Available online at www.CivileJournal.org

Civil Engineering Journal

(E-ISSN: 2476-3055; ISSN: 2676-6957)

Vol. 9, No. 07, July, 2023

Climate Forecasting Models for Precise Management Using Extreme Value Theory

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

[1] *Faculty of Interdisciplinary Studies, Nong Khai Campus, Khon Kaen University, Nong Khai 43000, Thailand.*

[2] *Department of Mathematics, Faculty of Science, Mahasarakham University, Maha Sarakham 41150, Thailand.*

Received 01 March 2023; Revised 17 June 2023; Accepted 25 June 2023; Published 01 July 2023

Abstract

The objective of this research was to develop a mathematical and statistical model for long-term prediction. The Extreme Value Theory (EVT) was applied to analyze the appropriate distribution model by using the peak-over-threshold approach with Generalized Pareto Distribution (GPD) to predict daily extreme precipitation and extreme temperatures in eight provinces located in the upper northeastern region of Thailand. Generally, each province has only 1–2 meteorological stations, so spatial analysis cannot be performed comprehensively. Therefore, the reanalysis data were obtained from the NOAA Physical Sciences Laboratory. The precipitation data were used for spatial analysis at the level of 25 square kilometers, which comprises 71 grid points, whereas the temperature data were used for spatial analysis at the level of 50 square kilometers, which includes 19 grid points. According to the analysis results, GPD was appropriate for the goodness of fit test with Kolmogorov-Smirnov Statistics (KS Test) according to the estimation for the return level in the annual return periods of 2 years, 5 years, 10 years, 25 years, 50 years, and 100 years, indicating the areas with daily extreme precipitation and extreme temperatures. The analysis results would be useful for supplementing decision-making in planning to cope with risk areas as well as in effective planning for resources and prevention.

Keywords: Extreme Value Theory; Generalized Pareto Distribution; Precipitation; Temperature.

1. Introduction

Climate change causes variations in temperature and precipitation, which can be seen from a higher average temperature in the summer and a higher average precipitation in the rainy season. Therefore, the accuracy of spatial climate prediction is important. The development of climate models in numerous countries indicates that severe floods will occur by the year 2100 at many times over, with a chance to occur 3–6 times in a 100-year period. This would be unlike the past, when it occurred once per 100 years. In Southeast Asia, natural disasters tend to occur more frequently with profound effects on human life; they can cause damage to houses and agricultural products, including human food sources, and affect residences. In addition, the global warming phenomenon is occurring continuously because the earth cannot effectively dissipate heat from solar radiation. As a result, the climate has changed globally, causing the average global temperature to rise partly as a consequence of more melting of polar icebergs and more water in the rivers and seas, which will impact the world's living creatures. These factors cause global climate change, as seen by higher average

* Corresponding author: nipaporn.c@msu.ac.th

http://dx.doi.org/10.28991/CEJ-2023-09-07-014

temperatures in the summer and higher average precipitation in the rainy season. Thailand has been affected continuously by climate change in the form of droughts, floods, and heat waves. It can be seen that the total amount of accumulated rainfall is not much different from year to year, but the rainfall pattern has changed from raining throughout the areas in the rainy season to raining heavily and intermittently in narrow areas. In the past, the northeastern region was affected by climate change, particularly in the upper northeastern region.

In the past 10 years, this region has been greatly affected by an increase in annual average temperature and accumulated precipitation during the monsoon period. Therefore, the long-term prediction of extreme values is important and necessary. Many researchers have applied the extreme value theory to develop models for long-term prediction, such as in insurance by Chavez-Demoulin et al. [1], who presented an extreme value approach for modeling operational risk losses depending on covariates. Chavez-Demoulin & Guillou [2] studied extreme quantile estimation for β-mixing time series and insurance applications. Einmahl et al. [3] studied limits to human life span through the EVT. Diawara et al. [4] applied the EVT to determine extreme claims in the automobile insurance sector. Researchers have applied the EVT in finance, such as de Haan et al. [5], who applied extreme value statistics to financial time series dealing with bias and serial dependence.

Gkillas & Katsiampa [6] studied the application of extreme value theory to cryptocurrency. Drees et al. [7] studied extreme value estimation for discretely sampled continuous processes. Researchers have applied the EVT to risk management. For example, Tamošaitienė et al. [8] studied Project portfolio construction using the EVT. Opala et al. [9] proposed modeling the tail-dependence of crypto assets with EVT from the perspective of risk management in banks. Embrechts et al. [10] propose a modern EVT at the interface of risk management. Külekci et al. [11] studied the assessment of dependent risk using EVT in a time-varying framework. Yousof et al. [12] proposed a new reciprocal Weibull extension for modeling extreme values with risk analysis. The researchers applied the extreme value theory in medicines, such as Canton Enriquez et al. [13], applying the probabilistic models for extreme values to the COVID-2019 epidemic daily dataset. Chutiman et al. [14] studied the risk area assessment of re-emerging diseases in elderly people by using the extreme value theory. Moreover, the EVT was applied to machine learning as proposed by Einmahl et al. [15] and Vignotto & Engelke [16].

Most hydrological research uses the data from meteorological stations as follows. Croitoru & Piticar [17] studied changes in daily extreme temperatures in the extra-Carpathian regions of Romania by using the data sets of daily minimum and maximum temperatures recorded in 14 weather stations. Olmo et al. [18] investigated the influence of atmospheric circulation on the temperature and precipitation of individual and compound daily extreme events by using meteorological station datasets. Tošić et al. [19] explored extreme temperature events in Serbia concerning atmospheric circulation by using the data for daily maximum and minimum temperatures observed at 11 stations in Serbia during the period between 1949–2018. Busababodhin et al. [20] conducted extreme value modeling of daily maximum temperature with the r-Largest Order Statistics. These data sets were collected between 1984 and 2018 by Thai Meteorological stations. Gong et al. [21] observed changes in extreme temperature and precipitation indices on the Qinghai-Tibet Plateau during 1960–2016 by basing their observations on daily observations from 1960–2016 at 94 meteorological stations. Chiangpradit et al. [22] performed a probability analysis for maximum temperature in Northeast Thailand by using the annual highest temperature data from January 2013 to December 2020 at 24 meteorological stations.

In Thailand, each province has only 1–2 meteorological stations, so comprehensive spatial analysis cannot be performed. Therefore, this research used satellite data as reanalysis data from the NOAA Physical Sciences Laboratory (*https://psl.noaa.gov/*). The precipitation data were used for spatial analysis at a level of 25 square kilometers, whereas the temperature data were used for spatial analysis at a level of 50 square kilometers. These data were used for developing a mathematical and statistical model for long-term prediction based on the EVT to predict precipitation and temperature in the upper northeastern provinces of Loei, Nong Khai, Bueng Kan, Nong Bua Lampu, Udon Thani, Nakhon Phanom, Mukdahan, and Sakon Nakhon. Moreover, the return level was estimated for extreme precipitation and extreme temperatures in the long term. The results would be useful for planning and preparing resources as well as effective prevention.

2. Domain of Experiment and Methodology

2.1. Domain of Experiment

The study area comprised provinces in the upper northeastern region of Thailand, including Loei, Nong Khai, Bueng Kan, Nong Bua Lampu, Udon Thani, Nakhon Phanom, Mukdahan, and Sakon Nakhon. These provinces are economic areas with international trade at the borders with neighboring countries. These areas were affected by climate change more than other areas in the northeastern region.

