution, and vice versa. The statistical value of the AD test can be determined using the following equation [48]:

$$A = -\frac{1}{m}\sum_{i=1}^{m}\left(2i-1\right)\left\{\ln\left[F_Y\left(y_i\right)\right] + \ln\left[1 - F_Y\left(y_{m-i+1}\right)\right]\right\} - m. \tag{35}$$

If the statistical value of the AD test is greater than the critical value $\left(A_{\alpha}\right)$, the estimated GPD parameter is not fit for describing the data distribution, and vice versa. With a significance level of 0.05, the statistical values of the KS and AD tests and their respective critical values for the estimated GPD parameters for each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ are presented in Table VI, which shows that the statistical values of the KS and AD tests for each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ were smaller than their respective critical values. Therefore, the GPD parameters of each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ were formally fit for describing the data distribution.

TABLE VI
THE KS AND AD STATISTICAL TEST VALUES AND THEIR RESPECTIVE CRITICAL VALUES OF THE ESTIMATED GPD PARAMETERS FOR THE CONVERGE-ITERATIONS

Iteration	τ	$\tau_{0.05}$	A	$A_{0.05}$
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{1,t}\end{bmatrix}$	0.0624		0.3162	
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{2,t}\end{bmatrix}$	0.0627	0.1460	0.3164	2.5018
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{4,t}\end{bmatrix}$	0.0625		0.3163	

D. The Necessary Conditions for Reaching the Maximum Solution for the Estimated GPD Parameters

The estimated GPD parameters for each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ must meet the necessary conditions to reach the maximum solution of the log-likelihood function (4). The maximum local solution of the GPD log-likelihood function (4) is reached at the critical point $\boldsymbol{\alpha}$ if the Hessian matrix of (4), denoted by $\mathbf{H}\left(\boldsymbol{\theta}\,|\,\mathbf{y}\right)$, at $\boldsymbol{\alpha}$ is a negative definite, namely [49]:

$$\forall \boldsymbol{a}=\begin{bmatrix}a_1\\ a_2\end{bmatrix}\in\mathbb{R}^2-\{0\},\boldsymbol{a}^T\mathbf{H}\left(\boldsymbol{\alpha}|\mathbf{y}\right)\boldsymbol{a}<0. \tag{36}$$

The results from reviewing the necessary conditions for reaching the maximum solution for the estimated GPD parameters for each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ are presented in Table VII, which delineates that the $\boldsymbol{a}^{\mathrm{T}}\mathbf{H}\left(\boldsymbol{\alpha}\mid\mathbf{y}\right)\boldsymbol{a}$ value of each

TABLE VII
CHECKING THE HESSIAN MATRIX OF THE GPD LOG-LIKELIHOOD FUNCTION AT $\boldsymbol{\alpha}$ OF THE CONVERGE-ITERATIONS

Iteration	$\boldsymbol{a}^{\mathrm{T}}\,\mathbf{H}\left(\boldsymbol{\alpha}\mid\mathbf{y}\right)\boldsymbol{a}$ Value
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{1,t}\end{bmatrix}$	$-\left[1.4338\left(a_1+2.9331a_2\right)^2+29.7677a_2^2\right]<0$
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{2,t}\end{bmatrix}$	$-\left[1.4315\left(a_1+2.9473a_2\right)^2+30.0156a_2^2\right]<0$
$\begin{bmatrix}\sigma_{1,t}\\ \xi_{4,t}\end{bmatrix}$	$-\left[1.4327\left(a_1+2.9384a_2\right)^2+29.8554a_2^2\right]<0$

iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ was negative for each $\boldsymbol{a}\in\mathbb{R}^2-\{0\}$. Therefore, every $\boldsymbol{\alpha}$ in each iteration $\begin{bmatrix}\sigma_{1,t}\\ \xi_{q^*,t}\end{bmatrix}$ satisfied the maximum solution of the log-likelihood function (4).

VI. CONCLUSION

This study presented the design of fixed-point iteration equations for estimating GPD parameters using the ML method. The fixed-point iteration method was used because it is the simplest numerical method for determining a function's solution. This method was also used to determine the solution for the first derivative of the GPD log-likelihood function, which does not have a closed form.

