An Estimation Approach of Classical Tail Index and an Application to Hydrology Assessment

Jutamas Boonradsamee,Winai Bodhisuwan, and Uraiwan Jaroengeratikun

***Abstract*—The estimators commonly used to estimate the tail index is Hill's estimator, which has a strong bias estimation of the tail index parameter. For heavy-tailed distributions, Hill's estimator is widely used to model long-tail phenomena occurring in many disciplines, including finance, insurance, telecommunication, meteorology and hydrology. However, Hill's estimators are sensitive to the number of order statistics, which are important for estimating. Therefore, choosing the most suitable value of order statistic k amount will lead to an effective estimation of Hill's estimator. In this research, we investigated the method for selecting k in Hill's estimator by using the method of the quantile. A simulation study on Pareto distributed data was conducted to compare the efficiency of this Hill's estimator with the Maximum likelihood (MLE) method and the criterion was the root mean square error (RMSE). The results found that Hill's estimator yielded RMSE lower than MLE. Finally, an application to hydrology data was presented. We applied the method of quantile estimator for selecting k in Hill's estimator for the approximation of the shape parameter at the tail end of the daily maximum volume of water flowing into the Khun Dan Prakan Chon Dam in Thailand in 2017 for 104 days. The results obtained from an application study were consistent with the simulation study results, therefore Hill's estimator by using a quantile method was efficient for an estimate of the shape parameter of hydrology data.**

***Index Terms*—Extreme Value Theory, Pareto Distribution, Tail Index Estimator.**

I. INTRODUCTION

THE Extreme Value Theory (EVT) is proposed for administering with modelling the effect of extreme events. In one studies the distribution of the maximum and minimum values of random variables as the sample size increases. It is based on statistical methods which are designed to estimate probabilities of extreme events, which methods use a limited range of data of such events that occurred in the past and are suitable for predicting events even more extreme than those previously observed. Therefore it gives the necessary guidance to evaluate the possibility of risk in improbable events. EVT is widely used to estimate the probability and return period of high water levels. The probability of extreme values event is relatively large are characteristics of phenomena of the heavy tails, namely heavy-tailed distribution.

Manuscript received March 23, 2021; revised April 8, 2021. This work was supported in part by by graduate college, King Mongkuts University of Technology North Bangkok, Bangkok, 10800, Thailand.

J.Boonradsamee is a PhD candidate of the department of Applied Statistics,Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangsue, Bangkok, 10800,Thailand (e-mail: jutamas.r@rmutsv.ac.th)

W.Bodhisuwan is an Associate Professor of the department of Statistics, Faculty of Science, Kasertsart University, Chatuchak, Bangkok, 10900, Thailand (e-mail: fsciwnb@ku.ac.th)

U.Jaroengeratikun is an Assistant Professor of the department of Applied Statistics,Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangsue, Bangkok, 10800, Thailand (corresponding author, e-mail: uraiwan.j@sci.kmutnb.ac.th)

Let $X_1, X_2, ..., X_n$ be a sample of independent and identically distribution (iid.) function, therefore the distribution of this function is right heavy tail when,

$$P(X > x) = \overline{F}(x) = 1 - F(x) = L(x)x^{-\alpha}, \quad x > 0, \qquad (1)$$

where $\alpha > 0$ is the index of regular variation and $L(x) > 0$ is a slowly varying function [5].

$$\lim_{x \to 0} \frac{L(tx)}{L(t)} = 1 \qquad (2)$$

The tail $(\alpha > 0)$ exponent controls the rate of decay of $F(x)$ and hence characterizes its tail behaviour. Another way to express equation (1) is that $1 - F(x)$ is regularly varying with index α . A distribution $F(x)$, is concentrated in $[0, \infty)$. As of α, $\alpha > 0$, if

$$\lim_{x \to 0} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}, \quad x > 0 \qquad (3)$$