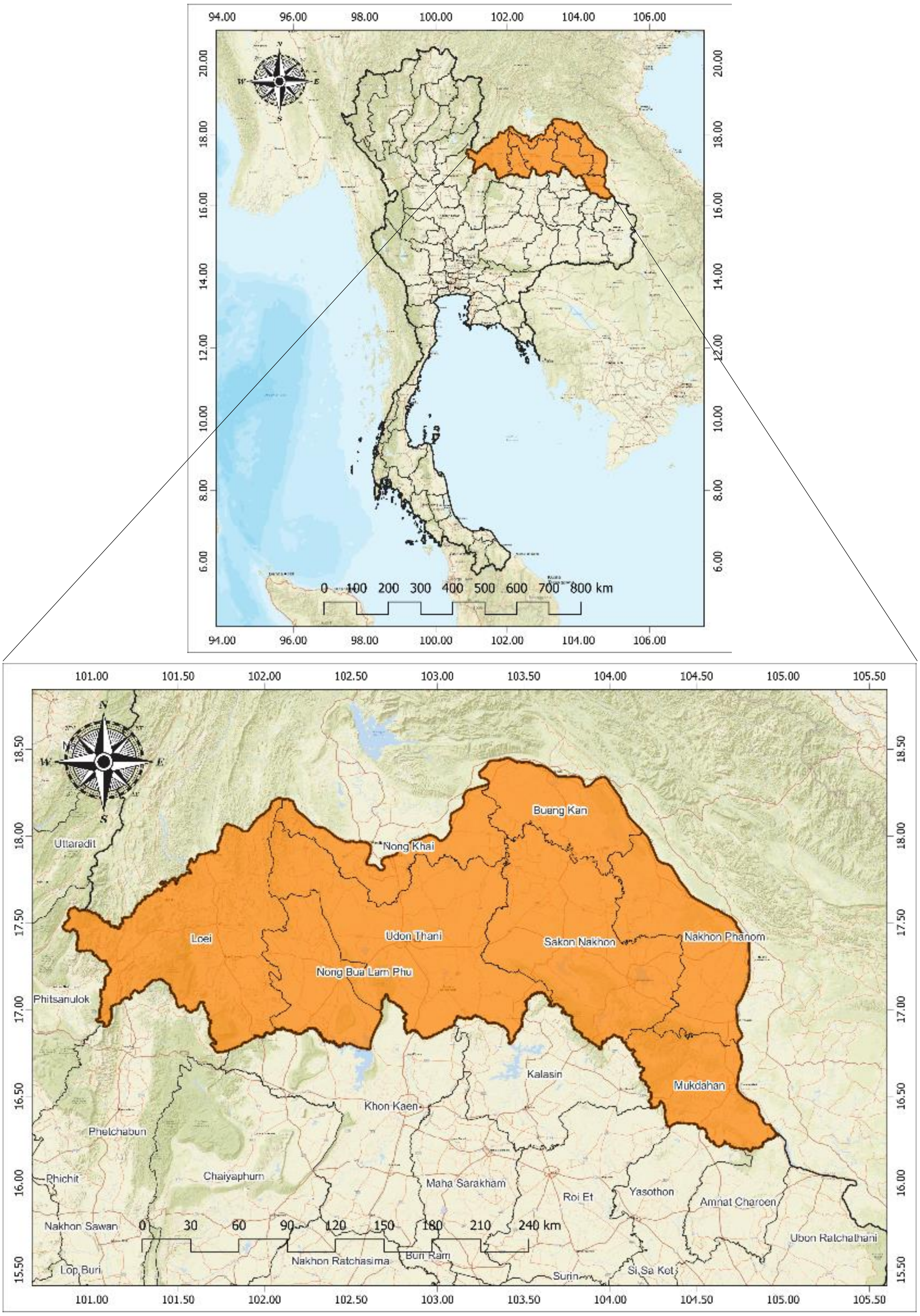

Figure 1. Study area covering provinces in the upper northeastern region of Thailand

Data Preparation

Data for precipitation and temperature in the grids were selected from the satellite data of NOAA Physical Sciences Laboratory (*https://psl.noaa.gov/*) as follows.

- Daily accumulated precipitation data with a spatial resolution of 0.25-degree latitude by starting from 16 - 21 degrees north longitude and 100–105 degrees east during 2012–2022.
- Daily maximum temperature data with a spatial resolution of 0.5-degree latitude starting from 16 - 21 degrees north longitude and 100–105 degrees east during 2012–2022.

2.2. Methodology

2.2.1. Extreme Value Theory

$X_1, X_2, \ldots, X_n$ were set as continuous random variables with an n size, which were independent and identically distributed (i.i.d.) with the same cumulative distribution function of $F(x) = Pr(X_i \leq x)$. The maximum values for the random variables were represented by M_n i.e., $M_n = \max(X_1, X_2, \ldots, X_n)$, where X_i represented the data selected from the blocks with the same period such as the high level of seawater every hour, average daily temperature, annual accumulated precipitation, etc. Therefore, M_n represented the maximum values for the studied variables in each time block of the process. If n represented the number of observation values in one year, M_n would be the annual maximum value.

According to the probability distribution function of M_n , the probability of the maximum value was calculated by considering the probability of all n values (Coles & Nadaraja [23]) as follows:

$$Pr\{M_n \leq z\} = Pr\{X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z\} = Pr(X_1 \leq z) \times Pr(X_2 \leq z) \times \ldots \times Pr(X_n \leq z) = [F(z)]^n \qquad (1)$$

Because the cumulative distribution function (F) of the random samples was unknown, the statistical method was used to estimate the random samples for the study. Although the error from the F estimation was minimal, it could lead to error for F^n with higher values. According to the study on F^n property, when $n \to \infty$ but some conditions did not cover the property, the random variable of M_n was converted to linear renormalization in the form of M_n^*, i.e. $M_n^* = \frac{M_n - b_n}{a_n}$. When the values were constant at $\{a_n > 0\}$ and $\{b_n\}$, appropriate choices for the $\{a_n\}$ and $\{b_n\}$ stabilized the location and scale of M_n^* as n increased, avoiding the difficulties that arise with the variable M_n. Therefore, the authors sought limited distribution for M_n^*, with appropriate choices for $\{a_n\}$ and $\{b_n\}$, rather than M_n.

Theorem 1: Extremal Types Theorem (Coles & Nadaraja [23]).

If the constant sequence values were $\{a_n > 0\}$ and $\{b_n\}$, and where $n \to \infty$, it would be found as follows:

$$Pr\left\{\frac{M_n - b_n}{a_n} \leq z\right\} \to G(z) \qquad (2)$$

In the case of non-degenerate distribution, $G(z)$ the property would be distributed in one of the following distributions.

$$\text{(I) } G(z) = exp\left\{-exp\left(-\left(\frac{z-b}{a}\right)\right)\right\}, \; -\infty < z < \infty;$$

$$\text{(II) } G(z) = \begin{cases} 0, & z \leq b, \\ exp\left\{-\left(-\left(\frac{z-b}{a}\right)^{-\alpha}\right)\right\}, & z > b; \end{cases} \qquad (3)$$

$$\text{(III) } G(z) = \begin{cases} exp\left\{-\left(-\left(\frac{z-b}{a}\right)^{-\alpha}\right)\right\}, & z < b, \\ 1, & z \geq b, \end{cases}$$

For parameter $a > 0, b \in \mathbb{R}$ and in the case of families II and III, $\alpha > 0$.

These three distribution types of the extreme value theory are known as Gumbel, Fréchet, and Weibull. For the distributions of Gumbel and Fréchet, $G(z)$ without a lower limit refers to a value higher than the final value, i.e. ∞ . On the other hand, for the Weibull distribution, $G(z)$ has an upper limit, and the limit of M_n cannot be identified by the extreme value theory, even though the distribution type is specified. The distribution type for the random samples is identified with the central limit theorem so the authors are not interested in what previous distribution type of F. The peak-over-threshold approach is another method for analyzing extreme values by basing on the following concepts. The random variables $X_1, X_2, \ldots, X_n$ were independently and identically distributed. The distribution function peaked over the threshold when X was greater than the threshold u i.e., $F_u(X) = Pr(X - u | X > u)$ and u was set at a fairly high value. The conditional distribution $F_u(X)$ was estimated by Generalized Pareto distribution (GPD). GPD was the peak-over-threshold approach suitable for the extreme value analysis with daily data. The data with a peak over threshold were selected for analysis according to the following concepts.

If X was set as a random variable of GPD, it was represented by $X \sim GPD\ (\sigma, \xi)$, where σ is the scale parameter and ξ is the shape parameter. Normally, when having a peak over threshold or u, the cumulative distribution function (CDF) of $x - u$ is conditional, i.e. $x > u$, as in the following Equation 4.

$$F(x-u) = H(x) = 1 - \left(1 + \frac{\xi x}{\tilde{\sigma}}\right)^{\frac{-1}{\xi}} \tag{4}$$

where $x > 0$, $\tilde{\sigma} = \sigma + \xi(u - \mu)$ and μ is the location parameter.

According to Equation 4, the distribution was in the same category as GPD, where $\tilde{\sigma}$ was a scale parameter for the specified value $u > u_0$. From equation 4, it was found that $\tilde{\sigma} = \sigma + \xi(u - u_0)$ so the scale parameter changed with the exception of where $\xi = 0$ the scale parameter was adjusted by $\sigma^* = \tilde{\sigma} + \xi u$. Moreover, u_0 was selected from the minimum value of u where the estimators of $\sigma^*$and ξ were constant, and the probability distribution function (PDF) was calculated as in the following Equation 5:

$$h(x) = 1 + \xi\left[\left(\frac{x-u}{\sigma}\right)\right]^{-\frac{1}{\xi}} \text{ where } \sigma > 0 \text{ and } -\infty < \xi < \infty \tag{5}$$

where σ represents a scale parameter and ξ represents a shape parameter. In GPD, the case of $\xi \to 0$ is called exponential distribution, while the case of $\xi > 0$ is called Pareto distribution, and the case of $\xi > 0$ is called gamma distribution.

The return level of GPD was calculated as in the following Equation 6:

$$\hat{z}_p = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}}\left[\frac{1}{T}\eta_y \lambda_\mu\right]^{\hat{\xi}} - 1 \text{ where } \hat{\sigma} > 0 \text{ and } -\infty < \hat{\xi} < \infty \tag{6}$$

where $\hat{z}_p$ refers to return level, $\hat{\mu}$ refers to location parameter, $\hat{\sigma}$ refers to scale parameter, $\hat{\xi}$ refers to shape parameter, T refers to the return period, η_y refers to average number of days per year, λ_μ refers to the estimated probability of the peak over the threshold.

2.2.2. Tests on Model Appropriateness with Kolmogorov-Smirnov Statistics (KS Test)

Kolmogorov-Smirnov Statistics (KS Test) is a statistical test for the appropriateness of distribution. Let X be a continuous random variable with distribution function $F(x)$, and $X_1, X_2, \ldots, X_n$ be a random sample from X with order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$. The authors wished to test 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 is a hypothesized distribution function to be tested.

$$KS = \left\{\sup_{t \in (-\infty,\infty)} |F_n(t) - F_0(t)|\right\}^2 = \left(\max_{1 \le i \le n}\left[max\left\{\frac{i}{n} - F_0(X_{(i)}), F_0(X_{(i)}) - \frac{i-1}{n}\right\}\right]\right)^2 \tag{7}$$

where KS is the Kolmogorov-Smirnov Statistics, the best-known statistics for goodness-of-fit-tests (Zhang [24]).