In the initial stage of designing the iterations, eight fixed-point iteration equations were obtained (see Table II). After this, the eight fixed-point iteration equations were reviewed to determine their unbiased estimators and convergence, with three of them satisfying these conditions, as presented in equations (30), (31), and (32). Thus, numerical proof was obtained for the assertion that an unbiased iteration equation of the GPD parameter is not necessarily convergent to the solution. Moreover, the closer to zero the value of the first derivative of the iteration at the solution, the fewer iterations it takes to converge. Moreover, it was found that the estimated GPD parameters obtained through the simulation have good accuracy and goodness of fit. The accuracy was demonstrated by the MAPE value of 17%, while the goodness of fit was determined using KS and AD tests, which resulted in values less than their respective critical values. Finally, the estimated GPD parameters met the necessary conditions for the maximum solution of the GPD log-likelihood function.

Therefore, the fixed-point iteration equations designed in this study can facilitate the estimation of GPD parameters when modelling the extreme distribution of random variables. By using these fixed-point iteration equations, professionals and practitioners no longer need to design fixed-point iteration equations when estimating GPD parameters and when reviewing unbiased estimators and convergence.

REFERENCES

[1] S. Thendo, N. Nethengwe, C. Sigauke and H. Chikoore, “Modelling of extremely high rainfall in Limpopo province of South Africa,” *Climate*, vol. 10, no. 3, p. 33, 2022.

[2] G. Salvadori, C. D. Michele, N. T. Kottegoda and R. Rosso, *Extremes in Nature: An Approach Using Copulas*. Dordrecht: Springer, 2007.

[3] M. A. J. van Montfort and J. V. Witter, “Testing exponentiality against generalised Pareto distribution,” *J. Hydrol.*, vol. 78, no. 3-4, pp. 305–315, 1985.

[4] J. Ferrez, A. C. Davison and M. Rebetez, “Extreme temperature analysis under forest cover compared to an open field,” *Agric. For. Meteorol.*, vol. 151, no. 7, pp. 992–1001, 2011.

[5] X. L. Wang, B. Trewin, Y. Feng and D. Jones, “Historical changes in Australian temperature extremes as inferred from extreme value distribution analysis,” *Geophys. Res. Lett.*, vol. 40, no. 3, pp. 573–578, 2013.

[6] F. Ashkar and C. N. Tatsambon, “Revisiting some estimation methods for the generalized Pareto distribution,” *Jour. Hydrol.*, vol. 346, no. 3-4, pp. 136–143, 2007.

[7] E. B. L. Mackay, P. G. Challenor and A. S. Bahaj, “A comparison of estimators for the generalized Pareto distribution,” *Ocean Eng.*, vol. 38, no. 11-12, pp. 1338–1346, 2011.

[8] P. Stansell, “Distributions of extreme wave, crest and trough heights measured in the North Sea,” *Ocean Eng.*, vol. 32, no. 8-9, pp. 1015–1036, 2005.

[9] L. Wei, L. Liu and J. Hou, “Pricing hybrid-triggered catastrophe bonds based on copula-EVT model,” *Quant. Fin. and Econ.*, vol. 6, no. 2, pp. 223–243, 2022.

[10] W. Chao and H. Zou, “Multiple-event catastrophe bond pricing based on CIR-Copula-POT model,” *Discrete Dyn. Nat. Soc.*, vol. 2018, 2018, Art. no. 5068480.

[11] A. A. Zimbidis, N. E. Frangos and A. A. Pantelous, “Modeling earthquake risk via extreme value theory and pricing the respective catastrophe bonds,” *ASTIN Bull.*, vol. 37, no. 1, pp. 163–183, 2007.

[12] J. Shao, A. Pantelous and A. D. Papaioannou, “Catastrophe risk bonds with applications to earthquakes,” *Eur. Actuar. J.*, vol. 5, no. 1, pp. 113–138, 2015.

