International Journal of Mathematics and
Computer Science, **17**(2022), no. 2, 589–599

Probability Analysis for Maximum Temperature in Northeast of Thailand

Monchaya Chiangpradit[1], **Piyapatr Busababodhin**[1], **Nipaporn Chutiman**[1], **Pannarat Guayjarernpanishk**[2], **Butsakorn Kong-ied**[1]

[1]Data Science and Sustainable Agriculture Research Unit
Department of Mathematics
Faculty of Science
Mahasarakham University
Maha Sarakham 41150, Thailand

[2]Faculty of Interdisciplinary Studies
Nong Khai Campus
Khon Kaen University
Nong Khai 43000, Thailand

email: butsakorn.k@msu.ac.th

(Received September 21, 2021, Accepted October 25, 2021)

Abstract

In this study, we investigate the climate change of northeastern region of Thailand. The annual extreme high temperature data was analyzed for figuring out the appropriate probability distribution. Three distributions including the Generalized Extreme Value Distribution (GEV), the Pearson Type III Distribution (P3), and the Log-Pearson Type III Distribution were conducted for estimating the return levels of return periods at 5 years, 10 years, 50 years, and 100 years. The analyzing results reveal that the Pearson Type III Distribution (P3) and the Log-Pearson Type III Distribution were appropriate according to the goodness of fit test with Kolmogorov-Smirnov Statistics (KS test). The return levels of the annual extreme high temperature

Key words and phrases: Temperature, Extreme Value, GEV Distribution, Pearson Type III Distribution, Log-Pearson Type III Distribution.
AMS (MOS) Subject Classifications: 60G70, 62P10, 62P25.
Butsakorn Kong-ied is the corresponding author.
ISSN 1814-0432, 2022, http://ijmcs.future-in-tech.net

in the return periods at 5 years, 10 years, 50 years, and 100 years are shown in the empirical results. The findings can help in making decisions and/or serve as guidelines for planning policy to cope with climate change.

1 Introduction

Climate change can have a serious effect on human beings and life in terms of food security, energy and water, as well as natural disaster [1]. According to the rapid increase of worlds population, the needs of consuming natural resources have also continuously increased. Nowadays, the operation of industrial sectors utilizing fossil fuel critically affects climate change. Daily life activities like transportation, household fuel and electric power utilization, as well as forest destruction, are also major causes of empowering greenhouse gas in the atmosphere leading to a greenhouse effect phenomenon or global warming condition. Greenhouse gas can absorb the heat wave or infrared ray initiated by the sun in order to balance the appropriateness of layers of the earths atmosphere. If the quantity of gas is larger than the appropriate level, then the temperature of the earth will be abnormally higher which will affect all life in the planet.

Thailand's climate has continuously and severely changed. This critical change will be gradually increased. For example, higher temperature occurs 30 to 60 days a year. Sea level, particularly at Andaman area, is higher due to the change of southwest monsoon. The frequent rate of raining is at the coastal area. Global warming will affect Chao Praya River in which the amount of water during November and December show a 40% increase compared with previous months due to amount of rainfall, water from the northern part, and storm surge. Sea level is higher causing flood in Bangkok. As a result, epidemic diseases are increasingly widespread due to higher global temperature. The Meteorological Department report [2] summarizing 2019 weather revealed that the beginning of the year (January and February) was cool especially in the northern and northeastern regions with the lowest temperature of 2.3 degree Celsius at Umpang District, Tak Province breaking the record. Mountain and hilltop weather very cold with the lowest temperature of -1.0 (minus) degree Celsius at the hilltop in Phuluang Wildlife Sanctuary and Phurue National Park.

Although the average temperature in February was lower than normal temperature, the temperature during daytime had increased to be hot in many areas. In the morning of early March, temperature increased and the weather was continuously hot throughout daytime. During March, April and May the weather in many areas of the upper part of Thailand was stuffy and very hot. The highest temperature broke the record at 44.6 degree Celsius on April 28 at Muang District, Maehongson Province. Average temperature in the upper part of Thailand during hot season was higher than other areas particularly during April and May which was higher than the normal temperature at 2-3 degree Celsius while 1-2 degree Celsius in the southern areas.

