Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 8, No. 5, October, 2024

Air Pollution Forecasting in a Regional Context for Sustainable Management

Pannarat Guayjarernpanishk [1*], Nipaporn Chutiman [2], Narumol Piwpuan [1], Butsakorn Kong-ied [2], Monchaya Chiangpradit [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 44150, Thailand.*

Abstract

The aim of this research was to develop and apply a statistical model that can be used to forecast long-term daily maximum particulate matter with a diameter of less than 2.5 microns (PM2.5) concentrations. In order to predict the daily maximum PM2.5 concentrations in the northeastern region of Thailand, the extreme value theory was analyzed, and an appropriate distribution model was identified by employing the Generalized Pareto distribution (GPD). The data of daily maximum PM2.5 concentrations during the years 2021–2023 obtained from six stations was used. These stations are located in Khon Kaen, Loei, Nakhon Ratchasima, Nong Khai, Nakhon Phanom, and Ubon Ratchathani provinces. The results of this study reveal that the GPD is appropriate based on the results of Kolmogorov-Smirnov Statistics Test. Estimating the return levels during the following return periods: 2 years, 5 years, 10 years, 25 years, 50 years, and 100 years showed that the area in the upper northeastern region, particularly Loei and Nakhon Phanom, has daily maximum PM2.5 concentrations above 500 micrograms per cubic meter. These results can also be used as information to support decision-making when conducting response planning in high-risk areas, which can be helpful for efficient resource planning and prevention actions.

Keywords:

Extreme Value Theory;
Peak Over Threshold;
Generalized Pareto Distribution;
Air Pollution; PM2.5.

Article History:

Received:	01	April	2024
Revised:	19	August	2024
Accepted:	03	September	2024
Published:	01	October	2024

1- Introduction

Over the past decade, the general air quality in Thailand has shown a consistent decline, with many regions still experiencing levels of air pollutants exceeding standard thresholds. The main pollutants that are still a major problem include "particulate matter with a diameter of less than 2.5 microns". Each area of the country has its own unique characteristics when it comes to the sources of air pollution, which are mostly related to energy-intensive operations in the transportation, electricity generation, and industrial sectors. Natural resource depletion, increasing environmental and pollution issues, and the ongoing trend of rising greenhouse gas emissions pose significant challenges to meeting the country's sustainable economic growth.

A number of industries in the northeastern region of Thailand contribute to the comparatively high levels of air pollution there, including rice mills, tapioca starch factories, jute bale factories, and sugarcane mills. Additionally, biomass burning is a common source of air pollution in this region, notably from January to April [1-3], accompanied by substantial greenhouse gas emissions from such activity. Moreover, human activities that release particulate matter pollution largely affect air quality [4-6]. The northeastern region is divided from other regions by mountains, with the

[*] **CONTACT**: panngu@kku.ac.th

DOI: http://dx.doi.org/10.28991/ESJ-2024-08-05-024

Korat Basin and Sakon Nakhon Basin, which slope from west to east. The general weather is hot and humid, alternating with the dry season, which influences the strength of the wind and the dispersion of air pollution. In addition, the region is also affected by transboundary haze. During the summer, more than 7,000 hotspots in the Great Mekong Subregion countries, including Laos, Myanmar, Vietnam, and Cambodia, cause transboundary haze. Furthermore, border trade is currently growing exponentially. There are three significant border trade routes and border crossings from the northeastern region to Lao PDR, Southern China, and Vietnam. As a result, there are more industrial facilities in the region and a large amount of goods being shipped to neighboring countries, which inevitably leads to an increase in air pollution. Thus, air pollution forecasting is crucial. The forecast's results show how much pollution there will be in the future, which is helpful in resource planning for efficient prevention. Using a faulty model would result in highly inaccurate forecasts, which would cause planning decisions to fluctuate and become inconsistent with future situations. Many researchers have developed models to forecast PM2.5 concentrations by creating prediction models with the Box–Jenkins method [7-11]. Additionally, some researchers have studied the effects of meteorological factors on PM2.5 concentrations [12-18], but these seek to produce short-term forecasts. Long-term forecasting is therefore crucial and required. Extreme value theory has been utilized by numerous researchers to develop models that predict long-term extreme values. These models have been applied in the fields of meteorology [19-25], medicine [26-29], insurance [30-32], risk management [33, 34], and machine learning [35, 36].

The researcher is interested in studying the model development to forecast long-term PM2.5 concentrations for the northeastern region of Thailand using the extreme value theory. Generally, PM2.5 concentrations are collected from ground stations, and spatial interpolation such as the nearest-neighbor method [37] and the kriging interpolation method [38] are used to extend from each point to a wider plane. Thus, in this research, the return levels of extreme PM2.5 concentrations are estimated, along with GIS, to help predict PM2.5 concentrations that can occur in the future. Relevant agencies therefore can use the information from the research results to work more efficiently. This will enable them to provide information to the public so that people can take preventive measures against air pollution and minimize acts that have an impact on it.

2- Domain of Experiment and Methodology

2-1- Domain of Experiment

There are 20 provinces in the northeastern region of the Thailand study area, including Amnat Charoen, Bueng Kan, Buriram, Chaiyaphum, Kalasin, Khon Kaen, Loei, Maha Sarakham, Mukdahan, Nakhon Phanom, Nakhon Ratchasima, Nong Bua Lamphu, Nong Khai, Roi Et, Sakon Nakhon, Sisaket, Surin, Ubon Ratchathani, Udon Thani, and Yasothon provinces, which are shown in Figure 1.

