EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 16, No. 4, 2023, 2461-2475
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

A New Class of Generalized Extreme Value Distribution and Application under Alpha Power Transformation Method

Pannawit Khamrot[1], Natthinee Deetae[2,*]

[1] *Department of Mathematics, Faculty of Science and Agricultural Technology, Rajamangala University of Technology Lanna Phitsanulok, Phitsanulok, Thailand.*
[2] *Department of Statistics, Faculty of Science and Technology, Pibulsongkram Rajabhat University, Phitsanulok, Thailand.*

Abstract. This paper presents an expansion of the generalized extreme value distribution to new distribution classes, specifically the Alpha Power Transformation Generalized Extreme Value (APTGEV) distribution. This extension is achieved by combining the Extreme Value theory and the alpha power transformation technique. We employ the maximum likelihood method in conjunction with the Newton-Raphson procedure to estimate the parameters in these proposed distributions. In the final stages of our research, we simulate these new distributions and apply them to real-world data. For this study, we have chosen extreme rainfall data from a weather station in the Si Samrong District Sukhothai Province of Thailand as our dataset. These extended distribution classes are designed to provide greater flexibility and adaptability in understanding complex data patterns, and their application to real-world data offers valuable insights into their effectiveness.

2020 Mathematics Subject Classifications: 62E10, 60G70, 62G32

Key Words and Phrases: Extremes value theory, Generalized extreme value distribution, Alpha Power Transformation Method, Return level, Maximum likelihood estimation

1. Introduction

Extreme Value Theory, also known as Extreme Value Analysis (EVA), is a specialized area within statistics that focuses on the significant departures from the median in probability distributions. Its purpose is to evaluate the likelihood of events surpassing any previously recorded extremities from a specified ordered sample of a particular random variable. EVA finds broad application in various fields, including meteorology, structural engineering, earth sciences, traffic forecasting, geological engineering, telecommunications,

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v16i4.4862

Email addresses: Pk_g@rmutl.ac.th (P. Khamrot), Natthinee@psru.ac.th (N. Deetae)

risk management, finance, insurance, economics, and hydrology. EVA involves estimating parameters from Extreme Value (EV) distributions, commonly regarded as boundary distributions for the highest or lowest values of independent and identically distributed random variables as the sample size grows. These estimates are derived from available historical data. The practical application of EVA can be seen in numerous studies on real-world data. [1, 2, 4–7, 9, 10, 12, 13, 16, 18, 19].

A random variable X follows the Generalized Extreme Value (GEV) distribution, $GEV(x;\mu,\sigma,\xi)$, if its cumulative distribution function (CDF) is given by

$$GEV(x;\mu,\sigma,\xi) = \begin{cases} \exp\{-[1+\xi(x-\mu)/\sigma]^{-1/\xi}\}, & \xi \neq 0, \\ \exp\{-\exp[-(x-\mu)/\sigma]\}, & \xi \to 0, \end{cases} \tag{1}$$

which is defined in the set $\{x : 1+\xi(x-\mu)/\sigma > 0\}$, where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter and $\xi \in \mathbb{R}$ is a shape parameter. The condition for a distribution possessing any of the extreme value distributions is given as follows: $\xi = 0$ for Gumbel distribution, $\xi > 0$ for Fréchet distribution and $\xi < 0$ for Weibull distribution. The corresponding probability density function (PDF) of GEV, $gev(x;\mu,\sigma,\xi)$, is then obtained as

$$gev(x;\mu,\sigma,\xi) = \begin{cases} \sigma^{-1}[1+\xi(x-\mu)/\sigma]^{-(1/\xi)-1}\exp\{-[1+\xi(x-\mu)/\sigma]^{-1/\xi}\}, & \xi \neq 0, \\ \sigma^{-1}\exp[-(x-\mu)/\sigma]\exp\{-\exp[-(x-\mu)/\sigma]\}, & \xi \to 0. \end{cases} \tag{2}$$

The estimates of extreme quantiles z_u^{GEV} of the maximum distribution, known as return level, are then obtained by inverting $GEV(x;\mu,\sigma,\xi)$,

$$z_u^{GEV} = \begin{cases} \mu + \dfrac{\sigma}{\xi}\left\{[-\log(u)^{-\xi} - 1]\right\}, & \xi \neq 0, \\ \mu - \sigma\log[-\log(u)], & \xi \to 0, \end{cases} \tag{3}$$

where $u \in [0,1]$, $u = 1 - T^{-1}$, and T is the number of recurrent times.