According to different phenomena affecting climate change in the northeastern region during the years 1951-2020 reported by Meteorological Department, it was found that the temperature showed an increase in terms of annual average temperature, annual highest temperature, and annual lowest temperature. The Meteorological Department studied the climate of northeastern region in 2012 and revealed that this region was affected by El Nino Phenomena. Some provinces had problems with temperature being higher than normal. The four provinces of highest increasing temperature were the provinces of Kalasin, Khon Kaen, Roi Et, and Maha Sarakham being not influenced by the southwest monsoon. If there were no tropical storm, then the region will basically face more serious drought.

The northeastern region of Thailand is the area facing higher temperature due to the landscape itself elevating the boundary with the central region. The boundary is clearly seen with different levels of high mountain which is the cause of drought in cold season. This region is situated between 120 and 400 meters above sea level and looks like a bowl lowering to the southeast. The edge of the mountain ridge is high. Phuphan Mountain Range lays across the western area while the south appears Sankamphaeng Mountain Range and Phanomdongrak Mountain Range sloping to Mekhong River towards the east. The average temperature of the region was 32.1 degree Celsius while the highest temperature of 43.8 degree Celsius was recorded in April in Udonthani.

Based on these problems, many researchers had tried to study, develop, and improve the methods of forecasting the weather. The accurate prediction of climate change is necessarily required for using as a tool and information for helping on explaining the climate change in the future. When analyzing data, the extreme high data is always found. This type of data is basically removed by analysts. It is not employed for creating the model due to its complexity and difficulty to analyze, e.g., daily highest and lowest rainfall, monthly highest wind speed, and daily highest and lowest temperature [3]. A good method of interpreting daily highest and lowest data is related to Extreme Value Theory. Many researchers had proposed this theory for applying to hydrology data. It is an effective method for explaining climate change.

Sabarish et al.[4] investigated the appropriate probability distribution of highest rainfall of 5-consecutive days at Tirushirappali City, Tamil Nadu State in India during a hundred year, it indicated that the Log-Pearson Type III (LP3) was the appropriate probability distribution for explaining the amount of consecutive rainfall. Husna et al.[5] studied the model of highest temperature with Generalized Extreme Value Distribution (GEV) and found that the highest temperature will rise to 35.6 degree Celsius at the next 10 years, as well as getting higher at the next 100 years. Senapeng and Bussababodhin [6] studied the highest temperature in northeastern region of Thailand using GEV Distribution and Generalized Pareto Distribution (GPD). They found that the highest temperature of all stations had increased when return period increased, particularly at Nongkhai Weather Station which showed the higher return level than other stations.

In this paper, we analyze the appropriate probability distribution of the annual highest temperature of the northeastern region. The Generalized Extreme Value Distribution (GEV), Pearson Type III Distribution (P3), and Log-Pearson Type III Distribution were conducted to predict the return level in the return period at 5 years, 10 years, 50 years, and 100 years. The estimation of the return level can inform us to determine the direction to follow to avoid or lessen the damage which may occur from climate change.

2 Data Preparation

This research used the annual highest temperature data during January 2013–December 2020 from the Meteorological Department, consisting of meteorological station and agrometeorological station in total 24 stations are shown in Table 1.

Table 1: The data from meteorological station and agrometeorological station

Province	Meteorological Station (M.S)	Agrometeorological Station (A.S)
Nong Khai	/	
Loei	/	/
Udon Thani	/	
Sakon Nakhon	/	/
Nakhon Panom	/	/
Khon Kaen	/	
Tha Phra		/
Mukdahan	/	
Kosum Phisai	/	
Chaiyaphum	/	
Roi Et	/	/
Ubon Ratchathani	/	/
Srisaket		/
Pak Chong	/	
Chokchai	/	
Surin	/	/
Tha Tum	/	
Nang Rong	/	

3 Methodology

3.1 GEV Distribution

Unlike the normal distribution that arises from the use of the central limit theorem on sample averages, the extreme value distribution arises from the limit theorem on extreme values or maxima in sample data.

Let x be a random variable represented by $x \sim \text{GEV}(\mu, \sigma, \xi)$ with the cumulative distribution as follows [7].

$$F(x) = exp\left\{-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\} \tag{3.1}$$

where $1+\xi\left(\frac{x-\mu}{\sigma}\right) > 0$, μ is a location parameter, σ is a scale parameter and ξ is a shape parameter respectively. The class of GEV distributions is very flexible with the tail shape parameter ξ controlling the shape and size of the tails of the three different families of distributions. The case of $\xi > 0$ is called the Frchet distribution, $\xi \longrightarrow 0$ is called the Gumbel distribution and $\xi < 0$ is called the Weibull distribution.