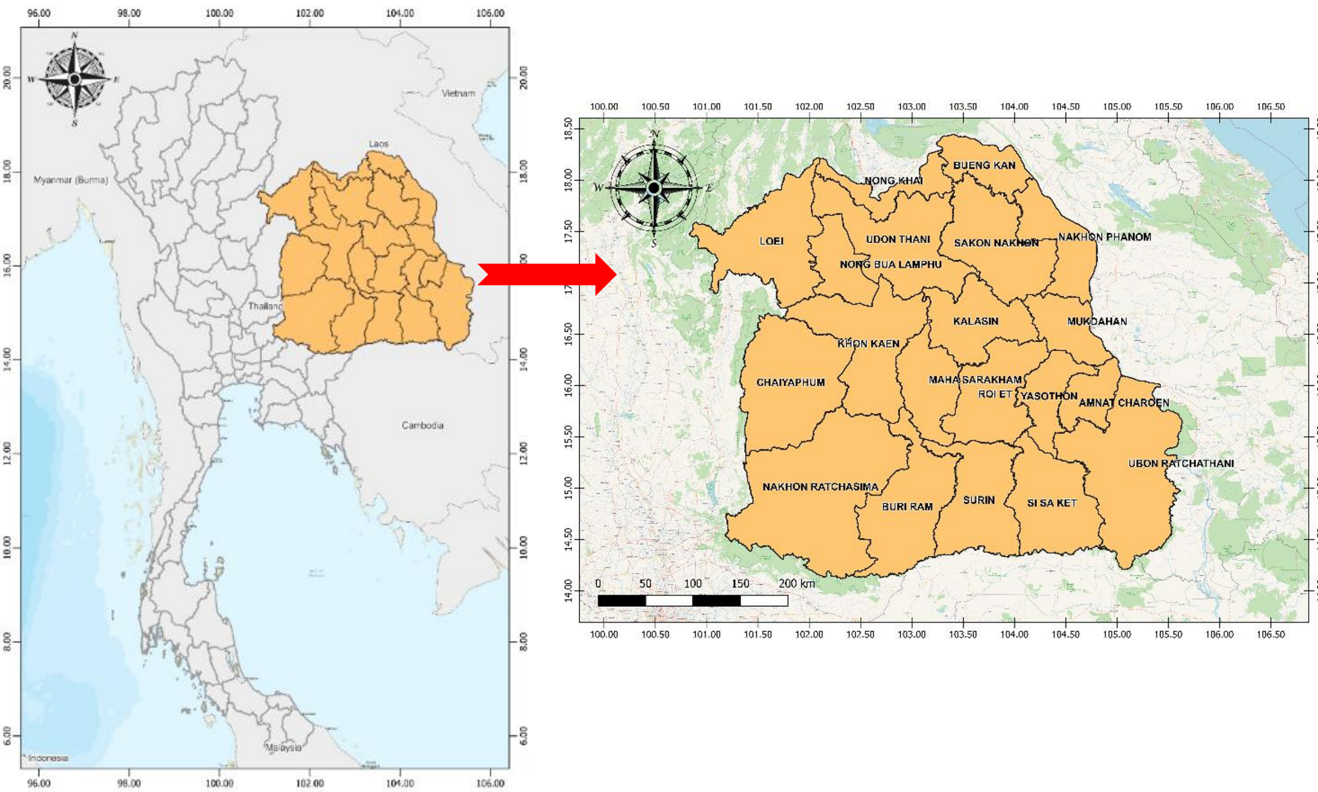

Figure 1. The provinces of the northeastern region of Thailand in the study area

2-2- Data Preparation

PM2.5 concentrations data is obtained from Thailand's Pollution Control Department. There are 11 stations monitoring concentrations of PM2.5 in the northeastern region, including Buriram, Chaiyaphum, Khon Kaen, Loei, Mukdahan, Nakhon Phanom, Nakhon Ratchasima, Nong Khai, Sakon Nakhon, Ubon Ratchathani, and Udon Thani provinces. However, due to data collection limitations, only six monitoring stations that have complete data for the years 2021-2023 will be used, which are those in Khon Kaen, Loei, Nakhon Phanom, Nakhon Ratchasima, Nong Khai, and Ubon Ratchathani provinces. Data of daily maximum PM2.5 concentrations are used for analysis.