The expression on the right characterizes regularly varying functions, which means they are power functions multiplied by slowly varying functions. The above stated power function has a negative index of variation $-1/\xi$ which says that the tail decays with rate, $\alpha = 1/\xi$. Therefore the parameter of interest is directly the tail index α which is the reciprocal of the shape parameter . Distributions for which $\xi > 0$ belong to the Frechet-Pareto class. They are often referred to as Pareto-type distribution or heavy-tailed distribution. The estimators of the tail index were presented in several studies such as Pickands estimator [9], the moment estimator [15], smooHill estimator, Hill's estimator in alt scale, smooHill estimator in alt scale [8], which one of the most popular estimators is Hill's estimator [3]. Hill's estimator of the tail index has stressed the importance to draw inference about the behaviour of the tails. The method is based on order statistics (k) and a non parametric approach is conditioned upon the values of order statistics which exceed a certain threshold and a problem of choosing the appropriate number k of upper order statistics to construct the estimator arises. This problem becomes particularly critical because the Hill's estimator has proved to be very sensitive to the choice of , which selects the optimal a value of k will lead to an effective estimate of the Hill's estimator.

This research used the modified Hill's estimator by a new approach to selecting optimal k for an alternative for choosing k presented by Boonradsamee et al. (2021) [12], which notion is using the method of quantile estimator on the interval of the stability region of Hill's plot. The information in this study drawn from a Pareto distribution

ISBN: 978-988-14049-2-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

in part of the simulation and the daily maximum volume of water flowing into the Khun Dan Prakan Chon Dam in Thailand for the application part. The goal of this research is to show the estimation of shape parameter at the tail end of the distribution of information by additional techniques by a new approach to selecting the optimal k of Hill's estimator.

II. Literature Review

In this section, some methods from the literature are presented which are used throughout the research.

A. The Generalised Extreme Value (GEV) Distribution

The generalized extreme value (GEV) distribution is a group of distributions for extreme value theory. The GEV can consists of the Gumbel, Fréchet and Weibull families of distributions. The distribution function of the GEV is given by

$$H_{\xi}(x) = \begin{cases} exp{-1}\,(1+\xi x)^{-1/\xi} & \text{if } \xi \neq 0 \\ exp\{-e^{-x}\} & \text{if } \xi = 0 \end{cases}$$

Hence, the support of H_{ξ} corresponds to
$x \in \mathbb{R}$ for $\xi = 0$, $x < -\frac{1}{\xi} = \alpha$ for $\xi < 0$, $x > -\frac{1}{\xi} = \alpha$ for $\xi > 0$

By defining $H_{\xi,\mu,\sigma} := H_{\xi}\left(\frac{x-\mu}{\sigma}\right)$ we obtain a three-parameter family with location parameter $\mu \in \mathbb{R}$, scale parameter $\sigma > 0$ and shape parameter $\xi \in \mathbb{R}$. The shape parameter ξ plays an important role as it characterizes the distribution of H_{ξ} as one of the three extreme value distributions named above; Fréchet $(\xi = \frac{1}{\alpha} > 0)$, Weibull $(\xi = -\frac{1}{\alpha} < 0)$, Gumbel $(\xi = 0)$. H_{ξ} is nothing else than a shifted and scaled version of these three distributions, which makes it easier to estimate parameters of a distribution of a sample of extremal random variables. Note that H_{ξ} is a continuous function of ξ if x is fixed.

B. The generalized Pareto distribution (GPD)

The GPD is often used to provide a goodfit to extreme of data. The the cumulative distribution function (CDF) of the GPD is defined as

$$G_{\xi}(x) = \begin{cases} 1-(1+\xi x)^{-1/\xi}, & \text{if} \quad \xi \neq 0 \\ 1-e^{-x}, & \text{if} \quad \xi = 0 \end{cases}$$

Where $x \geqslant 0$ for $\xi \geq 0$, and $0 < x \leq -\frac{1}{\xi}$ for $\xi < 0$

By defining $G_{\xi,\mu,\beta} := G_{\xi}\left(\frac{x-\mu}{\beta}\right)$ we get a three-parameter family with location parameter $\mu \in \mathbb{R}$, scale parameter $\beta > 0$ and shape parameter $\xi \in \mathbb{R}$.

In practice mostly assume $\mu = 0$ and therefore get,

$$G_{\xi,\beta}(x) = 1 - \left(1+\xi\frac{x}{\beta}\right)^{-1/\xi}$$

$$\text{where} \quad D = \begin{cases} [0,\infty), & \text{if} \quad \xi \geq 0 \\ \left[0, -\frac{\beta}{\xi}\right), & \text{if} \quad \xi < 0 \end{cases}$$

C. Parameter Estimation

Fitting the GEV to the data, we would like to mention a parameter estimation using the Maximum Likelihood Estimation (MLE) method [1]. This is a method that is accepted as a robust estimation and a wide variety of applications (see [13]-[14]). It is can be established as follows:

Step 1. consider the Probability Density Function (PDF) of GEV.