3.2 Pearson Type III Distribution

The probability density function (PDF) of a random variable X having the Pearson type III distribution (P3) may be written as [8].

$$f(x) = \frac{1}{a\Gamma(b)}\left(\frac{x-c}{a}\right)^{b-1} exp\left(-\frac{x-c}{a}\right) \tag{3.2}$$

where a , b and c are scale, shape and location parameters, respectively. If $a > 0$, then the distribution has a positive skewness and $x \geq c$. If $a < 0$, then the distribution has a negative skewness, $x \leq c$.

The cumulative distribution function (CDF) is defined as:

$$F(x) = \frac{1}{a\Gamma(b)}\int_0^{\infty}\left(\frac{x-c}{a}\right)^{b-1} exp\left(-\frac{x-c}{a}\right)dx \tag{3.3}$$

In addition, the gamma distribution is a special case of Pearson type III distribution when the location parameter equals to zero.

3.3 Log-Pearson Type III Distribution

The Log-Pearson type III distribution is a statistical technique for fitting distribution data. The random variable x has a Log-Pearson type III distribution (LP3) if the random variable y, which is $y = ln(x)$, has a Pearson type III distribution. The cumulative distribution function (CDF) of the Log-Pearson type III distribution is defined as [9].

$$F(x) = \frac{G\left(\alpha, \frac{x-\xi}{\beta}\right)}{\Gamma(\alpha)} \tag{3.4}$$

where $G(\alpha, x) = \int_0^x t^{\alpha-1}e^{-t}dx$.

The probability density function (PDF) of the Log-Pearson type III distribution is defined as:

$$f(x) = \frac{1}{x\beta\Gamma(\alpha)}\left[\frac{ln(x)-\xi}{\beta}\right] exp\left[-\frac{(ln(x)-\xi)}{\beta}\right] \tag{3.5}$$

where α , β and ξ are shape parameter, scale parameter and location parameter, respectively.

3.4 Kolmogorov-Smirnov Statistics

Kolmogorov-Smirnov Statistics (KS Test) is the best-known statistics for goodness-of-fit-tests on distribution. Let X be a continuous random variable with distribution function $F(x)$, and let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from X with order statistics $X_{(1)}, X_{(2)}, X_{(3)}, \ldots, X_{(n)}$. The null hypothesis $H_0 : F(x) = F_0(x)$, for all $x \in (-\infty, \infty)$ against the general alternative $H_1 : F(x) \neq F_0(x)$, for some $x \in (-\infty, \infty)$, where $F_0(x)$ is a hypothesized distribution function to be tested.

$$
\begin{aligned}
KS^2 &= \left\{ \sup_{t \in (-\infty,\infty)} \left| F_n(t) - F_0(t) \right| \right\}^2 \\
&= \left(\max_{1 \le i \le n} \left[\max \left\{ \frac{i}{n} - F_0(X_{(i)}), F_0(X_{(i)}) - \frac{i-1}{n} \right\} \right] \right)^2,
\end{aligned}
$$

where KS is the Kolmogorov-Smirnov Statistics [8].