The procedures for processing this research are shown in Figure 2.

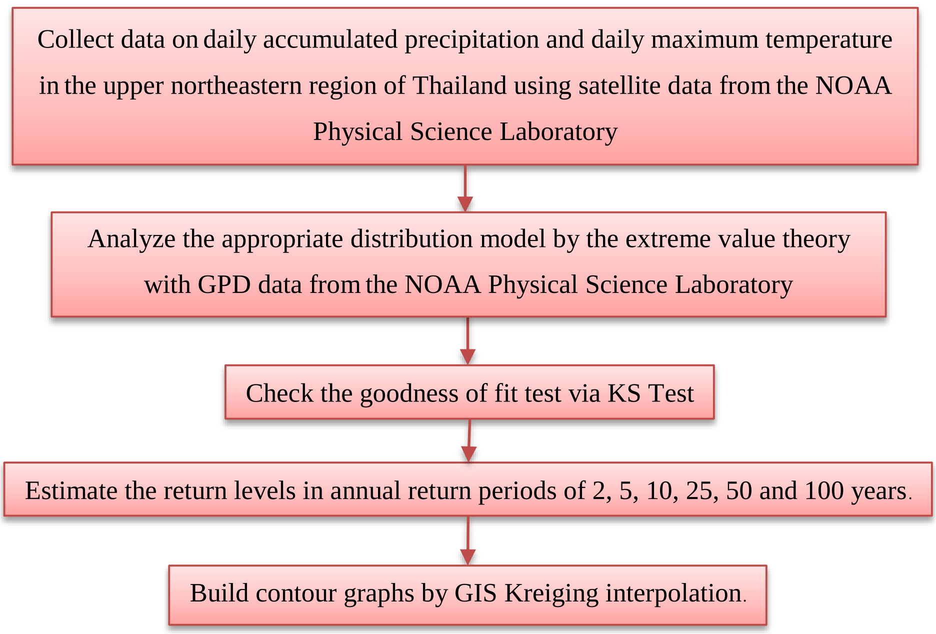

Figure 2. Procedures for the research methodology

3. Results and Discussion

For data analysis by using the extreme value theory, the appropriate model was calculated by Generalized Pareto Distribution (GPD) to analyze the daily extreme values in a daily period and the return level of the daily maximum precipitation, as shown in Table 1.

Table 1. Threshold values for precipitation and the appropriate distribution of precipitation in each grid

Province	Lat, long	u	Distribution	KS	Province	Lat, long	u	Distribution	KS
Loei	17.625,101.125	21.509	Exponential	0.9776	Nong Bua Lampu	17.375,102.125	24.770	Exponential	0.7373
	17.875,101.125	21.467	Exponential	0.9972		17.625,102.125	25.802	Exponential	0.8176
	17.375,101.375	21.719	Exponential	0.9931		17.875,102.125	25.933	Exponential	0.7623
	17.625,101.375	21.694	Exponential	0.6490		18.125,102.125	26.611	Exponential	0.9077
	17.125,101.625	21.873	Exponential	0.8294		17.625,102.375	26.494	Gamma	0.2983
	17.375,101.625	22.950	Exponential	0.6613		17.875,102.375	27.005	Exponential	0.7029
	17.625,101.625	23.473	Exponential	0.8978		18.125,102.375	27.317	Exponential	0.6688
	17.875,101.625	23.675	Exponential	0.2341		18.125,102.625	27.843	Exponential	0.7317
	17.125,101.875	22.524	Exponential	0.9893	Bueng Kan	16.625,103.375	23.556	Gamma	0.7430
	17.375,101.875	23.813	Gamma	0.9925		16.875,103.375	24.235	Exponential	0.8151
	17.625,101.875	24.955	Exponential	0.9440		16.625,103.625	23.331	Exponential	0.6900
	17.875,101.875	25.586	Exponential	0.9014		17.125,103.875	26.927	Exponential	0.8876
	18.125,101.875	25.307	Exponential	0.4139		16.875,104.125	27.341	Exponential	0.3868
Udon Thani	17.125,102.125	22.434	Exponential	0.4463	Sakon Nakhon	17.625,103.375	29.011	Exponential	0.7929
	17.125,102.375	23.189	Exponential	0.8609		17.125,103.625	26.736	Exponential	0.8163
	17.375,102.375	25.802	Exponential	0.9627		17.375,103.625	27.652	Exponential	0.9779
	17.625,102.625	27.418	Exponential	0.9866		17.625,103.625	28.668	Exponential	0.8966
	17.875,102.625	27.972	Exponential	0.9749		17.375,103.875	28.322	Exponential	0.7770
	17.375,102.875	26.525	Exponential	0.5291		17.625,103.875	29.543	Exponential	0.9946
	17.625,102.875	27.980	Exponential	0.8015		17.875,103.875	30.482	Exponential	0.7096
	17.875,102.875	28.615	Exponential	0.9615		18.125,103.875	31.365	Exponential	0.9331
	18.125,102.875	31.151	Exponential	0.1956		17.625,104.125	30.607	Exponential	0.7323
	17.375,103.125	26.998	Exponential	0.8521		17.875,104.125	30.438	Exponential	0.4818
	17.625,103.125	28.056	Exponential	0.6345		17.625,104.375	29.863	Exponential	0.9243
	17.875,103.125	29.301	Exponential	0.9365		17.875,104.375	30.354	Exponential	0.4998
	17.375,103.375	27.226	Exponential	0.9084	Nakhon Phanom	17.125,104.125	27.744	Exponential	0.6501
	17.875,103.375	29.470	Exponential	0.6974		17.375,104.125	29.914	Exponential	0.8450
	18.125,103.375	31.071	Exponential	0.9785		17.375,104.375	29.889	Exponential	0.9904
	17.875,103.625	30.309	Exponential	0.9122		18.125,104.375	28.924	Exponential	0.9667
Nong Khai	16.875,102.125	21.929	Exponential	0.6780		17.875,104.625	28.936	Exponential	0.5178
	17.125,102.625	24.423	Exponential	0.9987		18.125,104.625	29.640	Exponential	0.3652
	17.125,102.875	24.446	Exponential	0.9138	Mukdahan	18.375,104.375	28.634	Exponential	0.8131
	16.875,103.125	23.962	Exponential	0.9504		18.625,104.375	31.044	Exponential	0.9955
	17.125,103.125	25.423	Exponential	0.8948		18.375,104.625	30.414	Exponential	0.2963
	17.125,103.375	25.903	Exponential	0.9391		18.625,104.625	32.025	Exponential	0.7799
						18.625,104.875	35.573	Exponential	0.6295

Table 1 shows the provinces in the upper northeastern region of Thailand, including Loei, Nong Khai, Bueng Kan, Nong Bua Lampu, Udon Thani, Nakhon Phanom, Mukdahan, and Sakon Nakhon. In the analysis, there were a total of 71 grid points. There were only 3 grid points in 3 provinces, i.e. Loei, Nong Bua Lampu, and Bueng Kan that distributed precipitation in gamma form. For the other 68 grid points, precipitation was distributed in exponential form. The prediction values for the return levels in 2 years, 5 years, 10 years, 25 years, 50 years, and 100 years were used to build contour graphs by using GIS Kreiging interpolation. The results are shown in Figures 3 to 8.

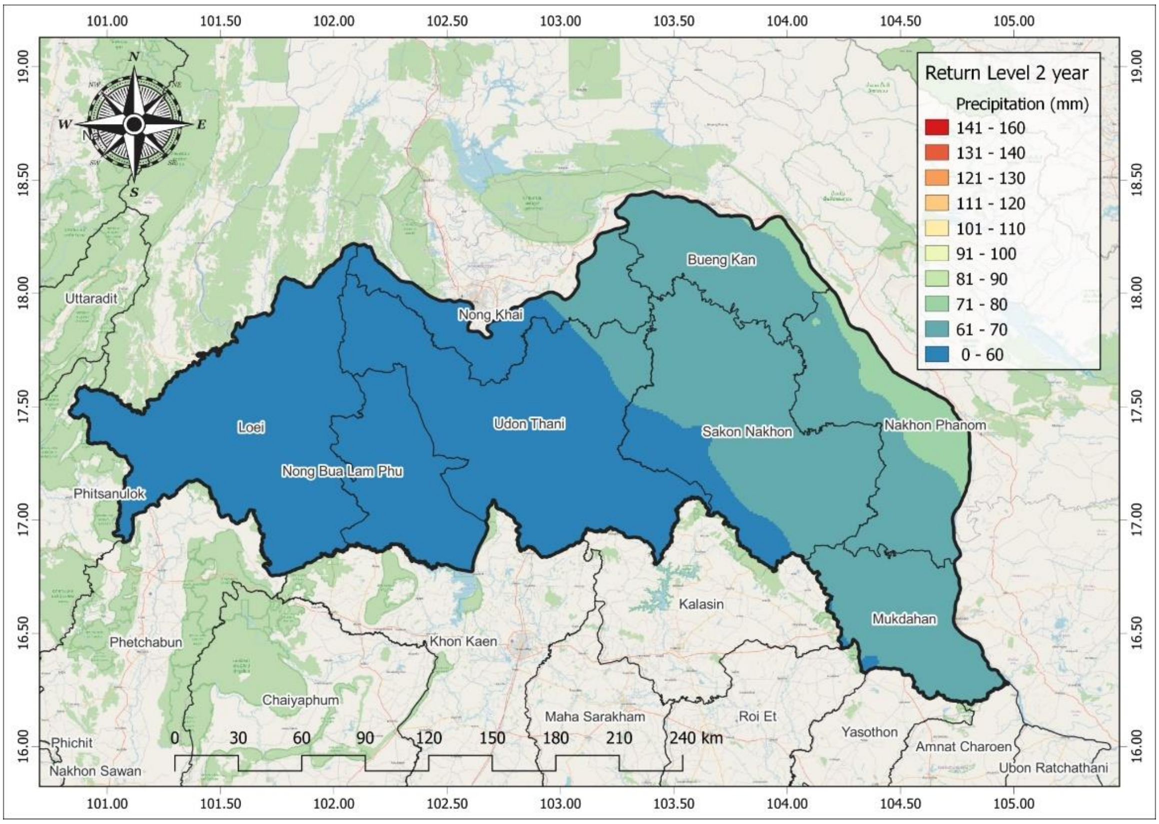

Figure 3. Estimated values for the return levels in 2 years with Generalized Pareto Distribution (GPD) of daily extreme accumulated precipitation