Step 2. Create the likelihood function of GEV for n sample sizes, $X_1, X_2, ..., X_n$. It can be written by,

$$L(\mu,\sigma,\xi) = \frac{1}{\sigma^n}\prod_{i=1}^{n}\left(1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right)^{\frac{-1}{\xi}-1}$$
$$\exp\left(-\sum_{i=1}^{n}\left(1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right)^{\frac{-1}{\xi}}\right) \quad (4)$$

Step 3. Take a log-likelihood function of the function. Explain the function obtained from the following steps.

$$l(\mu,\sigma,\xi) = n\log\sigma - \left(1-\frac{1}{\xi}\right)\sum_{i=1}^{n}\log\left(1-\xi\left(\frac{x_i-\mu}{\xi}\right)\right)$$
$$-\sum_{i=1}^{n}\left(1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right)^{\frac{-1}{\xi}} \quad (5)$$

Step 4. Using the Log- likelihood Function to find partial derivation to compare with each parameter is set to zero. Then solve the equation to get a parameter estimate

D. Hill's estimator

Hill's estimator was invented by a statistician Bruce M.Hill in 1975 [3]. This estimator has introduced a simple general approach to inference about the tail behaviour of the Pareto type distribution, which has a power law form with regularly varying tails as mentioned in equation (1) and (2). The Hill's estimator is one of the popular estimators of the extreme value index (EVI) or tail index which allowed exibility in the lower tail but that ensures the power law behaviour dominates the upper tail. Clearly, this model does not have such flexible upper tail behaviour as the GPD, but it is an important special case in many applications and since a wide range of techniques has been developed for both tail index and tail fraction estimation.

Let $X_1, X_2, ..., X_n$ are iid. random sample from a distribution $F(x)$ with sample size n and $X_{(1)} > X_{(2)} > ... > X_{(n)}$ are the order statistics of sample. The Hill's estimator used for $\xi = \dfrac{1}{\alpha}$, where $\xi > 0$, should be noticed that the estimator can be obtained by different approaches, e.g. the MLE approach, the regular variation approach or the mean excess function approach. This assures its functionality. The MLE approach will be shown here in a few steps. Again, we denote $X_{(1)} > X_{(2)} > ... > X_{(n)}$ as the ordered sample. From F is to be in the maximum domain of attraction of ϕ_{α}, $(F \in \text{MDA}(\phi_{\alpha}))$, where ϕ_{α} denote as distribution of X. We know that the tail of F is regularly varying and we assume that $c = 1$, then it holds for $X \sim F$:

$P(X > x) = \overline{F}(x) = x^{-\alpha}, x \geq 1, \Longrightarrow$ for $Y = \log X$ we have $P(Y > y) = e^{-\alpha y}$

This can be generalized by

ISBN: 978-988-14049-2-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$P(X>x)=\overline{F}(x)=cx^{-\alpha}=\left(\frac{x}{u}\right)^{-\alpha}$ for $x\geq u>0, c=u^{\alpha}$ and $Y=\log\left(\frac{x}{u}\right)$

Then Y is exponentially distributed with its parameter estimation, ξ, and by MLE we get,

$$\widehat{\xi}=\frac{1}{\widehat{\alpha}}=\frac{1}{n}\sum_{i=1}^{k}\log X_{(n-i+1)}-\log(u) \quad (6)$$

But as $X\sim F$ ofthen only holds for the upper extremes, we have to do a MLE on the k-th upper order statistics. If we choose either $u=X_k$ or $k=\{i\leq n; X_i\leq u\}$, we obtain the Hill's estimator for ξ is determined by

$$\widehat{\xi}^{Hill}=\frac{1}{k}\sum_{i=1}^{k}(\log X_{(n-i+1)}-\log X_{(n-k)}) \quad (7)$$