[13] L. Huyse, R. Chen and J. A. Stamatakos, “Application of generalized Pareto distribution to constrain uncertainty in peak ground accelerations,” *Bull. Seismol. Soc. Am.*, vol. 100, no. 1, pp. 87–101, 2010.

[14] R. A. Ibrahim, S. Sukono and R. Riaman, “Estimation of the extreme distribution model of economic losses due to outbreaks using the POT method with Newton Raphson iteration,” *Int. J. Quant. Res. Mod.*, vol. 2, no. 1, pp. 37–45, 2021.

[15] P. de Zea Bermudez and S. Kotz, “Parameter estimation of the generalized Pareto distribution—Part I,” *J. Stat. Plan. Inference,* vol. 140, no. 6, pp. 1353–1373, 2010.

[16] K. E. Atkinson, *An Introduction to Numerical Analysis*, 2nd ed. New Jersey: John Wiley & Sons, 1989.

[17] K. Atkinson and W. Han, *Elementary Numerical Analysis*, 3rd ed. New Jersey: John Wiley & Sons, 2004.

[18] S. Purwani, A. F. Ridwan, R. A. Hidayana and S. Sukono, “Secant method with Aitken extrapolation outperform Newton-Raphson method in estimating stock implied volatility,” *IAENG Int. J. Comput. Sci,*, vol. 50, no. 2, pp. 368–374, 2023.

[19] J. R. M. Hosking and J. R. Wallis, “Parameter and quantile estimators for generalized Pareto distribution,” *Technometrics*, vol. 29, no. 3, pp. 339–349, 1987.

[20] S. H. Moharram, A. K. Gosain and P. N. Kapoor, “A comparative study for the estimators of the generalized Pareto distribution,” *J. Hydrol.*, vol. 150, no. 1, pp. 169–185, 1993.

[21] V. P. Singh and H. Guo, “Parameter estimation for 2-parameter generalized Pareto distribution by POME,” *Stoch. Hydrol. Hydraul.*, vol. 9, pp. 81–93, 1995.

[22] W. T. Lin and W. Y. Wang, “Estimation for the generalized Pareto distribution with censored data,” *Commun. Stat. Sim. Comput.*, vol. 29, no. 4, pp. 1183–1213, 2000.

[23] P. de Zea Bermudez and M. A. A. Turkman, “Bayesian approach to parameter estimation of the generalized Pareto distribution,” *Test,* vol. 12, no. 1, pp. 259–277, 2003.

[24] S. F. Juarez and W. R. Schucany, “Robust and efficient estimation for the generalized Pareto distribution,” *Extremes,* vol. 7, no. 3, pp. 237–251, 2004.

[25] J. Zhang, “Likelihood moment estimation for the generalized Pareto distribution,” *Aust. N. Z. J, Stat.*, vol. 49, no. 1, pp. 69–77, 2007.

[26] J. Zhang and M. Stephens, “A new and efficient estimation method for the generalized Pareto distribution,” *Technometrics*, vol. 51, no. 3, pp. 316–325, 2009.

[27] P. de Zea Bermudez and S. Kotz, “Parameter estimation of the generalized Pareto distribution—Part II,” *J. Stat. Plan. Inference*, vol. 140, no. 6, pp. 1374–1388, 2010.

[28] J. Song and S. Song, “A quantile estimation for massive data with generalized Pareto distribution,” *Comput. Stat. Data Anal.,* vol. 56, no. 1, pp. 143–150, 2012.

[29] C. Wang and G. Chen, “A new hybrid estimation method for the generalized Pareto distribution,” *Comm. Stat. Theory Methods,* vol. 45, no. 14, pp. 4285–4294, 2016.

[30] R. M. El-Sagheer, T. M. Jawa and N. Sayed-Ahmed, “Inferences for generalized Pareto distribution based on progressive first-failure censoring scheme,” *Complexity*, vol. 2021, 2021, Art. No. 9325928.

[31] S. G. From and S. Ratnasingam, “Some efficient closed-form estimators of the parameters of the generalized Pareto distribution,” *Environ. Ecol. Stat.*, vol. 29, pp. 827–847, 2022.