3- Research Methodology

3-1- Extreme Value Theory

Let $X_1, X_2, \dots, X_n$ be a sequence of continuous random variables, which are independent and common distribution function of $F(x) = Pr(X_i \leq x)$ and define M_n as the maximum values for the random variables size n, i.e., $M_n = \max\{X_1, X_2, \dots, X_n\}$. According to the distribution function of M_n, the formula of distribution for the maximum value is as follows [39].

$$\begin{aligned} Pr\{M_n \leq z\} &= Pr\{X_1 \leq z, X_2 \leq z, \dots, X_n \leq z\} \\ &= Pr(X_1 \leq z) Pr(X_2 \leq z) \dots Pr(X_n \leq z) \\ &= [F(z)]^n \end{aligned} \tag{1}$$

where, z is a constant. Since the distribution function (F) of the random variables is unknown, therefore, the statistical theory is used to estimate the random variables. Although the error of estimation for F is minimized, it could lead to an error for F^n with increasing values. If we are interested in F^n behaviour, when $n \to \infty$, therefore the random variable of M_n is transformed to pattern of linear renormalization form, i.e., $M_n^* = \frac{M_n - b_n}{a_n}$, for sequence of constants $\{a_n > 0\}$ and $\{b_n\}$, the $\{a_n\}$ and $\{b_n\}$ stabilized the location and scale of M_n^* as n is increasing. Therefore, we seek to limit the distribution of M_n^*, with appropriate values for $\{a_n\}$ and $\{b_n\}$, more than M_n.

There are two methods for considering the F^n behaviour, when $n \to \infty$, i.e., the block maxima method and the Peak over Threshold method. The process of the block maxima method represented by M_n were the maximum values for the studied variables in each time block of the process and using the Generalized extreme value distribution (GEV). Whereas the process of the Peak over Threshold method represented by M_n. All values for the studied variables that are greater than or equal to the threshold and the Generalized Pareto distribution (GPD) is applied for analyzing.

3-2- The Block Maxima Method

Theorem 1: Extremal Types Theorem [39]:

If the constant sequence values are $\{a_n\}$ and $\{b_n\}$, when $a_n > 0$ as $n \to \infty$, it would be as follows

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

In the case of non-degenerate distribution, $G(z)$, the distribution is one of the following distributions

$$\begin{aligned} &\text{(I) } G(z) = exp\left\{-exp\left(-\left(\frac{z-b}{a}\right)\right)\right\}; \; -\infty < z < \infty, \\ &\text{(II) } G(z) = \begin{cases} 0 & ; \; z \leq b, \\ exp\left\{-\left(-\left(\frac{z-b}{a}\right)^{-\alpha}\right)\right\} & ; \; z > b, \end{cases} \\ &\text{(III) } G(z) = \begin{cases} exp\left\{-\left(-\left(\frac{z-b}{a}\right)^{-\alpha}\right)\right\} & ; \; z < b, \\ 1 & ; \; z \geq b, \end{cases} \end{aligned} \tag{3}$$

for parameters $a > 0$, b and, in the case of families II and III, $\alpha > 0$.

These three distribution types of the extreme value theorem are known as Gumbel, Fréchet, and Weibull, respectively.

3-3- The Peak over Threshold Method

The random variables $X_1, X_2, \dots, X_n$ were independently and identically distributed. The distribution function peaked over the threshold when X was greater than the threshold u i.e., $H_u(y) = Pr(X - u | X > u)$ where $Y = X - u$ and u was set at a fairly high value. The conditional distribution $H_u(y)$ was estimated by GPD. GPD is the one of suitable distribution when using the Peak over Threshold method to analyze extreme event data. The concepts of selected data of the Peak over Threshold method were as follows.

Let Y be a random variable of GPD, $Y \sim \text{GPD}(\sigma, \xi)$, where σ is the scale parameter and ξ is the shape parameter. The cumulative distribution function (CDF) of $x - u$ is conditional, i.e. $x > u$, where u is a threshold, as in the following Equation 4.

$$H(y) = \begin{cases} 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-\frac{1}{\xi}}; \; y > 0, 1 + \frac{\xi y}{\sigma} > 0 \text{ and } \xi \neq 0 \\ 1 - exp\left(-\frac{y}{\sigma}\right); \; y > 0, \sigma > 0 \quad \text{ and } \xi = 0 \end{cases} \tag{4}$$

There are three special cases of GPD which depend on the shape parameter (ξ), the first, $\xi < 0$ is called gamma distribution, the second, $\xi \to 0$ is called exponential distribution, and the third, $\xi > 0$ is called Pareto distribution.

The return level of GPD formulas were as in the following Equation 5.

$$z = \begin{cases} u + \frac{\sigma}{\xi}\left[\left(\frac{1}{T}\eta_y \lambda_u\right)^{\xi} - 1\right]; \; \sigma > 0 \text{ and } \xi \neq 0 \\ u + \sigma log\left(\frac{1}{T}\eta_y \lambda_u\right) \quad ; \; \sigma > 0 \text{ and } \xi = 0 \end{cases} \tag{5}$$

where z refers to the return level, T refers to the return period, η_y refers to the average number of days per year, λ_u refers to the probability of the peak over threshold.

3-4- Kolmogorov-Smirnov Statistics Test

The Kolmogorov-Smirnov Statistics Test (*KS* Test) is a nonparametric goodness-of-fit test for distribution. Let $X_1, X_2, \ldots, X_n$ be a sample from a population with continuous and order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, but the cumulative distribution function (CDF) is unknown. Let $F_n(x)$ be the empirical CDF based on the sample.

To test the hypothesis $H_0: F(x) = F_0(x)$, for all x versus the alternative $H_1: F(x) \neq F_0(x)$. The Kolmogorov-Smirnov Statistics Test formula is

$$KS = \left\{\sup_{t \in (-\infty,\infty)} |F_n(t) - F_0(t)|\right\}^2 = \left(\max_{1 \le i \le n}\left[max\left\{\frac{i}{n} - F_0\left(X_{(i)}\right), F_0\left(X_{(i)}\right) - \frac{i-1}{n}\right\}\right]\right)^2 \tag{6}$$

where *KS* is the Kolmogorov-Smirnov Statistics, the well-known statistics for goodness-of-fit-tests [40].