Moreover, we get the estimation for c by $\widehat{c}=\frac{k}{n}X_k^{1/\widehat{\xi}^{Hill}}$, so the tail probability is given by

$$\widehat{\overline{F}}(x)=\frac{k}{n}\left(\frac{x}{X_{(k)}}\right)^{-\frac{1}{\widehat{\xi}^{Hill}}}$$

and the corresponding p-quantile by

$$\widehat{x}_p=\left(\frac{n}{k}(1-p)\right)^{\widehat{\xi}^{Hill}}X_{(k)} \quad (8)$$

The Hill's estimator strongly depends on the select for k. It is important to mention that $\widehat{\xi}$ is a consistent estimator for the tail index only if $k\to\infty$ and $k/n\to 0$ as $n\to\infty$ was proved by [11], and asymptotically normal with mean ξ and variance ξ^2/k, $N(\xi,\xi^2/k)$. If one uses a too small k, the estimator has large variance, however for too large k, the estimator is likely to be biased.

$$\sqrt{k}(\widehat{\xi}(n,k)-\xi)\xrightarrow{d}N(0,\xi^2) \quad (9)$$

Hill plot

In practice, we can select k in Hill's estimator by the graph of varying Hill shape estimates depending on the cut-off level, as already described by Hill [3] called Hill plot.The value of α or $1/\xi$ is obtained by identifying a stable region in graph of Hill plot. [7], The Hill plot is as follows

$$\{(k,\widehat{\xi}^{Hill}), 1\leq k\leq n-1\} \quad (10)$$

The stability of the estimate is desired thus a stable region in the graph will be considered. Markéta Pokorná,(2016) [4] explained that is suggested to look for the stable part in the upper $1-5\%$ of the order statistics or we can estimate a value of $\alpha=\widehat{\xi}^{-1}$ from the stable regime of this plot. How the decision of choice of the threshold is made is not specified by all authors. A generally accepted rule of thumb is that the tail should consists of $5-10\%$ of the entire sample and hence the threshold is set to be in that region. It sholud not be higher than $10-15\%$ and 10% seems to be a frequently used limit (Rocco, (2011), Nystrom and Skoglund (2002), McNeil and Frey (2000) , see [4]). Hill plot is helpful when the data comes directly from Pareto or close to Pareto distributed random variable.

The Hill plot provides a clear evidence of the value of the estimator (Drees et al., 2000) [2]. The problem is that the graph is volatile and it is not easy to decide what the estimate should be. The sample size may just be too small.

Hill plot is another tool in shape parameter determination. Hill (1975) [3] proposed an estimator of ξ when $\xi>0$ (Fréchet Case). A Hill plot is constructed such that estimated ξ is plotted as a function of k upper order statistics of the threshold u. A threshold is seclected from the plot where the shape parameter ξ is fairly stable. A difficulty of the Hill's estimator is the ambiguity of the value k.

In this paper, we used the new selecting k by using the quantile estimator type 8 or $\widehat{Q}_8(p)$ finding optimal k on Hill plot [12] . Hyndman and Fan (1996) [10] investigating their motivation and some of their properties. There conclude by recommending that the median-unbiased estimator is used because it has most of the desirable properties of a quantile estimator and can be defined independently of the underlying distribution.

The $\widehat{Q}_8(p)$ is the theoretical of being medain unbiased. To compute the pth quantile with method of $\widehat{Q}_8(p)$ use linear interpolation. For each method we choose a constant determined by the sample quantile type, m is $(p+1)/3$, see reference [10]. If j is chosen such that $\frac{j-m}{n}\leq p\leq\frac{j-m+1}{n}$, we can define the knots as (j,x_j) and $(j+1,x_{j+1})$. Since the value selection of the optimal k must be made from the first interval stationary region of the Hill plot, which k is order statistics so that we find the techniques for choosing the cut point dividing the range of the observations in a sample. Therefore, the quantile estimator can be used as well and by the properties of quantiles estimate we use the $\widehat{Q}_8(p)$ for selecting k on Hill plot, which we will describe adoption in section 3.