In contemporary research, there is a growing interest in generalized distributions. These distributions are considered to be more adaptable as they incorporate one or more additional parameters, typically bearing a specific relationship with other distributions. The method of theoretical advancement often involves generating a new distribution derived from a baseline or existing distribution. This process allows for enhanced flexibility and versatility in statistical modeling, providing a broader scope for researchers to understand and interpret complex data patterns. In this study, the GEV distribution is used as a baseline to seek for new classes of distribution to provide a better fit for extreme data. Mahdavi and Kundu [11] introduced the alpha power transformation (APT) method, which is based on adding a parameter to a family of distributions to improve their flexibility. The APT-G family's cumulative distribution function (CDF) is defined as follows:

$$F(x;\alpha) = \begin{cases} \dfrac{\alpha^{G(x)} - 1}{\alpha - 1}, & \alpha > 0, \alpha \neq 1 \\ G(x), & \alpha > 0, \alpha = 1. \end{cases} \tag{4}$$

The probability density function related to (4) is

$$f(x;\alpha) = \begin{cases} \dfrac{\log \alpha}{\alpha - 1} g(x)\alpha^{G(x)}, & \alpha > 0, \alpha \neq 1 \\ g(x), & \alpha > 0, \alpha = 1. \end{cases} \tag{5}$$

Alpha power has been applied to various distribution theories. In 2017, Nassar et al.[15] present Alpha power Weibull distribution: Properties and applications. Which is a new distribution, and two real data sets are used to illustrate the importance of the proposed distribution. Nadarajah et al.[14] introduced a novel three-parameter distribution and illustrated an application to an ozone data set. In 2019, Ihtisham et al.[8] studied a new distribution referred to as the Alpha-Power Pareto distribution, which is an extra parameter, and two real datasets have been considered to examine the usefulness of the proposed distribution. In 2022, Shrahili et al.[17] introduced two-parameter alpha power transformed moment exponential (APTME) distribution. To examine the practical significance of the APTME distribution and real-world datasets. The above research found that the alpha power method was extended in the distribution theory to obtain a more efficient distribution.

This study proposes new variations of the GEV distribution, referred to as the Alpha Power Transformation Generalized Extreme Value (APTGEV) distribution. The application of these proposed GEV distributions to accurate data appears encouraging, specifically with the inclusion of additional shape parameters. The structure of this paper is organized as follows. Initially, we introduce the concept of an extended distribution. Subsequently, for the proposed distributions, we elaborate on the cumulative distribution function, probability density function, and return level and provide a detailed depiction of the maximum likelihood estimates (MLEs) of all parameters. In the third part, we present a concise study that uses Monte Carlo simulation to evaluate the efficiency and consistency characteristics of the MLEs for the parameters of the APTGEV distribution. In the fourth section, we apply the proposed distributions to analyze rainfall data collected from weather station 373301 in Si Samrong District Sukhothai Province, Thailand, from 1987 to 2021. This analysis is followed by a goodness-of-fit test to verify the model's accuracy. The paper concludes with a summary and interpretation of the results obtained from the study

2. Methodology

In this section, we derived the forms of the cumulative distribution function, probability density function, and the return level for the proposed distributions: alpha power transformation generalized extreme value distribution, including the maximum likelihood estimates (MLEs) of all parameters and confidence interval of return level.

2.1. Alpha power transformation generalized extreme value distribution

A random variable X is said to have APTGEV distribution, denoted by $G(\mu, \sigma, \xi, \alpha)$, with location parameter μ, shape parameter σ, and scale parameter ξ, α, if the PDF and CDF of X for $x \geq 0$ and $\xi \neq 0$ are given by

$$G(x; \mu, \sigma, \xi, \alpha) = \begin{cases} \dfrac{\alpha^{\exp\{-[1+\xi(x-\mu)/\sigma]^{-1/\xi}\}} - 1}{\alpha - 1}, & \alpha > 0, \alpha \neq 1 \\ \exp\{-[1 + \xi(x-\mu)/\sigma]^{-1/\xi}\}, & \alpha > 0, \alpha = 1. \end{cases} \tag{6}$$

The probability density function related to (6) is

$$g(x; \mu, \sigma, \xi, \alpha) = \begin{cases} \dfrac{\log \alpha}{\alpha - 1} \sigma^{-1}[1 + \xi(x-\mu)/\sigma]^{-(1/\xi)-1} & \\ \exp\{-[1 + \xi(x-\mu)/\sigma]^{-1/\xi}\} \alpha^{\exp\{-[1+\xi(x-\mu)/\sigma]^{-1/\xi}\}}, & \alpha > 0, \alpha \neq 1 \\ \sigma^{-1}[1 + \xi(x-\mu)/\sigma]^{-(1/\xi)-1} \exp\{-[1 + \xi(x-\mu)/\sigma]^{-1/\xi}\}, & \alpha > 0, \alpha = 1, \end{cases} \tag{7}$$

and In this case $\xi \to 0$, we have

$$G(x; \mu, \sigma, \alpha) = \begin{cases} \dfrac{\alpha^{\exp\{-\exp[-(x-\mu)/\sigma]\}} - 1}{\alpha - 1}, & \alpha > 0, \alpha \neq 1 \\ \exp\{-\exp[-(x-\mu)/\sigma]\}, & \alpha > 0, \alpha = 1. \end{cases} \tag{8}$$

The probability density function (pdf) corresponding to (8) is given by

$$g(x; \mu, \sigma, \alpha) = \begin{cases} \dfrac{\log \alpha}{\alpha - 1} \sigma^{-1} \exp[-(x-\mu)/\sigma] \exp\{-\exp[-(x-\mu)/\sigma]\} & \\ \alpha^{\exp\{-\exp[-(x-\mu)/\sigma]\}}, & \alpha > 0, \alpha \neq 1 \\ \sigma^{-1} \exp[-(x-\mu)/\sigma] \exp\{-\exp[-(x-\mu)/\sigma]\}, & \alpha > 0, \alpha = 1. \end{cases} \tag{9}$$

The APTGEV distribution includes three types of distribution as $\xi = 0$ is alpha power transformation gumbel distribution (APTGD), $\xi > 0$ is alpha power transformation fréchet distribution (APTFD) and $\xi < 0$ is alpha power transformation Weibull distribution (APTWD).