The procedures for the research methodology are shown in Figure 2.

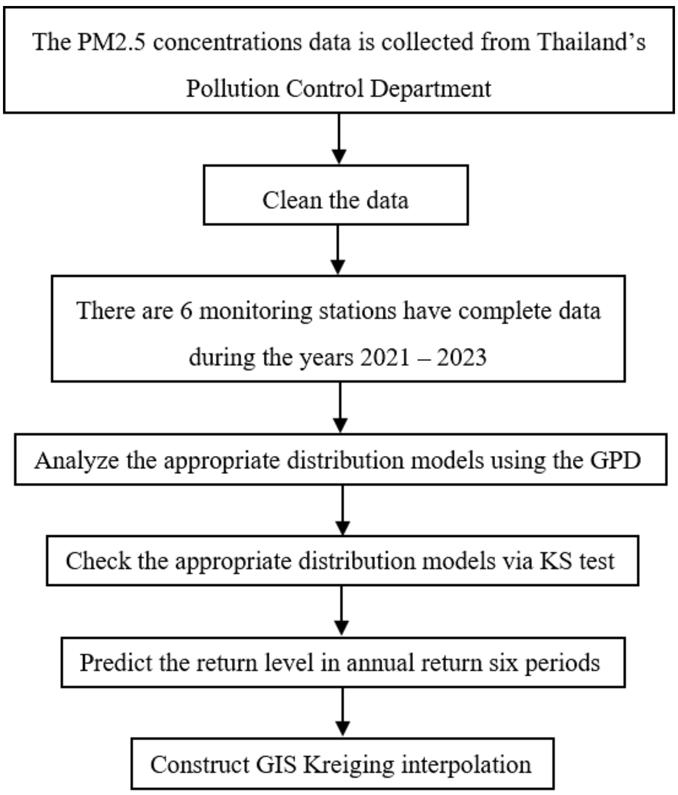

Figure 2. Procedures for the research methodology

4- Results and Discussion

The data used in this research was daily maximum PM2.5 concentrations at six stations during the years 2021–2023. The daily maximum PM2.5 concentrations at each station are shown in Table 1.

According to Table 1, the highest of daily maximum PM2.5 concentrations was found in Nakhon Phanom with 430 micrograms per cubic meter, which occurred in late March 2023. This was followed by Nong Khai with daily maximum PM2.5 concentrations of 343 micrograms per cubic meter, which occurred in late March 2023, and then Loei with the daily maximum PM2.5 concentrations of 308 micrograms per cubic meter, which occurred in early April 2023. The mean of daily maximum PM2.5 concentrations for all stations is approximately between 34 and 47 micrograms per cubic meter, which is above Thailand's standard criteria of 37.5 micrograms per cubic meter for all stations except Loei station.

Table 1. Daily maximum PM2.5 concentrations at the six stations (Unit: micrograms per cubic meter)

Stations	Median	Mean	Min	Max	1st Quartile	3rd Quartile
Khon Kaen	31.00	38.96	0.00	154.00	21.00	57.50
Nakhon Ratchasima	32.00	38.71	12.00	114.00	24.00	50.00
Loei	23.00	34.82	0.00	308.00	11.90	50.00
Nong Khai	28.00	45.99	5.00	343.00	13.00	64.00
Ubon Ratchathani	26.00	39.22	0.00	226.00	16.00	52.00
Nakhon Phanom	33.00	46.33	0.00	430.00	19.00	57.00

In data analysis using the extreme value theory, finding a proper model was done by using the GPD to analyze extreme values and return levels. The daily maximum concentrations of PM2.5 results are presented in Table 2.

Table 2. Threshold values of PM2.5 and a proper distribution of PM2.5 concentrations at each station (Unit: micrograms per cubic meter)

Stations	Threshold	Distribution	*KS* test	Return levels					
				2 years	5 years	10 years	25 years	50 years	100 years
Khon Kaen	69	Gamma	0.4669	137.379	145.727	151.261	157.678	161.932	165.730
Nakhon Ratchasima	60	Gamma	0.6979	110.203	112.641	113.928	115.120	115.749	116.211
Loei	63	Exponential	0.9037	264.079	322.596	372.467	446.788	510.127	580.336
Nong Khai	88	Exponential	0.4914	282.637	315.549	339.456	369.803	391.846	413.133
Ubon Ratchathani	72	Exponential	0.2994	219.232	241.590	257.355	276.787	290.489	303.388
Nakhon Phanom	77.7	Exponential	0.9056	343.838	410.327	464.436	541.405	604.042	670.755

Table 2 shows the gamma distribution of PM2.5 concentrations in Khon Kaen and Nakhon Ratchasima and the exponential distribution of PM2.5 concentrations in the four other provinces (Loei, Nakhon Phanom, Nong Khai, and Ubon Ratchathani). The 2 years estimated values of the return levels for daily maximum PM2.5 concentrations are between 110 and 344 micrograms per cubic meter. The 5 years estimated values of the return levels for daily maximum PM2.5 concentrations are between 112 and 411 micrograms per cubic meter. Meanwhile, the 100 years estimated values of the return levels for daily maximum PM2.5 concentrations are between 116 and 671 micrograms per cubic meter. Additionally, the outcomes of using the estimated values for the return levels in 2, 5, 10, 25, 50 and 100 years to create a contour graph using GIS Kreiging interpolation are shown in Figures 3 to 8.