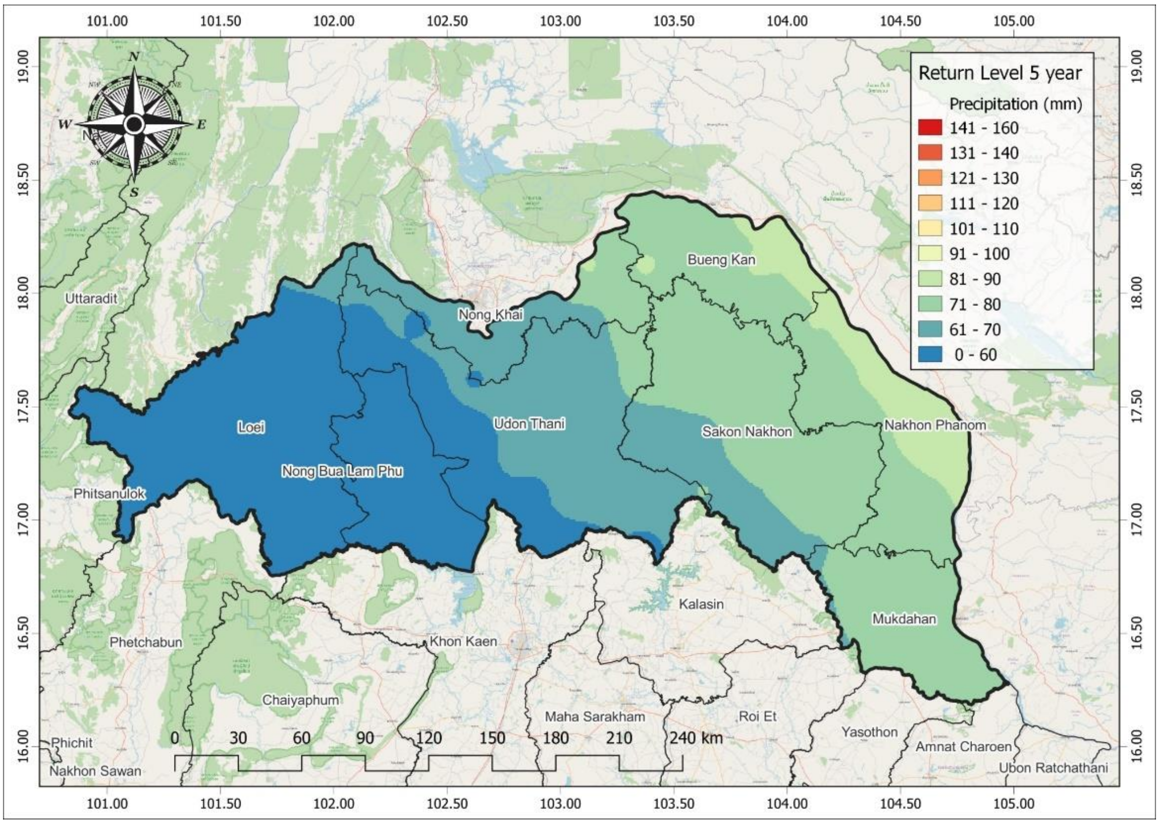

Figure 4. Estimated values for the return levels in 5 years with Generalized Pareto Distribution (GPD) of daily extreme accumulated precipitation

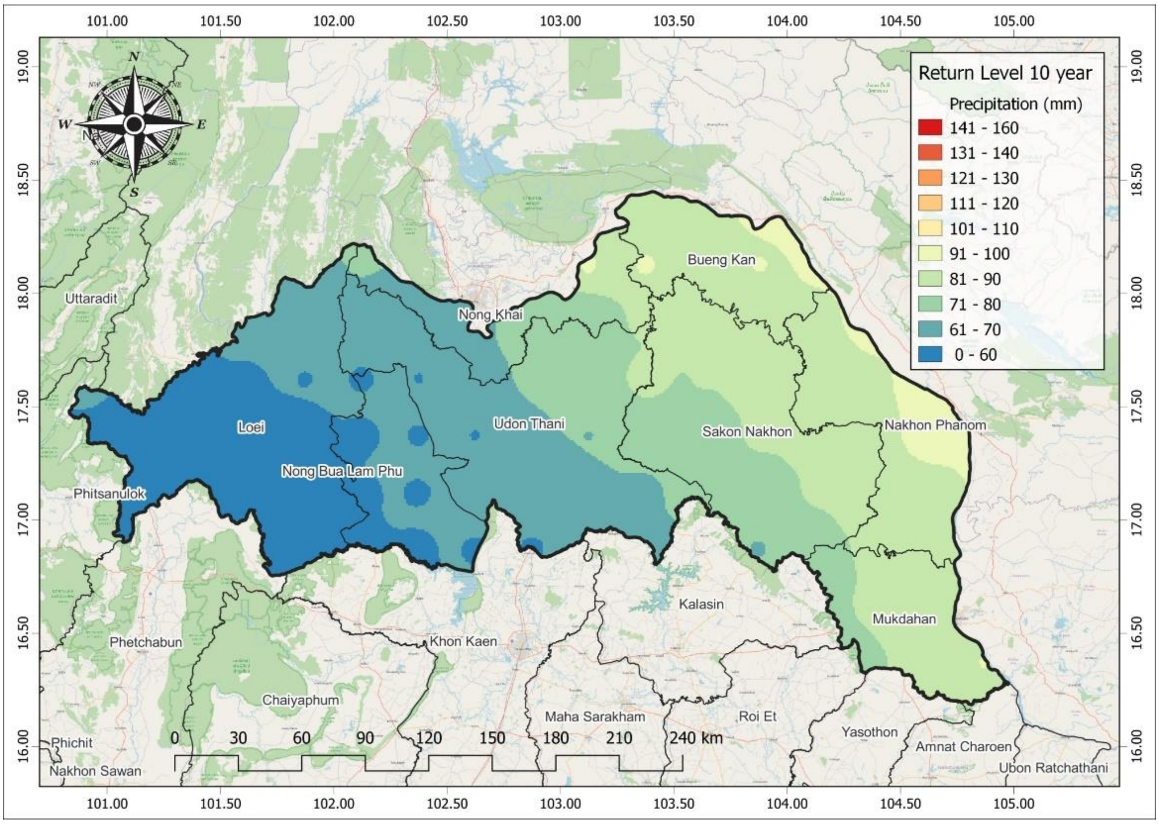

Figure 5. Estimated values for the return levels in 10 years with Generalized Pareto Distribution (GPD) of daily extreme accumulated precipitation

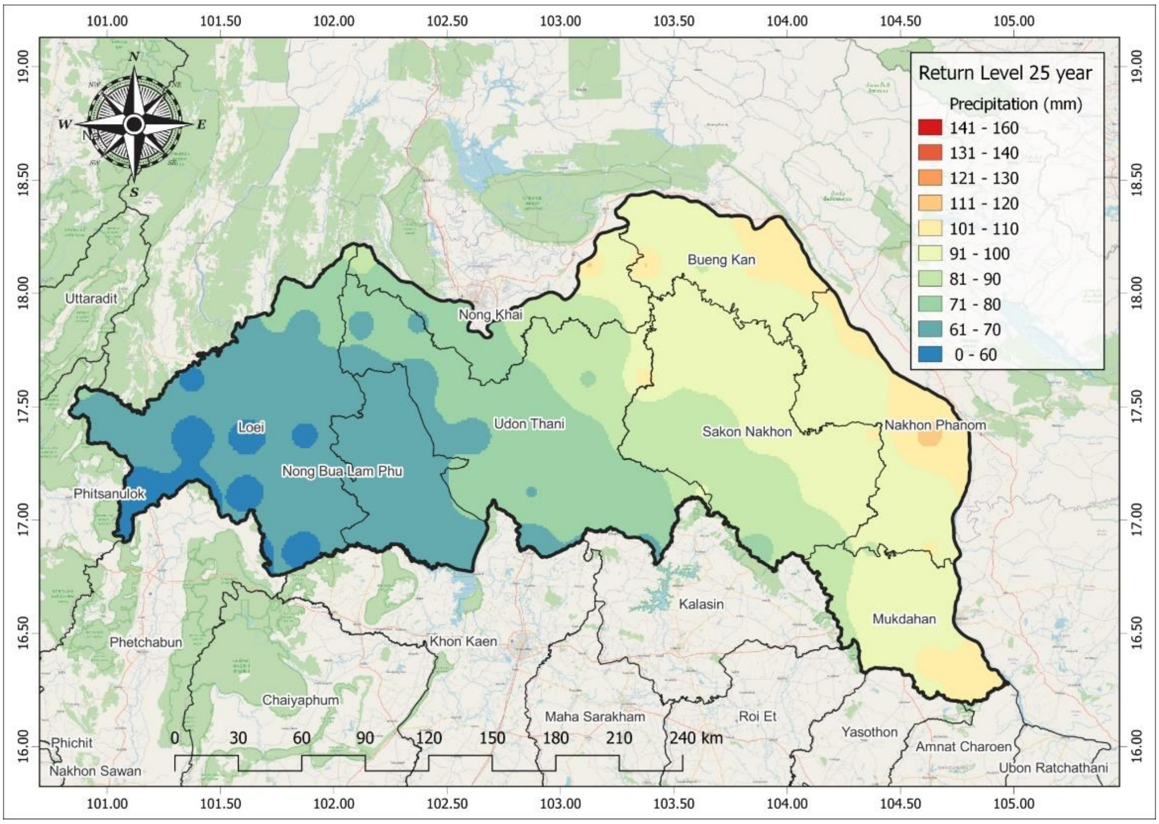

Figure 6. Estimated values for the return levels in 25 years with Generalized Pareto Distribution (GPD) of daily extreme accumulated precipitation