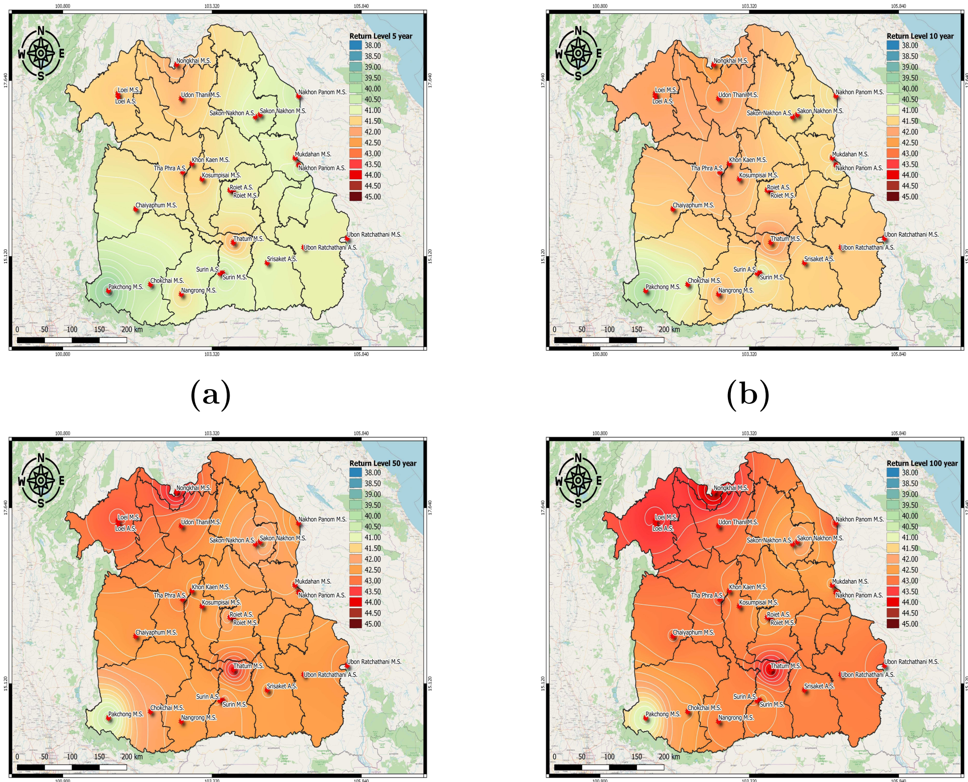
4 Empirical Results

The design and development of appropriate model for the annual highest temperature in the Northeastern region at 24 meteorological stations through GEV Distribution, Pearson type III distribution (P3), and Log-Pearson type III distribution (LP3), were analyzed by the KS test as a testing criterion for the appropriateness of distribution. the Maximum Likelihood Estimation (MLE) is the estimation parameter method. The result is presented in Table 2 which shows that 12 stations were fitted with Pearson Type III Distribution (P3) and the remaining 12 stations were fitted with Log-Pearson Type III Distribution. The extreme values of the temperature in different return levels were transformed into contour graphs by using GIS Kreiging interpolation in Geographic Information System (GIS). The results are illustrated in Figure 1.

Table 2: Probability analysis for annual highest temperature of meteorological stations

Stations	Fitted Distribution	Return levels (years)			
		5	10	50	100
Nong Khai M.S.	P3	41.94	42.58	43.80	44.27
Loei M.S.	LP3	41.63	42.23	43.33	43.73
Loei A.S.	LP3	41.54	42.10	43.13	43.51
Udon Thani M.S.	P3	41.79	42.22	42.91	43.14
Sakon Nakhon M.S.	LP3	40.89	41.20	41.60	41.70
Sakon Nakhon A.S.	LP3	40.93	41.38	42.17	42.44
Nakhon Panom M.S.	LP3	40.58	41.27	42.62	43.15
Nakhon Panom A.S.	LP3	40.58	41.21	42.40	42.86
Khon Kaen M.S.	P3	41.48	41.84	42.46	42.68
Tha Phra A.S.	P3	41.61	42.06	42.88	43.19
Mukdahan M.S.	LP3	41.25	41.59	42.14	42.32
Kosum Phisai M.S.	P3	41.53	41.91	42.57	42.80
Chaiyaphum M.S.	P3	41.08	41.60	42.61	42.99
Roi Et M.S.	P3	41.12	41.63	42.55	42.88
Roi Et A.S.	LP3	41.02	41.36	41.88	42.04
Ubon Ratchathani M.S.	LP3	41.12	41.66	42.71	43.11
Ubon Ratchathani A.S.	P3	41.08	41.60	42.58	42.94
Srisaket A.S.	LP3	41.01	41.51	42.44	42.78
Pak Chong M.S.	P3	39.54	39.91	40.51	40.71
Chok Chai M.S.	P3	40.35	40.89	41.97	42.40
Surin M.S.	LP3	40.19	40.74	41.87	42.33
Surin A.S.	LP3	41.14	41.66	42.64	43.01
Tha Tum M.S.	P3	41.88	42.47	43.59	44.01
Nang Rong M.S.	P3	41.35	41.78	42.54	42.80

M. S. stands for Meteorological Station
A. S. stands for Agrometeorological Station



(a) **(b)**

(c) **(d)**

Fig. 1: The estimated return levels in (a) 5 years, (b) 10 years, (c) 50 years and (d) 100 years with the fitted distribution of annual highest temperature in 24 stations