Figures 3 to 8 show that as the return periods increase, the return levels will also increase. The provinces bordering Laos in the northern area of the northeastern region have higher daily PM2.5 concentrations than other areas. Additionally, there is a potential that daily maximum PM2.5 concentrations will exceed 400 micrograms per cubic meter within the next 50-100 years.

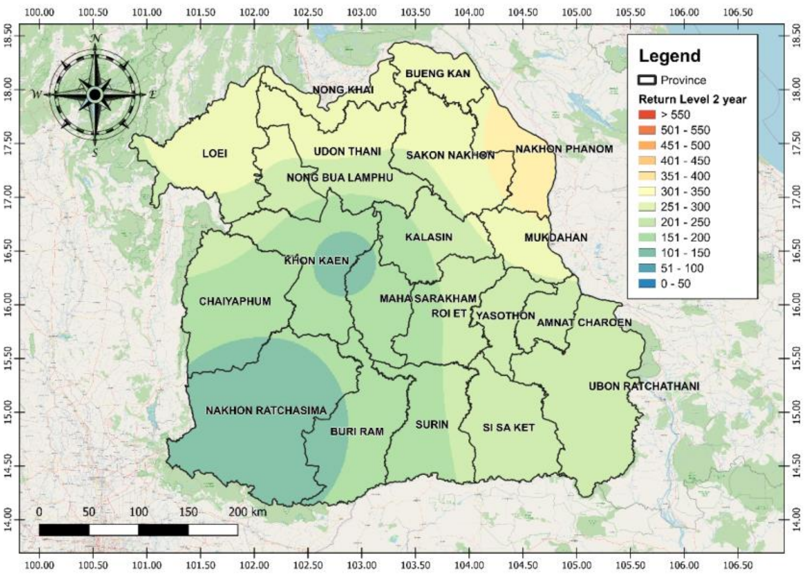

Figure 3. **Forecasted daily maximum PM2.5 concentrations for 2 years**

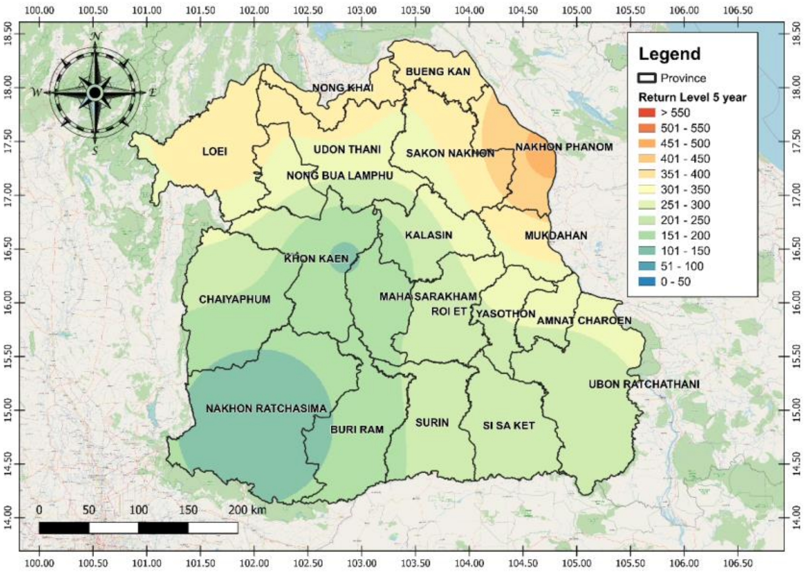

Figure 4. **Forecasted daily maximum PM2.5 concentrations for 5 years**

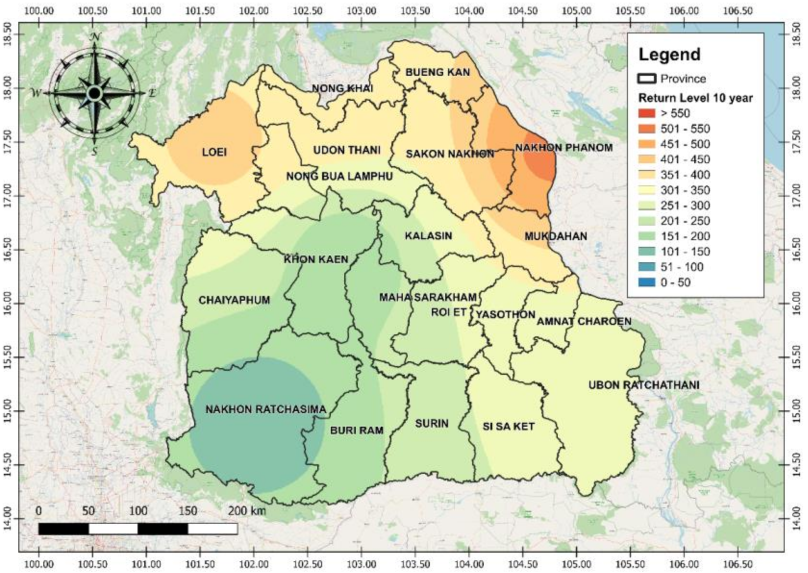

Figure 5. **Forecasted daily maximum PM2.5 concentrations for 10 years**

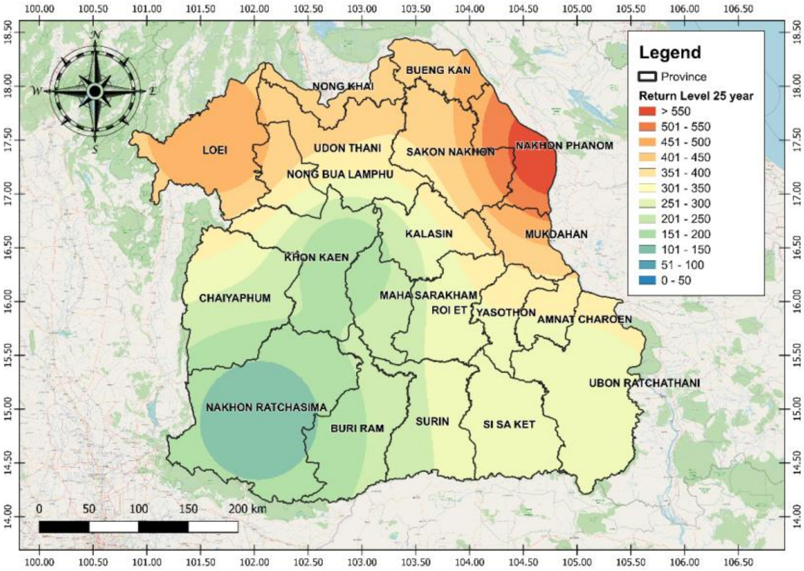

Figure 6. Forecasted daily maximum PM2.5 concentrations for 25 years

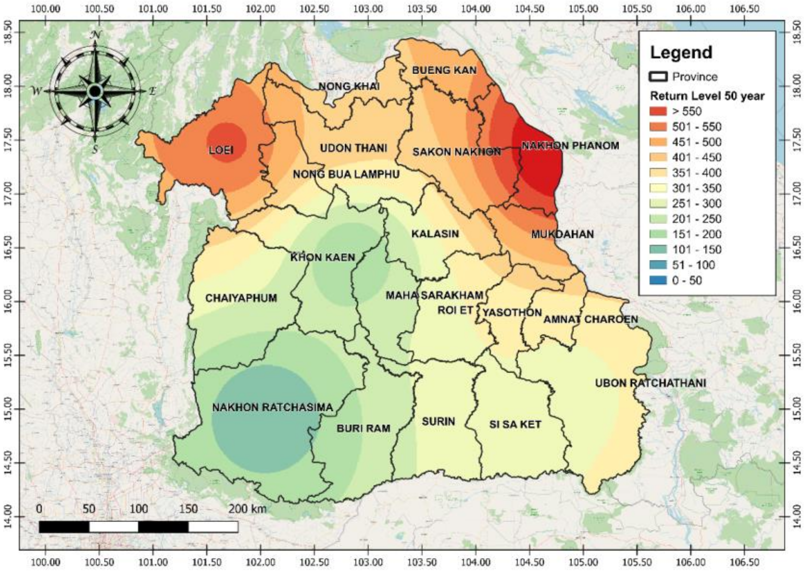

Figure 7. Forecasted daily maximum PM2.5 concentrations for 50 years

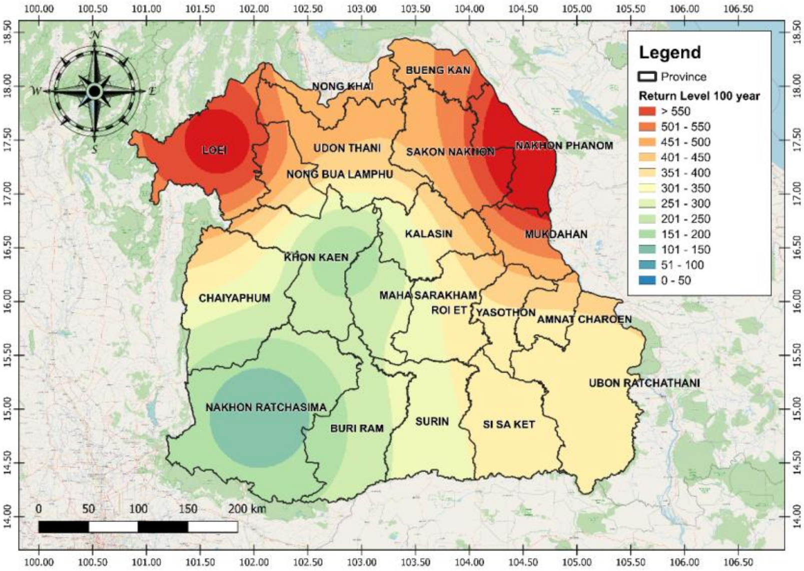

Figure 8. Forecasted daily maximum PM2.5 concentrations for 100 years

5- Conclusion

The concentrations of PM2.5 are usually high as winter gives way to summer. The factors contributing to high concentrations of PM2.5 are temperature and pressure. The temperature is hot throughout the day and cool during the evening. The weather is one factor contributing to particulate matter that fails to disperse, causing the phenomenon of temperature inversion. When air pressure is high during cold weather, the ground quickly releases heat, which cools the air above it. This causes heated air to rise and become trapped in between cool air layers, inhibiting vertical ascent of air and resulting in the accumulation of particulate matter in the atmosphere. In the northeastern region of Thailand, the period from January to February marks the change of seasons from winter to summer. However, over the last 3–4 years, the El Niño phenomenon has occurred, causing the seasonal transition period between winter and summer to extend to March, resulting in higher PM2.5 concentrations. Currently, climate change is becoming increasingly severe every year. Long-term climate forecasts are therefore important and interesting. In this research, the methods for long-term PM2.5 forecasting are studied in order to plan and deal with long-term PM2.5 changes and human lifestyle behavior modification. This differs from existing studies that seek to make short-term PM2.5 predictions [7-18].