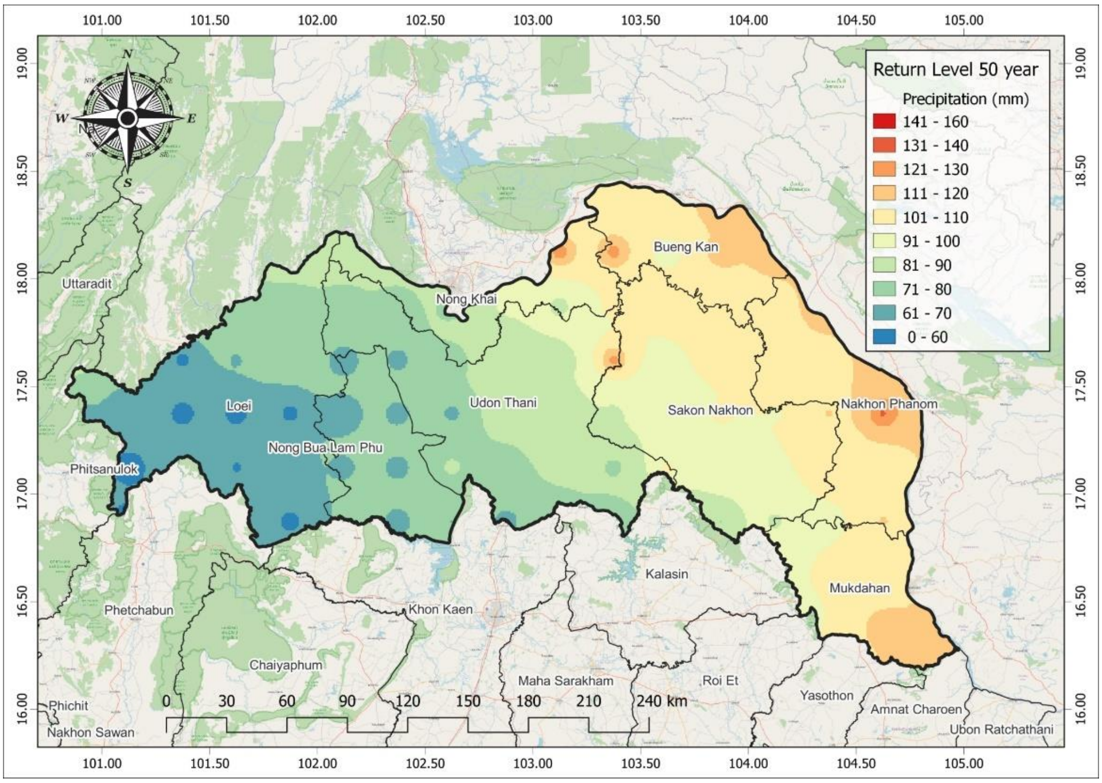

Figure 7. Estimated values for the return levels in 50 years with Generalized Pareto Distribution (GPD) of daily extreme accumulated precipitation

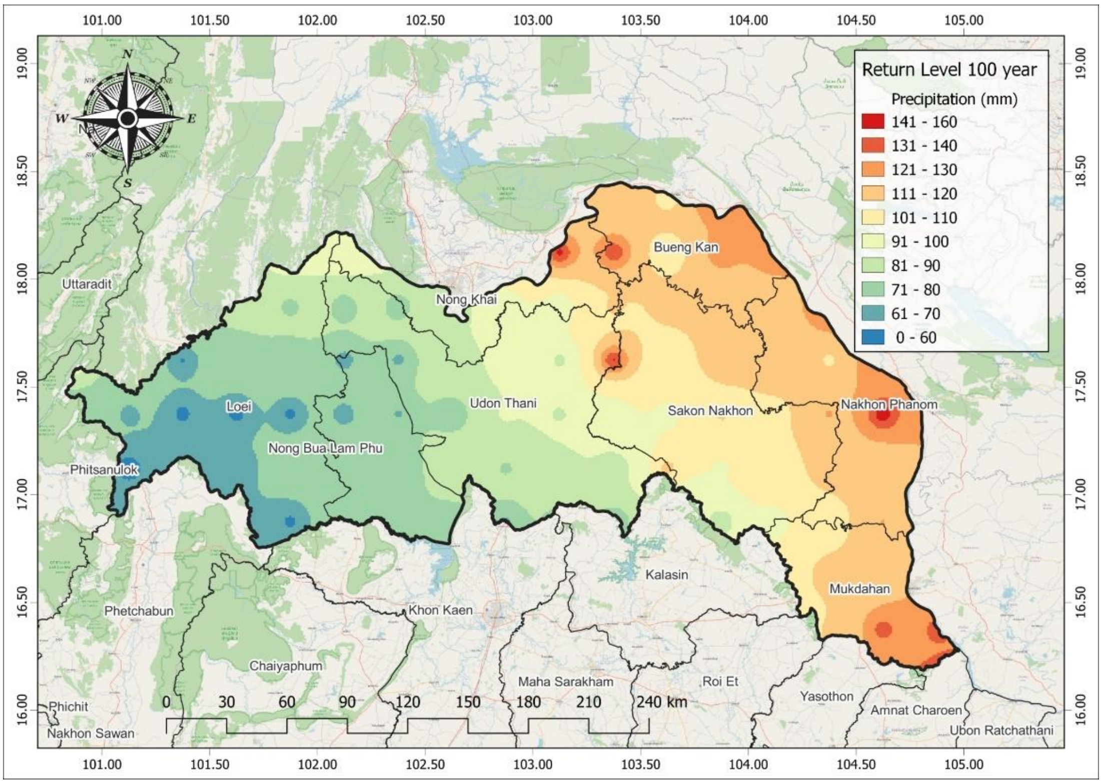

Figure 8. Estimated values for the return levels in 100 years with Generalized Pareto Distribution (GPD) of daily extreme accumulated precipitation

Figures 3 to 8 show that the areas of Bueng Kan, Nakhon Phanom, and Mukdahan, which are adjacent to Laos, have extreme precipitation higher than the other provinces in the upper northeastern provinces. In the next 25–100 years, the daily extreme precipitation in these provinces has a chance to exceed 100 mm. In the data analysis using the extreme value theory, the appropriate model was obtained through the generalized Pareto Distribution (GPD), analysis of the daily extreme value, and analysis of the return level for the daily maximum temperature. The results are shown in Table 2.

Table 2. Threshold values for temperature and appropriate temperature distribution in each grid

Province	Lat, long	u	distribution	KS
Loei	18.25,101.75	37.062	Gamma	0.7481
	17.75,101.25	37.677	Gamma	0.8774
	17.75,101.75	38.139	Gamma	0.5890
Udon Thani	18.25,102.25	36.717	Gamma	0.3035
	18.25,103.25	36.648	Gamma	0.5802
	17.75,102.75	38.238	Gamma	0.7203
	17.75,103.25	37.766	Gamma	0.5906
Nong Khai	18.25,102.75	36.203	Gamma	0.4011
	18.75,103.25	32.515	Exponential	0.8875
Nong Bua Lampu	17.75,102.25	37.634	Gamma	0.5450
Bueng Kan	18.75,103.75	621.34	Exponential	0.7269
Sakon Nakhon	18.25,103.75	36.677	Gamma	0.9865
	17.75,103.75	37.607	Gamma	0.8338
	17.75,104.25	37.173	Gamma	0.8339
Nakhon Phanom	18.25,104.25	35.496	Gamma	0.8742
	17.75,104.75	35.528	Gamma	0.8529
Mukdahan	17.25,104.25	37.598	Gamma	0.9348
	17.25,104.75	37.181	Gamma	0.5687
	16.75,104.75	37.390	Gamma	0.8286

Table 2 shows the provinces in the upper northeastern region of Thailand, including Loei, Nong Khai, Bueng Kan, Nong Bua Lampu, Udon Thani, Nakhon Phanom, Mukdahan, and Sakon Nakhon. In the analysis, there were a total of 19 grid points. There were only 2 grid points in 2 provinces, i.e., Nong Khai and Bueng Kan, which distributed temperature in exponential form. For the other 17 grid points, the temperature was distributed in gamma form. The prediction values for the return levels in 2 years, 5 years, 10 years, 25 years, 50 years, and 100 years were used to build contour graphs by using GIS Kreiging interpolation. The results are shown in Figures 9 to 14.