Plots of the APTGEV probability density function for specific values of parameters are displayed in Figure 1.

The quantile function of the APTGEV distribution for $\alpha > 0, \alpha \neq 1$ is

$$x_p = \begin{cases} \mu + \dfrac{\sigma}{\xi}((\log(\log(\alpha)) - \log(\log(p(\alpha - 1) + 1)))^{-\xi} - 1), & \xi \neq 0 \\ \mu - \sigma \log(\log(\log(\alpha)) - \log(\log(p(\alpha - 1) + 1))), & \xi \to 0. \end{cases} \tag{10}$$

The estimates of extreme quantiles of the maximum distribution, known as return level $R_T(u)$, are then obtained by inverting of $G(x; \mu, \sigma, \alpha)$ for $\alpha > 0, \alpha \neq 1$ is then obtained as

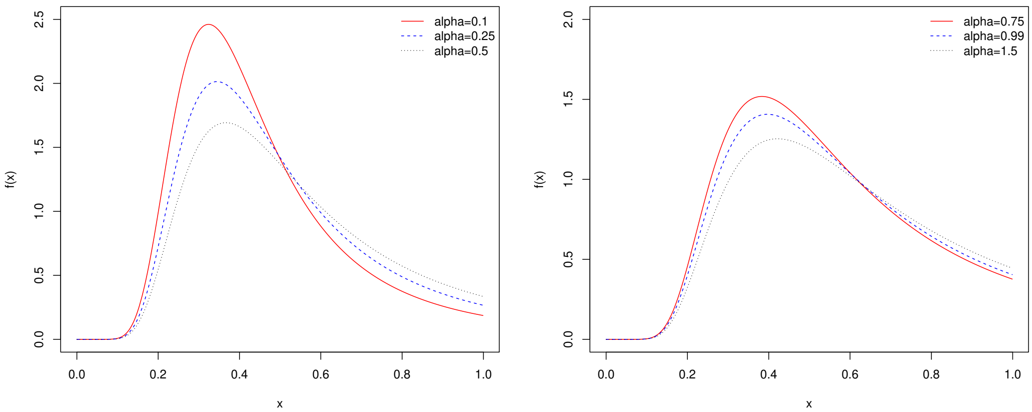

Figure 1: Plots of the APTGEV probability density function as $\mu = 0.5$, $\sigma = 0.3$, and $\xi = 0.5$

$$R_T(u) = \begin{cases} \mu + \dfrac{\sigma}{\xi}((\log(\log(\alpha)) - \log(\log(u(\alpha-1)+1)))^{-\xi} - 1), & \xi \neq 0 \\ \mu - \sigma \log(\log(\log(\alpha)) - \log(\log(u(\alpha-1)+1))), & \xi \to 0. \end{cases} \tag{11}$$

where $u \in [0,1]$, $u = 1 - T^{-1}$, and T is the number of recurrent times.

2.2. Distribution properties

In this section, we derive the expressions for some essential properties of the APTGEV distribution. The power series can be represented by

$$\alpha^w = \sum_{i=0}^{\infty} \frac{(\log \alpha)^i}{i!} w^i. \tag{12}$$

Hence, inserting (12) in PDF (7), then

$$\begin{aligned} g(x;\mu,\sigma,\xi,\alpha) =& \frac{\log\alpha}{\alpha-1}\sigma^{-1}[1+\xi(x-\mu)/\sigma]^{-(1/\xi)-1}\exp\{-[1+\xi(x-\mu)/\sigma]^{-1/\xi}\} \\ & \sum_{i=0}^{\infty}\frac{(\log\alpha)^i}{i!}\exp\{-i[1+\xi(x-\mu)/\sigma]^{-1/\xi}\} \\ =& \sum_{i=0}^{\infty}\frac{(\log\alpha)^{i+1}}{(\alpha-1)i!}\sigma^{-1}[1+\xi(x-\mu)/\sigma]^{-(1/\xi)-1}\exp\{-(i+1)[1+\xi(x-\mu)/\sigma]^{-1/\xi}\} \\ =& \sum_{i=0}^{\infty} W_i\sigma^{-1}u^{\xi+1}e^{-(i+1)u} \end{aligned}$$

where $W_i = \dfrac{(\log\alpha)^{i+1}}{(\alpha-1)i!}$ and $u = [1+\xi(x-\mu)/\sigma]^{-(1/\xi)}$.