[32] J. Pickands, “Statistical inference using extreme order statistics,” *The Ann. Stat.*, vol. 3, no. 1, pp. 119–131, 1975.

[33] A. Balkema and L. de Haan, “Residual life time at great age,” *Ann. Probab.*, vol. 2, no. 5, pp. 792–804, 1974.

[34] R. L. Smith, “Threshold Methods for Sample Extremes,” in *Statistical Extremes and Applications*, J. T. de Oliveira (Ed.), Dordrecht: Springer, 1984, pp. 621–638.

[35] A. C. Davison, “Modelling Excesses over High Thresholds, with an Application,” in *Statistical Extremes and Applications*, J. T. de Oliveira (Ed.), Dordrecht: Springer, 1984, pp. 461–482.

[36] R. L. Smith, “Maximum likelihood estimation in a class of nonregular cases,” *Biometrika*, vol. 72, no. 1, pp. 67–90, 1985.

[37] H. Nagatsuka and N. Balakrishnan, “Efficient likelihood-based inference for generalized Pareto distribution,” *Ann. Instit. Stat. Math.*, vol. 73, no. 6, pp. 1153–1185, 2021.

[38] P. Worapun and A. Kangtunyakarn, “An approximation method for solving fixed points of general system of variational inequalities with convergence theorem and application,” *IAENG Int. J. Appl. Math.*, vol. 51, no. 3, pp. 751-756, Sep. 2021.

[39] Mohammed Barrada, Reda Benkhouya, Mohammed Lahmer, and Idriss Chana, "On the Global Convergence of a New Super Halley's Family for Solving Nonlinear Equations," *Engineering Letters*, vol. 29, no.4, pp1407-1413, 2021.

[40] D. Shah, and M. Sahni, “DMS way of finding the optimum number of iterations for fixed point iteration method,” *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering 2018*, 4-6 July, 2018, London, U.K., pp. 87-89.

[41] P. C. Chang, Y. W. Wang and C. H. Liu, “The development of a weighted evolving fuzzy neural network for PCB sales forecasting,” *Expert Syst. Appl.*, vol. 32, no. 1, pp. 86–96, 2007.

[42] R. A. Ibrahim, Sukono, H. Napitupulu, R. I. Ibrahim, M. D. Johansyah and J. Saputra, “Estimating flood catastrophe bond prices using approximation method of the loss aggregate distribution: Evidence from Indonesia,” *Decis. Sci. Lett.*, vol. 12, no. 2, pp. 179-190, Mar. 2023.

[43] M. Gerson, “The techniques and uses of probability plotting,” *J. R. Stat. Soc.*, vol. 24, no. 4, pp. 235–257, 1975.

[44] A. Luceño, “Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators,” *Comput. Stat. Data Anal.,* vol. 51, no. 2, pp. 904–917, 2005.

[45] M. Raschke, “Opportunities of the minimum Anderson–Darling estimator as a variant of the maximum likelihood method,” *Commun. Stat. Simul. Comput,* vol. 46, no. 9, pp. 6879–6888, 2017.

[46] Sukono, R.A. Ibrahim, M. P. A. Saputra, Y. Hidayat, H. Juahir, I. G. Prihanto and N.B.A. Halim, “Modeling multiple-event catastrophe bond prices involving the trigger event correlation, interest, and inflation rates,” *Mathematics*, vol. 10, no. 24, pp. 4685, Dec. 2022.

[47] R. A. Ibrahim, Sukono and H. Napitupulu, “Multiple-trigger catastrophe bond pricing model and its simulation using numerical methods,” *Mathematics*, vol. 10, no. 9, pp. 1363, Apr. 2022.

[48] Z. G. Ma and C.Q. Ma, “Pricing catastrophe risk bonds: A mixed approximation method,” *Insur. Math. Econ.*, vol. 52, no. 2, pp. 243–254, 2013.

[49] R. A. Ibrahim, Sukono, H. Napitupulu and R. I. Ibrahim, “How to price catastrophe bond for sustainable earthquake funding? A

systematic review of the pricing framework," *Sustainability*, vol. 15, no. 9, pp. 7705, May. 2023.