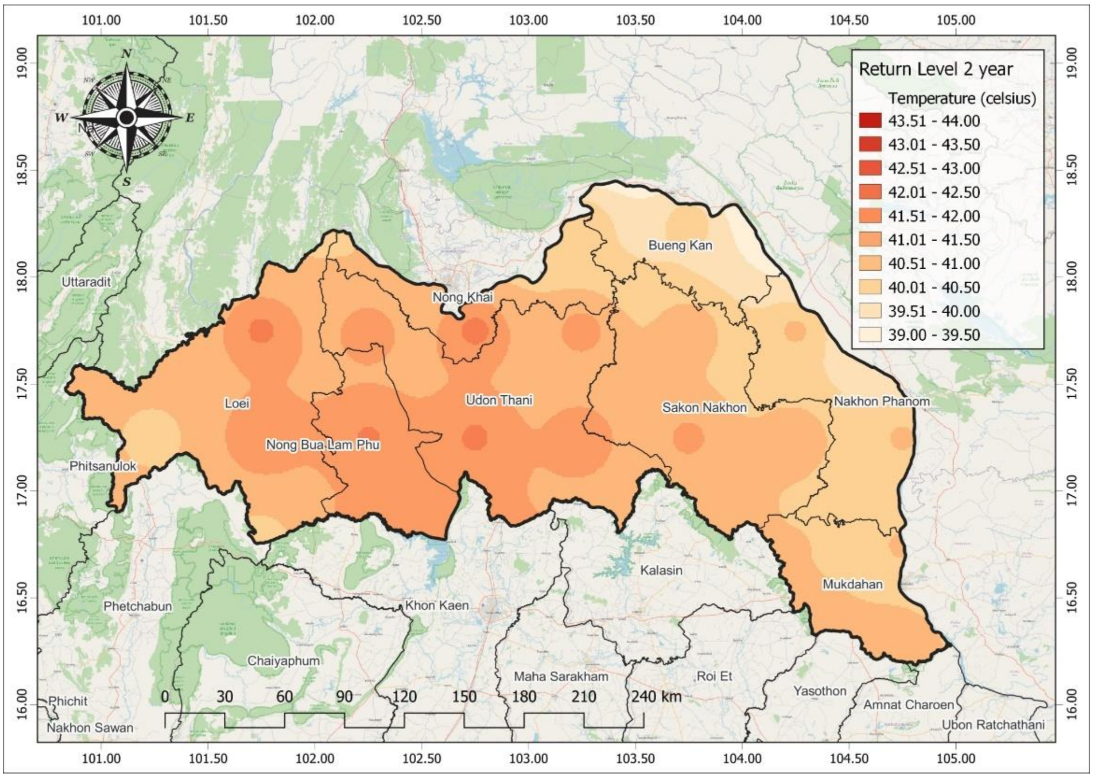

Figure 9. Estimated values for the return levels in 2 years with Generalized Pareto Distribution (GPD) of daily extreme temperature

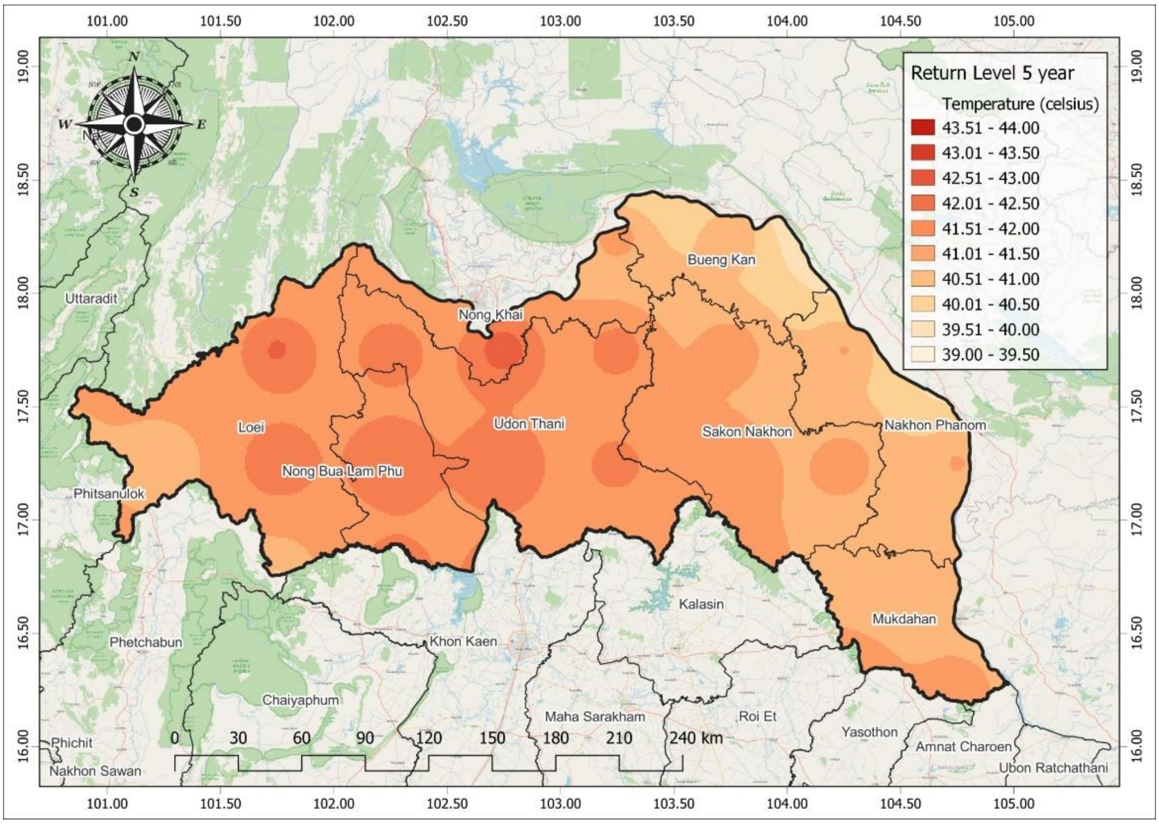

Figure 10. Estimated values for the return levels in 5 years with Generalized Pareto Distribution (GPD) of daily extreme temperature

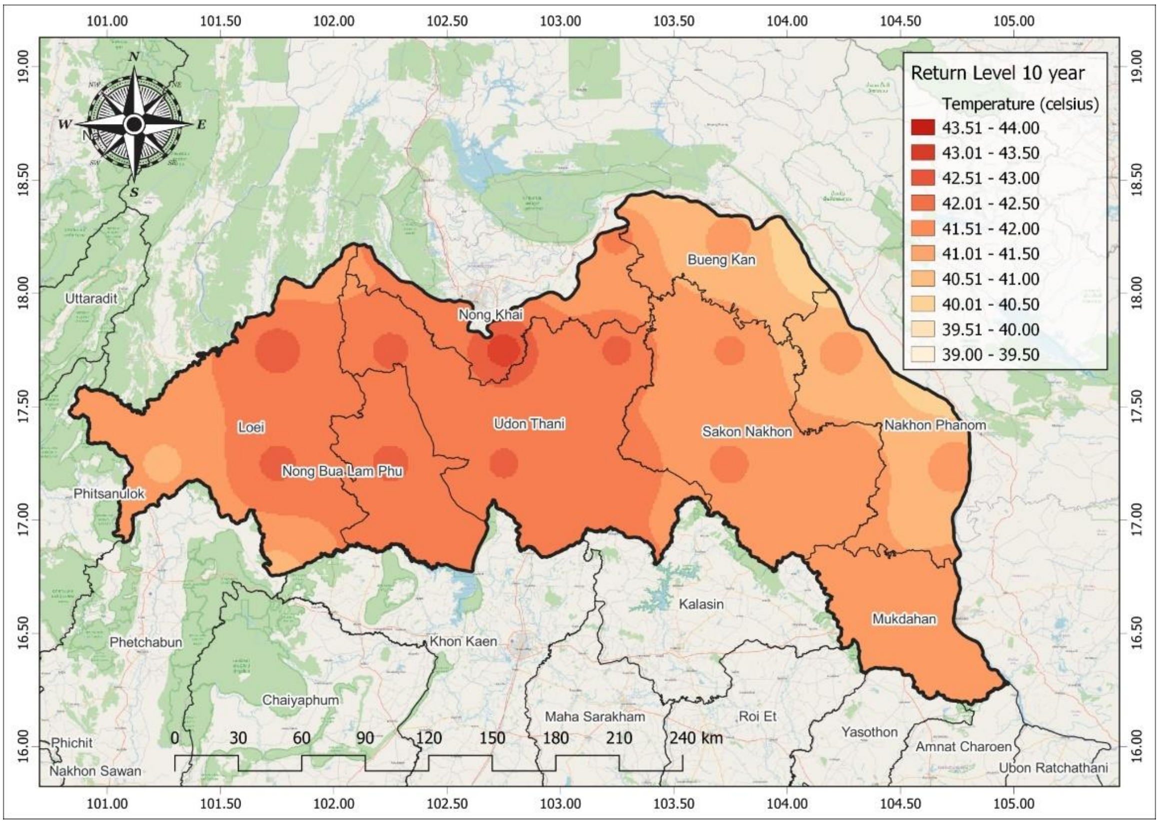

Figure 11. Estimated values for the return levels in 10 years with Generalized Pareto Distribution (GPD) of daily extreme temperature