If X has the PDF, then its k^{th} moment, we used binomial expression can be obtained as follows.

$$\begin{aligned}\mu'_k &= \sum_{i=0}^{\infty} W_i \sigma^{-1} \int_0^{\infty} x^k u^{\xi+1} e^{-(i+1)u} dx \\ &= \sum_{i=0}^{\infty} \sum_{j,m=0}^{k} \frac{\mu^{k-j}(-1)^{k-m-1}\sigma^k k!k!}{(k-j)!(k-m)!j!m!\xi^k} W_i \int_0^{\infty} u^{-\xi m} e^{-(i+1)u} du \\ &= \sum_{i=0}^{\infty} \sum_{j,m=0}^{k} \tau_{i,j,m}(i+1)^{(\xi m-1)}\Gamma(1-\xi m),\end{aligned}$$

where $\tau_{i,j,m} = \dfrac{\mu^{k-j}(-1)^{k-m-1}\sigma^k k!k!}{(k-j)!(k-m)!j!m!\xi^k} W_i$ and $\Gamma(.)$ stands for the gamma function. The first four moments about zero can be found by setting $k = 1, 2, 3$ and 4 in μ'_k. Also, the APTGEV distribution's moment generating function can be found as follows:

$$M_x(t) = E(e^{tx}) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{i=0}^{\infty} \sum_{j,m=0}^{k} \tau_{i,j,m}(i+1)^{(\xi m-1)}\Gamma(1-\xi m).$$

The central moments (μ_k) of APTGEV distribution can be obtained from

$$\mu_k = E(X-\mu'_1)^k = \sum_{n=0}^{k} (-1)^n \binom{k}{n} (\mu'_1)^n \mu'_{k-n}.$$

Some simulations presented how the simple mean ($\bar{x}$), standard deviation (SD), skewness values (S.V.), and kurtosis values (K.S.) of the APTGEV distribution. The result is changed for different α values, and random samples with 1,000 are generated 1,000 times by the quantile function of the APTGEV distribution, shown in Table 1

2.3. The maximum likelihood estimates (MLEs)

All parameters of the APTGEV distributions are estimated from a complete set of samples using the maximum likelihood method, accompanied by the Newton-Raphson procedure. When the random sample X's follows APTGEV distribution, the likelihood and log-likelihood functions for $\xi \neq 0$ can be respectively written as

$$\begin{aligned}L(\mu,\sigma,\xi,\alpha;x) = \prod_{i=1}^{n} &\frac{\log\alpha}{\alpha-1}\sigma^{-1}[1+\xi(x_i-\mu)/\sigma]^{-(1/\xi)-1} \\ &\exp\{-[1+\xi(x_i-\mu)/\sigma]^{-1/\xi}\}\alpha^{\exp\{-[1+\xi(x_i-\mu)/\sigma]^{-1/\xi}\}}\end{aligned}$$

Table 1: The simulation results for the APTGEV distribution with $\mu = 0.5$, $\sigma = 0.3$ and $\xi = 0.5$

α	Mean	SD	S.V.	K.V.
0.01	0.5594	0.1273	-2.2132	4.6335
0.05	0.4441	0.1116	-1.9081	3.3215
0.10	0.3678	0.1020	-1.6687	2.3618
0.30	0.1556	0.0685	-1.0045	0.2559
1.50	0.5030	0.2951	-0.0348	-1.1410
3.00	0.5031	0.2949	-0.0336	-1.1448
5.00	0.5031	0.2949	-0.0364	-1.1432
10.00	0.5032	0.2947	-0.0299	-1.1497
20.00	0.5032	0.2947	-0.0321	-1.1472
40.00	0.5032	0.2946	-0.0305	-1.1506

and

$$l(\mu,\sigma,\xi,\alpha;x) = n\log\left(\frac{\log\alpha}{\sigma(\alpha-1)}\right) - (1/\xi+1)\sum_{i=1}^{n}[1+\xi(x-\mu)/\sigma] - \sum_{i=1}^{n}([1+\xi(x-\mu)/\sigma]^{-1/\xi}) + \log\alpha\sum_{i=1}^{n}\exp\left[-1+\xi(x-\mu)/\sigma\right]^{-1/\xi}.$$

When the random sample X's are distributed as APTGEV, the respective likelihood and log-likelihood functions for $\xi \to 0$ can be described as

$$L(\mu,\sigma,\alpha;x) = \prod_{i=1}^{n}\frac{\log\alpha}{\alpha-1}\sigma^{-1}\exp[-(x_i-\mu)/\sigma]\exp\{-\exp[-(x_i-\mu)/\sigma]\} \alpha^{\exp\{-\exp[-(x_i-\mu)/\sigma]\}},$$

and

$$l(\mu,\sigma,\alpha;x) = n\log\left(\frac{\log\alpha}{\sigma(\alpha-1)}\right) + \sum_{i=1}^{n}(-(x_i-\mu)/\sigma) - \sum_{i=1}^{n}\exp(-(x_i-\mu)/\sigma) + \sum_{i=1}^{n}\exp(-\exp(-(x_i-\mu)/\sigma))\log(\alpha).$$