5 Conclusion and Discussion

The Result of analyzing the probability for annual highest temperature in northeastern region of Thailand when using MLE method indicated that the Pearson Type III Distribution (P3) and Log-Pearson type III distribution (LP3) were the fitted distribution. The 12 stations: Nongkhai Meteorological Station, Udon Thani Meteorological Station, Khon Kaen Meteorological Station, Tha Phra Agrometeorological Station, Kosum Phisai Meteorological Station, Chaiyaphum Meteorological Station, Roi Et Meteorological Station, Ubon Ratchathani Agrometeorological Station, Pak Chong Meteorological Station, Chokchai Meteorological Stationn, Tha Tum Meteorological Station, and Nang Rong Meteorological Station were fitted with Pearson type III distribution (P3). The remaining 12 stations: Loei Meteorological Station, Loei Agrometeorological Station, Sakon Nakhon Meteorological Station, Sakon Nakhon Agrometeorological Station, Nakhon Panom Meteorological Station, Nakhon Panom Agrometeorological Station, Mukdahan Meteorological Station, Roi Et Agrometeorological Station, Ubon Ratchathani Meteorological Station, Srisaket Agrometeorological Station, Surin Meteorological Station, Surin Agrometeorological Station were fitted with Log-Pearson type III distribution (LP3).

Regarding the return levels, the first rank was Nongkhai Meteorological Station, in Nongkhai Province and the second rank was Thatum Meteorological Station, in Surin, respectively. For Nongkhai province, a return period of 5 years, it was 0.2 of probability that annual highest temperature being 41.95 degree Celsius was once. For a return period of 10 years, it was 0.1 of probability that annual highest temperature being 42.58 degree Celsius was once. For a return period of 50 years, it was 0.02 of probability that annual highest temperature being 43.80 degree Celsius was once. For a return period of 100 years, it was 0.01 of probability that annual highest temperature being 44.27 degree Celsius was once. The result corresponding to [2] which reported that the change of highest temperature in Thailand under the change of greenhouse gas situation tended to increase in almost all regions. The temperature may be higher than 35 degrees Celsius during the hot season of the year. The prediction also indicated that the period of hot season will be 2-3 month longer in almost all regions of Thailand, particularly in northern Thailand. In addition, the results can inform us to determine the direction to follow to avoid or lessen the damage which may occur from climate change.

Acknowledgment. This work was financially supported by the National Research Council of Thailand and Mahasarakham University. The authors would like to thanks constructive comments and suggestions from reviewers.

References

[1] K. L. Ebi, D. M. Mills, J. B. Smith, A. Grambsch, Climate Change and Human Health Impacts in the United States: An Update on the Results of the U.S. National Assessment, Environ. Health Perspect, **114,** no. 9, (2006), 1318–1324.

[2] Thai Meteorological Department, Climate of Thailand, Retrieved July 12, 2021 from https://www.tmd.go.th/

[3] F. Ashoori, M. Ebrahimpour, A. Gharib, A. Bozorgnia, Minimal Generalized Extreme Value Distribution and its Application in Modeling of Minimum Temperature, International Journal of Statistics and Probability, **4,** no. 2, (2015), 87–94.

[4] R. Sabarish, N. R. Mani, A. R. Chandhru, C. R. Suribabu, J. Sudharsan, S. Nithiyanantham, Probability Analysis for Consecutive-day Maximum Rainfall for Tiruchirapalli City (south India, Asia), Applied Water Science, **2,** no. 7, (2014), 1033–1042.

[5] H. B. Hasan, F. B. Noor, A. Radi, S. B. Kassim, Modeling of Extreme Temperature using Generalized Extreme Value Distribution: A Case Study of Penang. In Proceeding of the World Congress on Engineering, London, **I,** (2012).

[6] P. Senapeng, P. Bussababodhin, Modeling of Maximum Temperature in Northeast Thailand, Burapha Science Journal, **22,** no. 1, (2017), 92–107.

[7] S. Coles, S. Nadaraja, An Introduction to Statistical Modeling of Extreme Values, Springer-Verlag, London, 2001.

[8] R. W. Vogel, D. E. mcMartin, Probability Plot Goodness-of-Fit and Skewness Estimation Procedures for the Pearson Type 3 Distribution, Water Resources Research, **27,** no. 12, (1991), 3149–3158.

[9] N. Millington, S. Das, S. P. Simonovic, The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models, Water Resources Research Report, Department of Civil and Environmental Engineering, The University of Western Ontario, (2011).