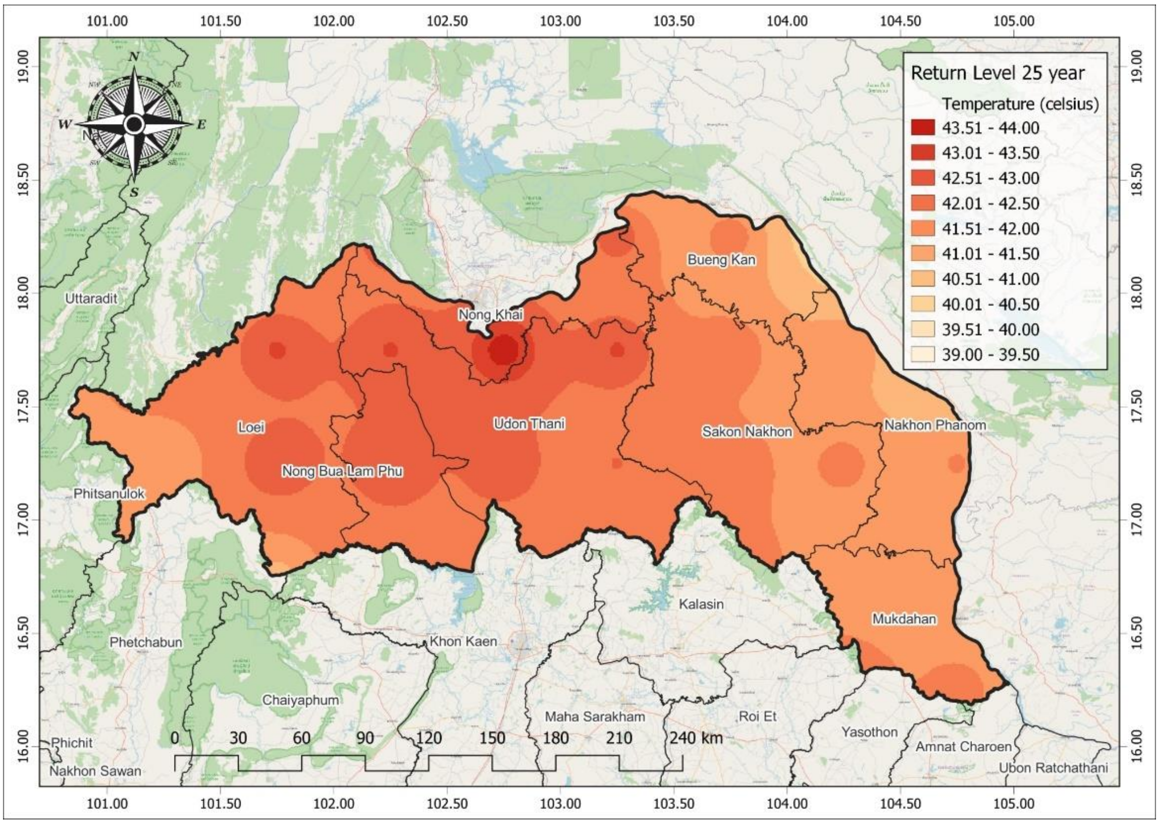

Figure 12. Estimated values for the return levels in 25 years with Generalized Pareto Distribution (GPD) of daily extreme temperature

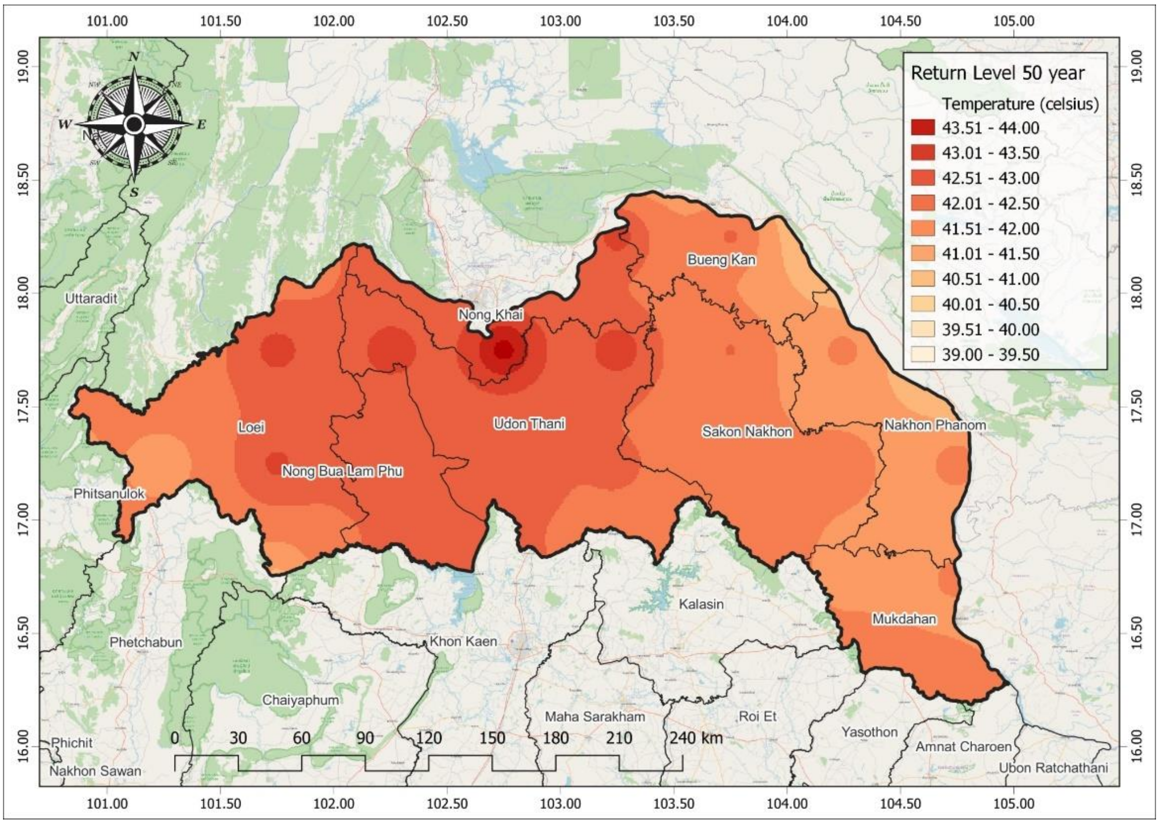

Figure 13. Estimated values for the return levels in 50 years with Generalized Pareto Distribution (GPD) of daily extreme temperature

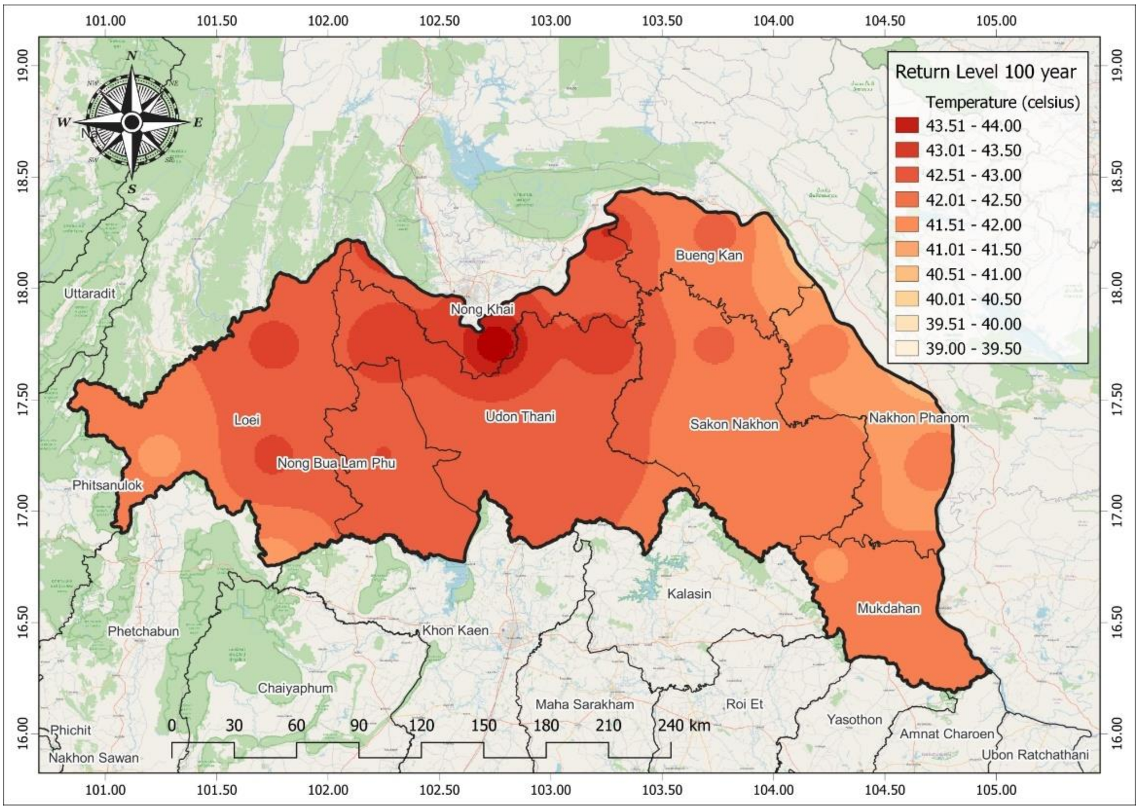

Figure 14. Estimated values for the return levels in 100 years with Generalized Pareto Distribution (GPD) of daily extreme temperature

Figures 9 to 14 show that the areas of Loei, Nong Khai, Nong Bua Lampu, and Udon Thani have higher daily extreme temperatures than the other provinces in the upper northeastern provinces. Nong Khai has a higher daily maximum temperature than the other provinces at all return levels.