The results of the analysis were attained by using the Peak over Threshold approach with the GPD of daily maximum PM2.5 concentrations at six stations in the northeastern region of Thailand. The goodness of fit test using Kolmogorov-Smirnov Statistics (KS Test) was then conducted to examine the appropriateness of the distribution that was obtained from the analysis. The analysis revealed that Khon Kaen and Nakhon Ratchasima provinces show the daily maximum PM2.5 concentration in a gamma distribution, while Loei, Nong Khai, Ubon Ratchathani, and Nakhon Phanom provinces display the daily maximum PM2.5 concentration in an exponential distribution. When considering estimates of the return levels of maximum PM2.5 concentrations, it is found that the northern part of the region has higher daily maximum PM2.5 concentrations than other areas. This is because it is a border area adjacent to Laos, which is affected by transboundary haze. Additionally, it is predicted that in the next 50-100 years, provinces in the upper northeastern region, especially Loei and Nakhon Phanom provinces, will have daily maximum PM2.5 concentrations above 500 micrograms per cubic meter based on the prediction of the return levels of the daily maximum PM2.5 concentrations found.

6- Declarations

6-1- Author Contributions

Conceptualization, P.G., N.C., N.P., B.K., and M.C.; methodology, P.G. and N.C.; software, M.C.; validation, P.G., N.P., and B.K.; formal analysis, M.C.; investigation, P.G., B.K., and M.C.; resources, P.G., N.P., and N.C.; data curation, P.G., N.P., and B.K.; writing—original draft preparation, N.C.; writing—review and editing, P.G., M.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.

6-2- Data Availability Statement

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

6-3- Funding and 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.