2.4. Confidence interval of return level

The confidence interval of the return level for APTGEV distribution is performed using the Delta method as $Var(R_T) \approx \bigtriangledown R_T^t V \bigtriangledown R_T$, where V is a covariance matrix of $(\mu,\sigma,\xi,\alpha)^t$ and

$$\bigtriangledown R_T^t = [\frac{\partial R_T}{\partial\mu}, \frac{\partial R_T}{\partial\sigma}, \frac{\partial R_T}{\partial\xi}, \frac{\partial R_T}{\partial\alpha}].$$

Where

$$
\begin{aligned}
\frac{\partial R_T}{\partial \mu} =&1,\\
\frac{\partial R_T}{\partial \sigma} =&\xi^{-1}((\log(\log\alpha)-\log(\log(u(\alpha-1))+1))^{-\xi}-1),\\
\frac{\partial R_T}{\partial \xi} =&-\sigma\xi^{-1}(\log(\log\alpha)-\log(\log(u(\alpha-1))+1))^{-\xi}\log(\log(\log\alpha)\\
&-\log(\log(u(\alpha-1))+1))-\sigma\xi^{-2}(\log(\log\alpha)-\log(\log(u(\alpha-1))+1))^{-\xi}+\sigma\xi^{-2},\\
\frac{\partial R_T}{\partial \alpha} =&-\sigma(\log(\log(\alpha))-\log(\log(u(\alpha-1)+1)))^{-\xi-1}\\
&\left(\frac{1}{\alpha\log(\alpha)}-\frac{u}{(u(\alpha-1)+1)\log(u(\alpha-1)+1)}\right).
\end{aligned}
$$

The Wald approach is a well-known technique regularly employed in constructing confidence intervals for a particular parameter of interest. When performing extreme value analysis, especially when dealing with APTGEV distributions, the Wald method remains relevant. It facilitates the computation of the confidence interval for a desired parameter, denoted by θ. The confidence interval is determined using the equation $\theta \pm Z_{\alpha/2} \times \sqrt{Var(R_T)}$. Here, $Z_{\alpha/2}$ represents the critical value from the standard normal distribution that cuts off the upper $\alpha/2$ area in the tails, and $\sqrt{Var(R_T)}$ is the square root of the estimated variance. It's worth noting that this estimated variance, denoted by $Var(R_T)$, is derived using the Delta method, a common statistical technique used for approximating the variance of a function of a random variable.

3. Simulation Study

This section discusses implementing a Monte Carlo simulation study to ascertain the efficacy of Maximum Likelihood Estimates (MLEs) applied to the Asymmetric APTGEV distribution. The Inverse Cumulative Distribution Function technique produces random samples corresponding to the APTGEV distribution. This study uses the R programming language to conduct simulations and generate the resulting data.

The simulation is performed $N = 1{,}000$ times on the artificial data. Multiple sample sizes $n = 25$, 50, 100, 500, and 1,000 are derived from the APTGEV distribution via the inverse CDF technique. Subsequently, the Newton-Raphson method is employed with the maximum likelihood estimation technique to compute the estimated parameter values. The entire process is repeated for all parameters, facilitating the calculation of the biases and Mean Squared Errors (MSEs) for the APTGEV distribution in Table 2.

Upon scrutinizing the simulation results, it becomes evident that the estimated values for all parameters closely match the pre-defined ones, signifying a remarkable level of accuracy. Moreover, an intriguing pattern is discernible when analyzing the mean squared errors (MSEs) for all parameters: they consistently converge toward zero. This consistent tendency suggests an increasing precision in the estimates, a trend vividly illustrated in

Table 2. The biases of μ, σ, and ξ tend to decrease as the sample size increases while the bias of α increases. The estimated values of all four parameters in Table 2 are rather consistent across sample sizes, and Figure 2 shows a histogram of one simulated sample (for example, n=500) and its estimated density.

Table 2: Simulation for APTGEV distribution for some parameter values with $\mu = 0.3$, $\sigma = 0.1$, $\xi = 0.5$, $\alpha = 3$

		Parameter			
n		μ	σ	ξ	α
25	MLE	0.31030	0.13808	0.37270	2.56052
	MSEs	0.00092	0.00331	0.05837	0.88913
	Biase	0.01030	0.03808	-0.12730	-0.43948
50	MLE	0.30244	0.11523	0.45811	2.96645
	MSEs	0.00023	0.00107	0.02139	0.38388
	Biase	0.00244	0.01523	-0.04189	-0.03355
100	MLE	0.30231	0.10770	0.47998	3.10213
	MSEs	0.00011	0.00052	0.01131	0.23723
	Biase	0.00231	0.00770	-0.02002	0.10213
500	MLE	0.30183	0.10122	0.50090	3.21792
	MSEs	0.00003	0.00008	0.00196	0.08718
	Biase	0.00211	0.00144	-0.00082	0.20959
1000	MLE	0.30215	0.10047	0.50376	3.21847
	MSEs	0.00002	0.00004	0.00104	0.06933
	Biase	0.00204	0.00064	0.00323	0.21631