4. Conclusion

The results of the analysis were obtained by using the peak-over-threshold approach with Generalized Pareto Distribution (GPD) for daily extreme precipitation and daily extreme temperature as reanalysis data. The goodness of fit test was performed to examine the appropriateness of the distribution from the analysis using Kolmogorov-Smirnov Statistics (KS Test). According to the study on the distribution of daily maximum accumulated precipitation in the upper northeastern region of Thailand, there were 71 grid points in total, with 68 grid points exhibiting an exponential distribution and 3 grid points exhibiting a gamma distribution. In the study on the distribution of daily maximum temperature in the upper northeastern region of Thailand, there were 19 grid points in total, with 17 grid points showing gamma distribution and 2 grid points showing exponential distribution.

Regarding the estimated return level of precipitation, it was found that the levels of precipitation in Bueng Kan Province, Nakhon Phanom Province, and Mukdahan Province were higher than in the other provinces because the rainy season in Thailand extends from around the middle of May to the middle of October every year. Therefore, these provinces tend to be affected by monsoon winds originating in the South China Sea and the Pacific Ocean in the northwest, which move through Vietnam, Cambodia, and Laos before entering Thailand. The upper northeastern provinces affected first include Nakhon Phanom Province, Mukdahan Province, and Bueng Kan Province. It was found that the areas with higher precipitation had lower maximum temperatures than other areas. The results of this research corresponded with those of Gong et al. [21] and Chiangpradit et al. [22], which revealed that the return level increased in correlation with precipitation increases. Similarly, the return level increased correspondingly with the temperature. In conclusion, an advantage of this research is the higher resolution of the data. Therefore, the forecasting model is more precise in the study area.

5. Declarations

5.1. Author Contributions

Conceptualization, N.C., P.G., B.K., and M.C.; methodology, P.G.; software, P.G.; validation, B.K. and M.C.; formal analysis, N.C.; investigation, N.C. and M.C.; resources, B.K. and P.G.; data curation, B.K. and P.G.; writing—original draft preparation, N.C.; writing—review and editing, N.C.; visualization, B.K.; project administration, P.G.; funding acquisition, P.G. All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

5.3. Funding

This research was supported by the Fundamental Fund of Khon Kaen University and the National Science, Research and Innovation Fund (NSRF).

5.4. Acknowledgements

This research was supported by the Fundamental Fund of Khon Kaen University and the National Science, Research and Innovation Fund (NSRF), and Mahasarakham University. The authors would like to thank the editor and the referees.

5.5. Conflicts of Interest

The authors declare no conflict of interest.

6. References

[1] Chavez-Demoulin, V., Embrechts, P., & Hofert, M. (2016). An Extreme Value Approach for Modeling Operational Risk Losses Depending on Covariates. Journal of Risk and Insurance, 83(3), 735–776. doi:10.1111/jori.12059.

[2] Chavez-Demoulin, V., & Guillou, A. (2018). Extreme quantile estimation for β-mixing time series and applications. Insurance: Mathematics and Economics, 83, 59–74. doi:10.1016/j.insmatheco.2018.09.004.

[3] Einmahl, J. J., Einmahl, J. H. J., & de Haan, L. (2019). Limits to Human Life Span Through Extreme Value Theory. Journal of the American Statistical Association, 114(527), 1075–1080. doi:10.1080/01621459.2018.1537912.

[4] Diawara, D., Kane, L., Dembele, S., & Lo, G. S. (2021). Applying of the Extreme Value Theory for determining extreme claims in the automobile insurance sector: Case of a China car insurance. Afrika Statistika, 16(3), 2883–2909. doi:10.16929/as/2021.2883.188.

[5] de Haan, L., Mercadier, C., & Zhou, C. (2016). Adapting extreme value statistics to financial time series: dealing with bias and serial dependence. Finance and Stochastics, 20(2), 321–354. doi:10.1007/s00780-015-0287-6.

[6] Gkillas, K., & Katsiampa, P. (2018). An application of extreme value theory to cryptocurrencies. Economics Letters, 164, 109–111. doi:10.1016/j.econlet.2018.01.020.

[7] Drees, H., de Haan, L., & Turkman, F. (2018). Extreme value estimation for discretely sampled continuous processes. Extremes, 21(4), 533–550. doi:10.1007/s10687-018-0313-0.

[8] Tamošaitienė, J., Yousefi, V., & Tabasi, H. (2021). Project portfolio construction using extreme value theory. Sustainability (Switzerland), 13(2), 1–13. doi:10.3390/su13020855.

[9] Opala, N., Fischer, A., & Svoboda, M. (2022). Modeling tail-dependence of crypto assets with extreme value theory: Perspectives of risk management in banks. Risk Governance and Control: Financial Markets and Institutions, 12(4), 67–77. doi:10.22495/rgcv12i4p5.

[10] Embrechts, P., Klüppelberg, C., Mikosch, T. (2023). Modern Extreme Value Theory at the Interface of Risk Management, Bayesian Networks and Heavy-Tailed Time Series. Mathematics Going Forward, Lecture Notes in Mathematics, 2313, Springer, Cham, Switzerland. doi:10.1007/978-3-031-12244-6_10.

[11] Külekci, B. Y., Karabey, U., & Selcuk-Kestel, A. S. (2023). Assessment of dependent risk using extreme value theory in a time-varying framework. Hacettepe Journal of Mathematics and Statistics, 52(1), 248–267. doi:10.15672/hujms.992699.

[12] Yousof, H. M., Tashkandy, Y., Emam, W., Ali, M. M., & Ibrahim, M. (2023). A New Reciprocal Weibull Extension for Modeling Extreme Values with Risk Analysis under Insurance Data. Mathematics, 11(4), 966. doi:10.3390/math11040966.

[13] Canton Enriquez, D., Niembro-Viceña, J. A., Muñoz Mandujano, M., Alarcon, D., Arcadia Guerrero, J., Gonzalez Garcia, I., Montes Gutierrez, A. A., & Gutierrez-Lopez, A. (2022). Application of probabilistic models for extreme values to the COVID-2019 epidemic daily dataset. Data in Brief, 40, 107783. doi:10.1016/j.dib.2021.107783.

[14] Chutiman, N., Guayjarernpanishk, P., Chiangpradit, M., Busababodhin, P., Rattanawan, S., Phoophiwfa, T., & Kong-ied, B. (2022). The Risk Area Assessment on Re-emerging Diseases in Elderly People by Using Extreme Value Theory. International Journal of Mathematics and Computer Science, 17(1), 243–253.

[15] Einmahl, J. H. J., Li, J., & Liu, R. Y. (2015). Bridging centrality and extremity: Refining empirical data depth using extreme value statistics. Annals of Statistics, 43(6), 2738–2765. doi:10.1214/15-AOS1359.

[16] Vignotto, E., & Engelke, S. (2020). Extreme value theory for anomaly detection – the GPD classifier. Extremes, 23(4), 501–520. doi:10.1007/s10687-020-00393-0.

[17] Croitoru, A. E., & Piticar, A. (2013). Changes in daily extreme temperatures in the extra-Carpathians regions of Romania. International Journal of Climatology, 33(8), 1987–2001. doi:10.1002/joc.3567.

[18] Olmo, M., Bettolli, M. L., & Rusticucci, M. (2020). Atmospheric circulation influence on temperature and precipitation individual and compound daily extreme events: Spatial variability and trends over southern South America. Weather and Climate Extremes, 29. doi:10.1016/j.wace.2020.100267.

[19] Tošić, I., Putniković, S., Tošić, M., & Lazić, I. (2021). Extreme temperature events in serbia in relation to atmospheric circulation. Atmosphere, 12(12), 1584. doi:10.3390/atmos12121584.

[20] Busababodhin, P., Chiangpradit, M., Papukdee, N., Ruechairam, J., Ruanthaisong, K., & Guayjarernpanishk, P. (2021). Extreme Value Modeling of Daily Maximum Temperature with the r-Largest Order Statistics. The Journal of Applied Science, 20(1), 28–38. doi:10.14416/j.appsci.2021.01.003.

[21] Gong, X., Wang, X., Li, Y., Ma, L., Li, M., & Si, H. (2022). Observed Changes in Extreme Temperature and Precipitation Indices on the Qinghai-Tibet Plateau, 1960–2016. Frontiers in Environmental Science, 10, 888937. doi:10.3389/fenvs.2022.888937.

[22] Chiangpradit, M., Busababodhin, P., Chutiman, N., Guayjarernpanishk, P., & Kong-ied, B. (2022). Probability Analysis for Maximum Temperature in Northeast of Thailand. International Journal of Mathematics and Computer Science, 17(2), 589–599.

[23] Coles, S., & Nadaraja, S. (2001). An Introduction to statistical modeling of extreme values. Springer, London United Kingdom.

[24] Zhang, J. (2002). Powerful Goodness-of-fit Tests Based on the Likelihood Ratio. Journal of the Royal Statistical Society Series B: Statistical Methodology, 64(2), 281–294. doi:10.1111/1467-9868.00337.