III. Simulation Study

For this study, we used the algorithm for finding a new approach to choose optimal k [12], which used to selecting shape parameter estimator (ξ). We have found a flexible way to choose ξ, it will be presented on this section. Fist of all we simulated datasets of event random (x) from Pareto distribution with scale parameter $\lambda=1$ and shape parameter $\xi=1$, there are three sizes such as small size is 50,medium size is 100, and the large size is 150. We estimated shape parameter of the random variables by using Hill's estimator ($\widehat{\xi}^{Hill}$). Subsequently, we plotted $\widehat{\xi}^{Hill}$ against order different k, in the equation (7). Then we consider the interval of the stable region and used quantiles estimator type 8 or $\widehat{Q}_8(p)$ with $p=0.75$ in each case for selected a suitable k. Afterwards, we found $\widehat{\xi}^{Hill}$.

For the Fig.1, we consider a random sample of three categorical sizes from Pareto distribution with parameters $\lambda=1$. The scatter plots, which are showing right-skewed and has a long tail. The Pareto distribution is the heavy-tailed distribution, see more in Natalia (2007) [6]. Afterwards, we can check a well model fits to simulated data through built-in diagnostic plots, shown in Fig.1, which the graphical diagnostic plots imply behaviour and hence that the data comes from non-normal distribution and fat tails, suggest that to GEV be a good approximation for tail of distribution and it can be estimated by the Hill's estimator. Thus, the logical next step is to estimate the shape parameter by Hill's estimator.

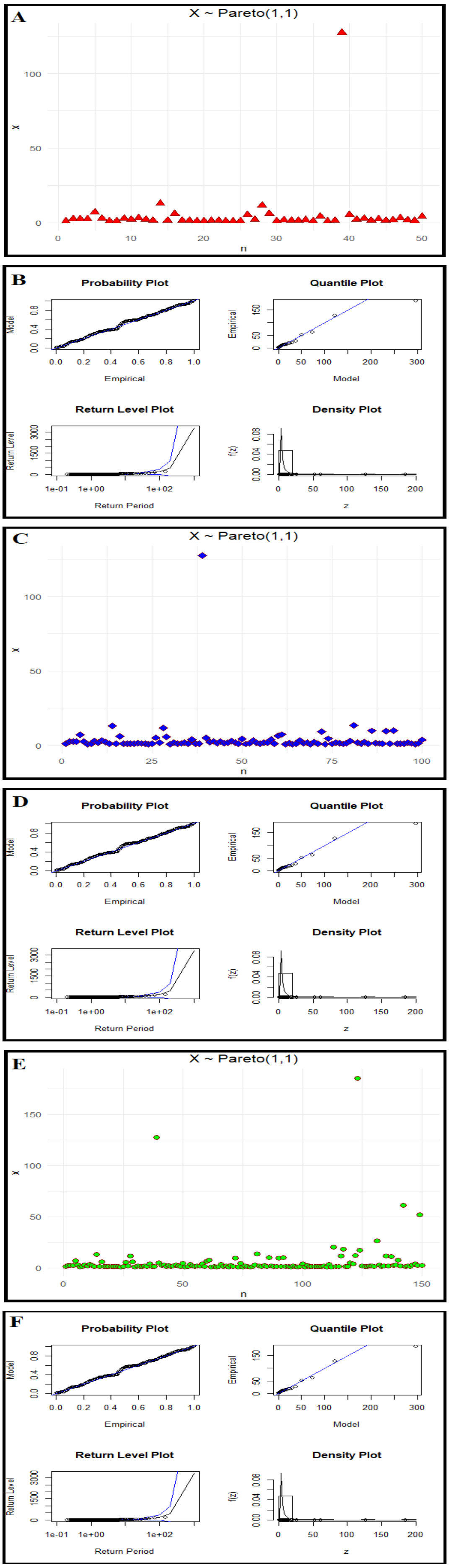

Fig. 1. *the scatter plot and the graph diagonostic of* $X \sim$ *Pareto* (λ, ξ) *where* $\lambda = 1$, $\xi = 1$ *for small size*, $n = 50$ *(A and B), for medium size*, $n = 100$ *(C and D), and for large size*, $n = 150$ *(E and F)*

After then a simulation study of different case is performed. Thus, we can choose the optimal k of the stable region of Hill plot by $\widehat{Q}_8(0.75)$ method, shown in Fig.2.