4. Application of rainfall

This study applies the proposed APTGEV distributions to the highest monthly rainfall data in Sukhothai province, Thailand, during 1987 – 2021. Weather station 373301-Si Samrong District Sukhothai Province collects data from the Thailand Meteorological Department, covering an area of 6,596 square kilometers (4,122,557.5 rai). The rainy season starts around May - October, which is the period when the southwest monsoon winds. Which is hot and humid, blowing from the Indian Ocean into Thailand. The rain began to fall a lot around the middle of May onwards. The rainiest month is September, with an average rainfall of 1,259.2 millimeters per year. The temperature varies in the range of 19°C to 37°C and very rarely below 16°C or above 39°C. With 176 observations, the maximum rainfall is 176.70 mm, with a means of 57.41 mm and a standard deviation of 25.42 mm. For parameter estimation in GEV and APTGEV distributions, we utilize the maximum likelihood method accompanied by Newton–Raphson procedure.

Next, we illustrate the results of parameter estimates in the GEV and APTGEV distributions in Table 3.

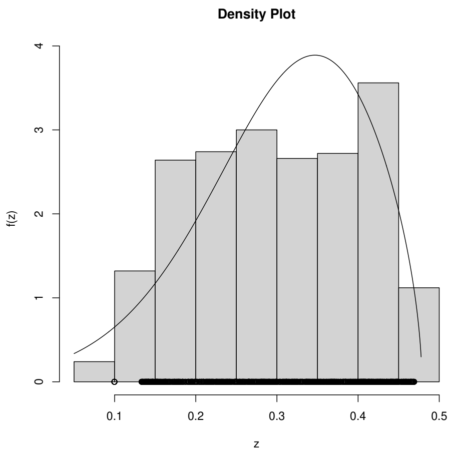

Figure 2: Plots of the APTGEV probability density function as simulated with sample size 500

Table 3: Maximum likelihood estimates, standard errors and 95% CI of all parameters in GEV and APTGEV distribution

Distribution	Parameter	Estimate	Standard error	95% CI
GEV	μ	44.34251	1.1199	(42.1476,46.5374)
	σ	12.5069	0.9905	(10.5654,14.4483)
	ξ	0.3744	0.0831	(0.2114,0.5373)
APTGEV	μ	40.5331	4.2223	(32.2574,48.8089)
	σ	9.79247	2.9535	(4.0036,15.5813)
	ξ	0.3992	0.0863	(0.2301,0.5683)
	α	2.9367	3.8754	(-4.6590,10.5324)

In Table 3, the parameters in GEV distribution are estimated using the maximum likelihood method. The estimated values are $(\mu, \sigma, \xi) = (44.34251,\ 12.5069,\ 0.3744)$, with standard errors (1.1199, 0.9905, 0.0831). The approximate 95% CI for the parameters are (42.1476,46.5374) for μ, (10.5654, 14.4483) for σ, and (0.2114,0.5373) for ξ. Since the confidence interval of $\xi > 0$, fréchet distribution is the optimal distribution for the GEV class.

Similarly, the APTGEV parameters are estimated using the maximum likelihood method. The estimated results, as shown in Table 3, are $(\mu, \sigma, \xi, \alpha) = (40.5331,\ 9.79247,\ 0.3992,\ 2.9367)$, with standard errors (4.2223, 2.9535, 0.0863, 3.8754). The approximate 95% CI for the parameters are thus (32.2574,48.8089) for μ, (4.0036,15.5813) for σ, (0.2301,0.5683) for ξ, and (-4.6590,10.5324) for α. Since zero lies within the confidence interval of ξ is contains 0, APTGD is the optimal distribution for the APTGEV class.

Table 4 shows the return level estimates at different return periods (T) of 10, 20, 50 and 100 years, based on two different distributions: GEV and APTGEV.

For the GEV distribution, the estimated return levels are higher as the return period

Table 4: Return level estimates (mm) at selected return periods (T) for the GEV and APTGEV distributions

Distribution	T	Return Level	95% CI
GEV	10	88.4916	(77.0835, 99.8997)
	20	112.4703	(91.1340, 133.8066)
	50	154.8332	(109.9808, 199.6857)
	100	197.8030	(123.4002, 272.2057)
APTGEV	10	88.8589	(77.1743, 100.5435)
	20	113.3955	(91.4084, 135.3826)
	50	157.5150	(110.6834, 204.3466)
	100	203.1066	(124.4146, 281.7987)

increases. For example, the estimated return level for a return period of 10 years is 88.4916 mm, while it is 197.8030 mm for a return period of 100 years. As a result, the lower and upper limits for confidence intervals also increase as the return period rises.

Similarly, the APTGEV distribution shows an increasing trend in the estimated return levels with increasing return periods. For example, the estimated return levels for return period of 10 years and 100 years are 88.8589 mm, and 203.1066 mm, respectively. As a consequence, the lower and upper limits for CIs also increase with the return period.