6-4- Institutional Review Board Statement

Not applicable.

6-5- Informed Consent Statement

Not applicable.

6-6- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

[1] Yin, S., Wang, X., Zhang, X., Guo, M., Miura, M., & Xiao, Y. (2019). Influence of biomass burning on local air pollution in mainland Southeast Asia from 2001 to 2016. Environmental Pollution, 254, 112949. doi:10.1016/j.envpol.2019.07.117.

[2] Amnuaylojaroen, T., Inkom, J., Janta, R., & Surapipith, V. (2020). Long range transport of Southeast Asian PM2.5 pollution to northern thailand during high biomass burning episodes. Sustainability (Switzerland), 12(23), 1–14. doi:10.3390/su122310049.

[3] Amnuaylojaroen, T. (2022). Prediction of PM2.5in an Urban Area of Northern Thailand Using Multivariate Linear Regression Model. Advances in Meteorology, 2022, 9. doi:10.1155/2022/3190484.

[4] Amnuaylojaroen, T., Barth, M. C., Emmons, L. K., Carmichael, G. R., Kreasuwun, J., Prasitwattanaseree, S., & Chantara, S. (2014). Effect of different emission inventories on modeled ozone and carbon monoxide in Southeast Asia. Atmospheric Chemistry and Physics, 14(23), 12983–13012. doi:10.5194/acp-14-12983-2014.

[5] Lee, H. H., Iraqui, O., Gu, Y., Yim, S. H. L., Chulakadabba, A., Tonks, A. Y. M., Yang, Z., & Wang, C. (2018). Impacts of air pollutants from fire and non-fire emissions on the regional air quality in Southeast Asia. Atmospheric Chemistry and Physics, 18(9), 6141–6156. doi:10.5194/acp-18-6141-2018.

[6] Lee, H. H., Iraqui, O., & Wang, C. (2019). The Impact of Future Fuel Consumption on Regional Air Quality in Southeast Asia. Scientific Reports, 9(1), 2648. doi:10.1038/s41598-019-39131-3.

[7] Zafra, C., Ángel, Y., & Torres, E. (2017). ARIMA analysis of the effect of land surface coverage on PM10 concentrations in a high-altitude megacity. Atmospheric Pollution Research, 8(4), 660–668. doi:10.1016/j.apr.2017.01.002.

[8] Zhang, L., Lin, J., Qiu, R., Hu, X., Zhang, H., Chen, Q., Tan, H., Lin, D., & Wang, J. (2018). Trend analysis and forecast of PM2.5 in Fuzhou, China using the ARIMA model. Ecological Indicators, 95, 702–710. doi:10.1016/j.ecolind.2018.08.032.