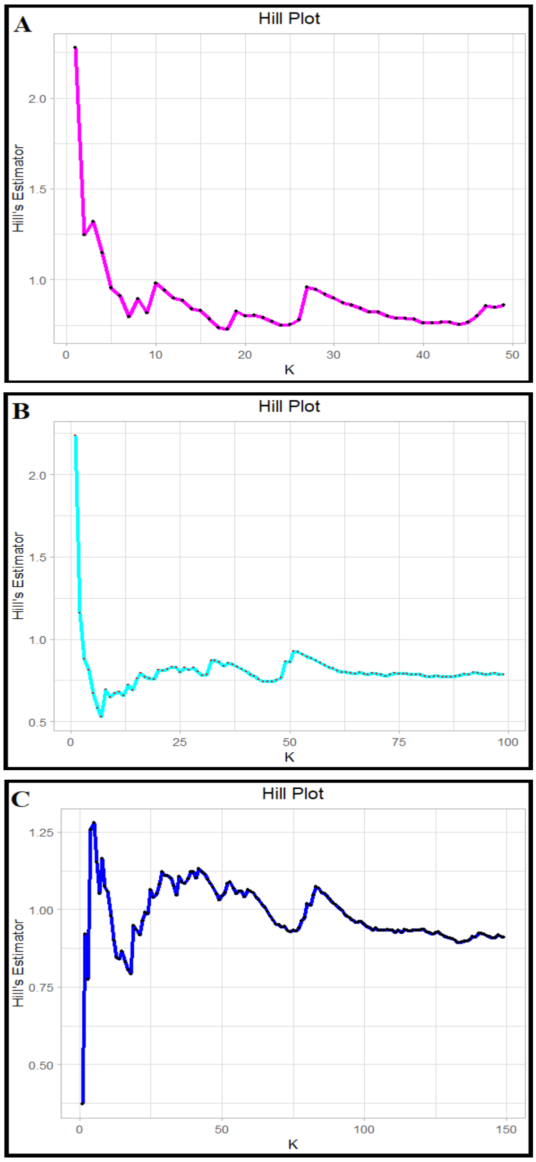

Fig. 2. *the Hill plots show the shape parameter* $\hat{\xi}$ *by Hill's estimator of* $X \sim$ *Pareto* (λ, ξ) *where* $\lambda = 1$, $\xi = 1$ *for small size*, $n = 50$ *(A), for medium size*, $n = 100$ *(B), and for large size*, $n = 150$ *(C).*

When considering the results of comparing the root mean square error (RMSE) and Bias values of the shape parameters of $X \sim$ Pareto (λ, ξ), where $\lambda = 1$, $\xi = 1$ obtained from the MLE and Hill's estimator, it was found that the estimation of the shape parameter by the Hill's estimator has the RMSE and Bias values lower than in all cases of the MLE method, shown in Table I.

IV. Application

Climate change has impacted all over the world and has affected rainfall, which has altered the pattern of rainfall and distribution of raindrops across different river basins in Thailand. As well as impacts from rapid climate change, rainfall and rivers are likely to change significantly from the past to the present, including the severity and frequency of flooding. The construction of the dam is a solution to the

ISBN: 978-988-14049-2-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I
SUMMARIZE THE STUDY ON THE SHAPE PARAMETERS OF $X \sim$ PARETO (λ, ξ), WHERE $\lambda = 1$, $\xi = 1$, OBTAINED FROM THE MLE AND HILL ESTIMATOR, WHICH WAS DETERMINED FROM THE RMSE AND BIAS VALUES

Number of Observation	RMSE		Bias	
	MLE	Hill	MLE	Hill
50	1.9816	0.1721	1.1969	-0.0413
100	2.0522	0.0980	0.5923	-0.5029
150	1.8292	0.0726	0.1142	-0.6674

problem of drainage system management. Therefore, dam building is important to many countries, including Thailand.

We consider an extreme value approach for the data set on a major Dam comprises the main dam and a secondary dam. It is the largest and longest roller compacted concrete (RCC) dam in the world with a total length of 2,720 meters and a height of 93 meters. It receives water flowing from Khao Yai National Park through Heo Narok Waterfall to the reservoir with a storage capacity of 224 million cubic meters. Benefit Water for agriculture, domestic consumption, solution of acid soil, being fish breeding ground and flood alleviation, being a new tourist attraction of Nakhon Nayok, tourists can appreciate Khao Yai National Park from the dam crest, to enjoy the spectacular scenery in front of the Dam and observe the landscape of Nakhon Nayok Province in Thailand behind the Dam.