4.1. Goodness-of-fit tests

We will apply structured goodness-of-fit examinations to ascertain the distribution that best aligns with the data. The statistical tools we'll utilize include the Cramér-von Mises (W^*) and Anderson-Darling (A^*) methods delineated by Chen and Balakrishnan [3]. In general, smaller values of these statistics indicate a better fit of the distribution to the data. Let $F(x;\theta)$ be a CDF, where the form of F is known, but θ is unknown. To obtain the statistics W^* and A^* as follows:

$$W^2 = \sum_{i=1}^{n} \left\{ u_i - \frac{2i-1}{2n} \right\}^2 + \frac{1}{12n}$$

and

$$A^2 = -\,n - \frac{1}{n}\sum_{i=1}^{n}\{(2i-1)\log(u_i) + (2n+1-2i)\log(1-u_i)\},$$

modify W^2 into $W^* = W^2(1+0.5/n)$ and A^2 into $A^* = A^2(1+0.75/n+2.25/n^2)$, where $u_i = \Phi\{(y_i-\bar{y})/s_y\}$, $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$, $s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i-\bar{y})^2$, $y_i = \Phi^{-1}(v_i)$, $v_i = F(x_i;\hat{\theta})$, the x_i's are in ascending order, $\Phi(\cdot)$ is the standard normal CDF and $\Phi^{-1}(\cdot)$ denotes its inverse.

The values of W^* and A^* for GEV and APTGEV distribution are given in Table 5 for goodness-of-fit tests on rainfall data. Based on the values of W^* and A^*, the proposed

Table 5: Goodness-of-fit tests for the GEV and APTGEV distributions

Distribution	W^*	p-value	A^*	p-value
GEV	0.0454	0.9040	0.3436	0.9020
APTGEV	0.0414	0.9262	0.3247	0.9184

APTGEV distribution fits the data better than GEV distribution. Therefore, it could be considered as an alternative to other distributions in the literature for positive real data.

Figure 3 presents the diagnostic plots for the APTGEV distributions. Most points in the probability and quantile plots closely align with the unit diagonal, indicating that the APTGEV distribution functions offer reliable fits. This aligns with the results from the Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistical tests. Furthermore, the density plot shows a strong concurrence between the established APTGEV distribution function and the empirical density.

5. Conclusions

This paper introduces a novel extreme value distribution, the Alpha Power Transformation generalized extreme value (APTGEV) distribution. Our research methodology elaborates on the maximum likelihood estimation for all parameters, quantile function, moment generating function, and returns levels of GEV and APTGVE distribution, including their confidence intervals. Owing to the complexity of likelihood functions, we employ the Newton–Raphson method to derive the estimated values for all parameters. Through Monte Carlo experiments in a simulation study, we examine the behavior of these maximum likelihood estimates, which closely align with their initial assigned values. All parameters' mean squared errors (MSEs) converge toward zero. Notably, the biases for parameters μ, σ, and ξ diminish with an increase in sample size, while the bias for α rises. Yet, across different sample sizes, the estimated values for all four parameters remain relatively consistent. Regarding practical applications, we apply the proposed distributions to rainfall data. The results reveal a commendable fit per the Cramér-von Mises and Anderson-Darling goodness-of-fit tests. Thus, these proposed distributions demonstrate their effectiveness in fitting extreme data sets. We anticipate this generalization will have extensive applicability across disciplines, including statistics, mathematics, biology, environmental science, engineering, and more. Moreover, our study provides valuable insights for meteorologists regarding the behavior of high rainfall events within the context of extreme value theory. Both governmental and non-governmental entities can leverage these insights to make informed decisions and devise contingency plans in preparation for the effects of extreme rainfall events. Such measures can aid in early warning systems, food security, poverty reduction, and disaster or risk management, among other areas. Beyond its practical applications, our study also enriches theoretical knowledge, expands the scope of extreme value theory, and makes other theories more efficient.

Acknowledgements

This work is a partially supported dataset from the Thailand Meteorological Department. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