[9] Zhao, L., Li, Z., & Qu, L. (2022). Forecasting of Beijing PM2.5 with a hybrid ARIMA model based on integrated AIC and improved GS fixed-order methods and seasonal decomposition. Heliyon, 8(12), 12239. doi:10.1016/j.heliyon.2022.e12239.

[10] Bhatti, U. A., Yan, Y., Zhou, M., Ali, S., Hussain, A., Qingsong, H., Yu, Z., & Yuan, L. (2021). Time Series Analysis and Forecasting of Air Pollution Particulate Matter (PM2.5): An SARIMA and Factor Analysis Approach. IEEE Access, 9, 41019–41031. doi:10.1109/ACCESS.2021.3060744.

[11] Agarwal, A., & Sahu, M. (2023). Forecasting PM2.5 concentrations using statistical modeling for Bengaluru and Delhi regions. Environmental Monitoring and Assessment, 195(4), 502. doi:10.1007/s10661-023-11045-8.

[12] Li, X., Feng, Y. J., & Liang, H. Y. (2017). The Impact of Meteorological Factors on PM2.5 Variations in Hong Kong. IOP Conference Series: Earth and Environmental Science, 78, 12003. doi:10.1088/1755-1315/78/1/012003.

[13] Miao, Y., Liu, S., Guo, J., Yan, Y., Huang, S., Zhang, G., Zhang, Y., & Lou, M. (2018). Impacts of meteorological conditions on wintertime PM2.5 pollution in Taiyuan, North China. Environmental Science and Pollution Research, 25(22), 21855–21866. doi:10.1007/s11356-018-2327-1.

[14] Amnuaylojaroen, T. (2022). Prediction of PM2.5in an Urban Area of Northern Thailand Using Multivariate Linear Regression Model. Advances in Meteorology, 2022, 9. doi:10.1155/2022/3190484.

[15] Saiohai, J., Bualert, S., Thongyen, T., Duangmal, K., Choomanee, P., & Szymanski, W. W. (2023). Statistical PM2.5 Prediction in an Urban Area Using Vertical Meteorological Factors. Atmosphere, 14(3), 15. doi:10.3390/atmos14030589.

[16] Jayamala, R., Shanmugapriya, N., Lalitha, K., & Vijayarajan, P. (2023). A deep learning model and optimization algorithm to forecasting environment monitoring of the air pollution. Global Nest Journal, 25(10), 47–55. doi:10.30955/gnj.004759.

[17] Barid, A. J., & Hadiyanto, H. (2024). Hyperparameter optimization for hourly PM2.5 pollutant prediction. Journal of Emerging Science and Engineering, 2(1), e15. doi:10.61435/jese.2024.e15.

[18] Mohammadi, F., Teiri, H., Hajizadeh, Y., Abdolahnejad, A., & Ebrahimi, A. (2024). Prediction of atmospheric PM2.5 level by machine learning techniques in Isfahan, Iran. Scientific Reports, 14(1), 2109. doi:10.1038/s41598-024-52617-z.

[19] Chutiman, N., Busababodhin, P., P. Senapeng, P., Phoophiwfa, T. & Chiangpradit, M. (2019). Analysis Extreme Weather Event of Thailand with Spatial Modeling Theory of Extreme values: Case study in Mekong, Chi, Mun River Basin. Advances and Applications in Mathematical Sciences, 19(2), 129–138.

[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] Guayjarernpanishk, P., Chiangpradit, M., Kong-Ied, B., & Chutiman, N. (2023). Climate Forecasting Models for Precise Management Using Extreme Value Theory. Civil Engineering Journal (Iran), 9(7), 1753–1767. doi:10.28991/CEJ-2023-09-07-014.

[23] Contzen, J., Dickhaus, T., & Lohmann, G. (2023). Long-term temporal evolution of extreme temperature in a warming Earth. PLoS ONE, 18(2), 280503. doi:10.1371/journal.pone.0280503.

[24] Liu, H., Yang, F., & Wang, H. (2023). Research on Threshold Selection Method in Wave Extreme Value Analysis. Water (Switzerland), 15(20), 3648. doi:10.3390/w15203648.

[25] Maruyama, F. (2023). Analyzing of the ENSO Index Using Extreme Value Theory. Journal of Geoscience and Environment Protection, 11(06), 96–105. doi:10.4236/gep.2023.116007.

[26] Canton Enriquez, D., Niembro-Ceceña, 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

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

[34] 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.

[35] 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.

[36] 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.

[37] Just, A. C., Wright, R. O., Schwartz, J., Coull, B. A., Baccarelli, A. A., Tellez-Rojo, M. M., Moody, E., Wang, Y., Lyapustin, A., & Kloog, I. (2015). Using High-Resolution Satellite Aerosol Optical Depth to Estimate Daily PM2.5 Geographical Distribution in Mexico City. Environmental Science & Technology, 49(14), 8576–8584. doi:10.1021/acs.est.5b00859.

[38] Ma, Z., Hu, X., Huang, L., Bi, J., & Liu, Y. (2014). Estimating ground-level PM2.5 in China using satellite remote sensing. Environmental Science & Technology, 48(13), 7436–7444. doi:10.1021/es5009399.

[39] Coles, S. and Nadaraja, S. (2001). An Introduction to statistical modeling of extreme values. Springer, London. United Kingdom. doi:10.1007/978-1-4471-3675-0.

[40] 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.