Fig. 3. *the Khun Dan Prakan Chon Dam in Thailand.*

Table II reports the summary data of the daily maximum volume of water flowing into the Khun Dan Prakan Chon Dam in Thailand in 2017 for 104 days, which shows for the skewness indicate data that are skewed right. In addition, kurtosis is a measure of whether the data are peaked of flat relative to a normal distribution. The data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. We denote x as the daily maximum volume of water flowing into the Dam, then the emirical distribution in Fig.4 (A) shows the scatter plot of the data. (B) shows the diagnostic plots imply behaviour and hence that the data comes from non-normal distribution and fat tails, QQ-plot can help an individual determine how good a fit general class distributions to loss data, which GEV be a good approximation for the tail of distribution and it can be estimated by the Hill's estimator. Thus, the logical next step is to estimate the shape parameter by MLE and Hill's estimator. For the method of Hill's estimator, we selecting the optimal k over the stable region of Hill plot by using

TABLE II
SUMMARY STATISTICS OF THE DAILY MAXIMUM VOLUME OF WATER FLOWING INTO THE KHUN DAN PRAKAN CHON DAM IN THAILAND IN 2017 FOR 104 DAYS

Mean	**S.D.**	**Kurtosis**	**Skew**	**Min**	**Max**	**N**
1097761	1426283	6.9346	2.3322	12930	7959355	104

TABLE III
SUMMARIZE THE STUDY ON THE SHAPE PARAMETERS OF WATER FLOWING INTO THE KHUN DAN PRAKAN CHON DAM IN THAILAND IN 2017 FOR 104 DAYS, OBTAINED FROM THE MLE AND HILL ESTIMATOR OF THE SHAPE PARAMETER

Sample Size	**Shape Parameter**	
	ξ_{MLE}	ξ_{Hill}
104	0.8181	0.5152

$\widehat{Q}_8(0.75))$ [12], which Hill's estimator is shown in the Hill plot in Fig. 5.

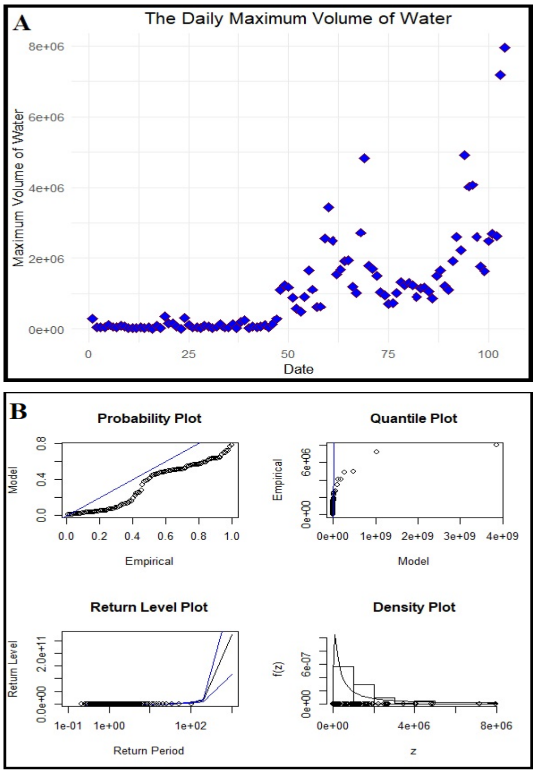

Fig. 4. *(A) the daily maximum volume of water flowing into the Khun Dan Prakan Chon Dam. (B) the daigonostic plot of the daily maximum volume of water flowing into the Khun Dan Prakan Chon Dam. Considering of 104 observations in 2017.*

When considering the results of comparing the RMSE and Bias values of the shape parameters of water flowing into the Khun Dan Prakan Chon Dam in Thailand in 2017 for 104 days obtained from the MLE and Hill's estimator, shown in Table III, it was found that the estimation of the shape parameter by the Hill's estimator gave the RMSE and Bias values lower than in all cases of the MLE method, in Table IV.