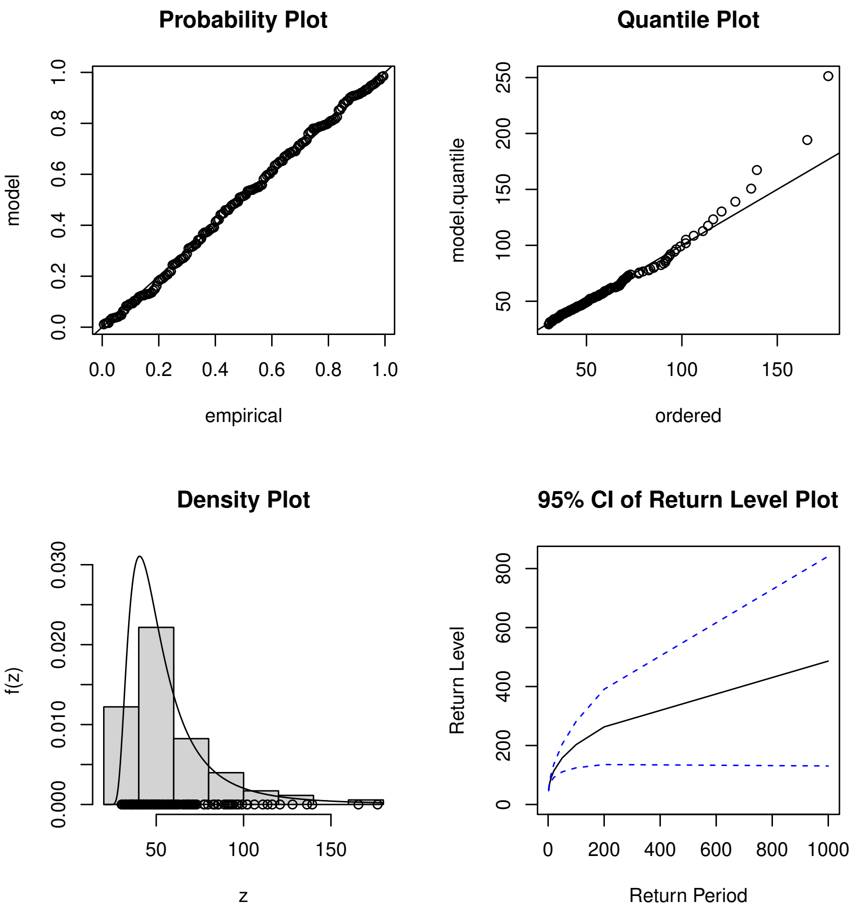

Figure 3: Diagnostic plots for APTGEV distribution, applying to rainfall data in Phitsanulok province, Thailand

References

[1] S. El Adlouni, T. B. M. J. Ouarda, X. Zhang, R. Roy, and B. Bobee. Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, 43:1–13, 2007.

[2] P. Busababodhin and A. Kaewmun. Extreme values statistics. *The Journal of King Mongkut's University of Technology North Bangkok*, 25(2):55–65, 2015.

[3] G. Chen and N. Balakrishnan. A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, 27(2):154–161, 1995.

[4] D. Chikobvu and R. Chifurira. Modelling of extreme minimum rainfall using generalised extreme value distribution for zimbabwe. *South African Journal of Science*, 111:1–8, 2015.

[5] N. Deetae. Analysis and mathematical modeling for flood surveillance from rainfall by extreme value theory for agriculture in phitsanulok province, thailand. *European Journal of Pure and Applied Mathematics*, 15(4):1797–1807, 2022.

[6] L. Gao, L. Zhang, and X. Xu. A global data set for economic losses of extreme hydrological events during 1960–2014. *Water Resources Research*, 5:5165–5175, 2019.

[7] E. Gilleland, M. Ribatet, and A. G. Stephenson. A software review for extreme value analysis. *Extremes*, 16(1):103–119, 2013.

[8] S. Ihtisham, A. Khalil, S. Manzoor, S. A. KhanID, and A. Ali. Alpha-power pareto distribution: Its properties and applications. *PloS one*, 14(6):e0218027, 2022.

[9] E. A. Ijigah and T. A. Akinyemi. Flood disaster: An empirical survey of causative factors and preventive measures in kaduna, nigeria. *International Journal of Environment and Pollution*, 3:53–56, 2015.

[10] C. Jones, D. E. Waliser, K. M. Lau, and W. Stern. Global occurrences of extreme precipitation and the madden–julian oscillation: Observations and predictability. *Journal of Climate*, 17:4575–4589, 2004.

[11] A. Mahdavi and D. Kundu. A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13):6543–6557, 2017.

[12] J.-L Martel, M. Alain, and B. Francois. Global and regional projected changes in 100-yr subdaily, daily, and multiday precipitation extremes estimated from three large ensembles of climate simulations. *Journal of Climate*, 33(3):1089–1103, 2020.

[13] M. M. Q. Mirza. Global warming and changes in the probability of occurrence of floods in bangladesh and implications. *Global Environmental Change*, 12:127–138, 2002.

[14] S. Nadarajah and I. E. Okorie. On the moments of the alpha power transformed generalized exponential distribution. *Ozone: Science and Engineering*, 46:273–285, 2017.

[15] M. Nassar, A. Alzaatreh, M. Mead, and O. Abo-Kasem. Alpha power weibull distribution: Properties and applications. *Communications in Statistics - Theory and Methods*, 46(20):10236–10252, 2017.

[16] F. C. Onwuegbuche, A. B. Kenyatta, S. B. Affognon, E. P. Enock, and M. O. Akinade. Application of extreme value theory in predicting climate change induced extreme rainfall in kenya. *International Journal of Probability and Statistics*, 8(4):85–94, 2019.

[17] M. Shrahili, A. S. Hassan, E. M. Almetwally, A. B. Ghorbal, and I. Elbatal. Alpha power moment exponential model with applications to biomedical science. *Scientific Programming*, 2022.

[18] S. B. Sunday, N. S. Agog, P. Magdalene, M. Abdullateef, and K. A. Gladys. Modeling extreme rainfall in kaduna using the generalised extreme value distribution. *Scientific World Journal*, 15(3):73–77, 2020.

[19] C. S. Withers and S. Nadarajah. Evidence of trend in return levels for daily rainfall in new zealand. *Journal of Hydrology (New Zealand)*, 39:155–166, 2000.