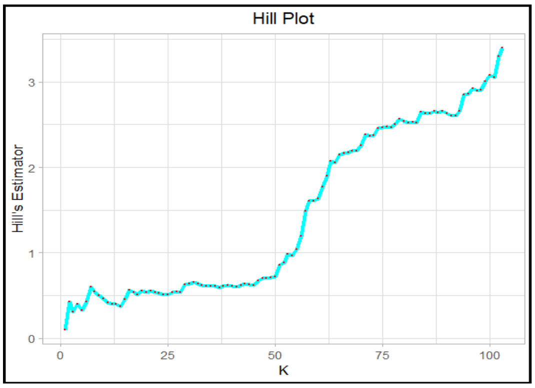

Fig. 5. *the Hill plot show the shape parameter $\widehat{\xi}$ of the daily maximum volume of water flowing into the Khun Dan Prakan Chon Dam*

TABLE IV
SUMMARIZE THE STUDY ON THE SHAPE PARAMETERS OF WATER FLOWING INTO THE KHUN DAN PRAKAN CHON DAM IN THAILAND IN 2017 FOR 104 DAYS, OBTAINED FROM THE MLE AND HILL ESTIMATOR, WHICH WAS DETERMINED FROM THE RMSE AND BIAS VALUES

N of Obs.	RMSE		Bias	
	MLE	Hill	MLE	Hill
104	1.9816	0.1721	1.1969	-0.0413

V. CONCLUSION

The aim of this research is used a new method by using graphics for selecting in Hill's estimator by using type 8 quantile estimator from the stable region of Hill plot, which presented by Boonradsamee et al. (2021) [12], which will be a more flexible alternative for use in Hill's estimator. Expected results: a new approach for select is compatible with Hill's estimator and another estimator, which depends on the order statistics for heavy-tailed distributions is lead to classifying behavior of the tail probability.This study will be an approximation of the shape parameter at the tail end of the distribution of hydrology data.

REFERENCES

[1] S. Coles,"An Introduction to Statistical Modeling of Extreme Values",*pringer,London*, 2001.
[2] H. Drees, L. De Haan, and S. Resnick, "How to make a hill plot",*The Annals of Statistics*, 2000, pp.1833-1855.
[3] B.M. Hill, "A simple general approach to inference about the tail of a distribution", *Annals of Statistics* vol.3, 1975, pp. 1163-1174.
[4] Markéta Pokorná, "Estimation and Application of the Tail Index", *Charles University in Prague*, 2016.
[5] A. J. McNeil, R. Frey, and P. Embrechts,"Quantitative risk management: concepts, techniques and tools-revised edition", *Princeton university press*, 2015.
[6] M. Natalia, "Nonparametric Analysis of Univariate Heavy-Tailed Data".*John Wiley and Sons Ltd, Chichester, England*, 2007.
[7] S.I. Resnick, "Heavy Tail Phenomena: Probabilistic and Statistical Modeling", *Springer, New York, Berlin, Heidelberg, London, Paris, Tokyo*, 2006.
[8] S.I. Resnick, "Discussion of the Danish data on large fire insurance losses", *ASTIN Bulletin: The Journal of the IAA*, vol.27, no.1, 1997, pp. 139-151.
[9] J. Pickands III, "Statistical inference using extreme order statistics", *Annals of statistics*, vol.3, no. 1, 1975 , pp.119-131.
[10] R. J. Hyndman, and Y. Fan, "Sample quantiles in statistical packages", *The American Statistician*, vol. 50, no.4, 1996, pp.361-365.
[11] D. M. Mason, "Laws of large numbers for sums of extreme values", *The Annals of Probability*, 1982, pp.754-764.
[12] J. Boonradsamee, W. Bodhisuwan,and U. Jaroengeratikun, "A New Selecting k Method of Hill's Estimator",*Thai Journal of Mathematics*, 2021, pp.153-163.
[13] S. Puangkaew, and T.Talangtam, "Peaks Over Threshold Model of Generalized Pareto Distributions in Non-life Insurance", 2017
[14] P. Amphanthong,"Modeling Trend of Maximum Rainfall with Probability Distribution of Tah Jene River in Suphanburi Province", *RMUTSB ACADEMIC JOURNAL*, vol.7, no.1, 2019, pp.1-19.
[15] A. L. Dekkers, J. H Einmahl, and L. De Haan, "A moment estimator for the index of an extreme-value distribution". *The Annals of Statistics*, 1989, pp. 1833-1855.