Estimating Air Pollution Index in Indonesia as Effort to Increase Life Expectancy

Fanny Novika[1]

[1] Trisakti School of Insurance, Jakarta 13210, ID

novikafanny@gmail.com

Abstract. High levels of unhealthy Air Pollution Index (API) can severely impact human health and the environment. Thus, effective preparation for such risk events relies on precise estimation of unhealthy API levels. This study proposes the Generalized Pareto Distribution (GPD) Method based on the Bayesian framework to obtain accurate estimation of API exceedance in Indonesia. This study will look at the capital city of provinces with the worst API as an effort to obtain great benefits if they succeed in cleaning the air effectively so that they can increase life expectancy. To produce the model, two parameters are determined in GDP, namely scale and frequency in each capital city for province in Indonesia that have a heavy tail. Unhealthy API are governed by the above the threshold u = 100. The data used in this study are data from IQAir in Indonesia and BMKG. Furthermore, the GDP parameters will be identified and the spatial and seasonal impacts. The results show that capital city of provinces that have unhealthy API for sensitive groups are Jambi, Palangkaraya and Banjarbaru. Besides, the capital city of provinces that have very unhealthy API is Palembang. The Generalized Pareto model shows that Palembang and Palangkaraya have a higher risk of experiencing unhealthy air conditions on a regular basis.

Keywords: Air Polution Index; Hierarchical Generalized Pareto Distribution; IQAir

1 Introduction

1.1 Background

Indonesia, a country of thousands of islands and a diverse population of over 270 million people [1], has extraordinary beauty and potential. However, amidst this potential lies a complex and persistent challenge: the air pollution problem. The average Indonesian could lose 1.2 years of life expectancy at current pollution levels, according to the Air Quality Index for Life (AQLI), as air quality fails to meet World Health Organization (WHO) guidelines for concentrations of fine particulate matter (PM2.5). The pollution index, developed by Michael Greenstone and colleagues at the Energy Policy Institute at the University of Chicago (EPIC), shows that health impacts are much greater in parts of the country with high particulate pollution. Residents of Indonesia's capital, Jakarta, for example, are estimated to lose 2.3 years of life expectancy if 2016 pollution levels persist throughout their lifetimes. In some areas,

T. A. Aziz et al. (eds.), *Proceedings of the 4th Science and Mathematics International Conference (SMIC 2024)*, Advances in Physics Research 11,
https://doi.org/10.2991/978-94-6463-624-6_2

life expectancy could be reduced by more than 4 years. It is known that AQLI data shows that air quality was not a pressing problem in Indonesia two decades ago, but air quality has declined drastically in recent decades with the sharpest decline since 2013 [2].

Indonesia currently has no national standard for air pollution levels [3]. As governments begin to recognize the problem of air quality, the AQLI suggests that Indonesia stands to gain significant health benefits if it succeeds in cleaning up its air [4]. Other countries in the Asia Pacific region provide useful benchmarks. If, for example, Indonesia were to achieve sustained improvements in air quality comparable to those achieved by China in the past 5 years, Indonesians could expect to live eight months longer. Those living in the most polluted areas would benefit even more, living up to 2.5 years longer. Over the past two decades, Indonesia has seen dramatic changes in its air quality. From 1998 to 2016, Indonesia went from being one of the cleanest countries in the world to being one of the twenty most polluted, as its concentrations of particulate matter increased by 171 percent [5]. The largest spikes have occurred in recent years. Pollution more than doubled from 2013 to 2016 alone, with at least some of the increase likely due to the devastating fires (See Figure 1). Regardless of the causes, 80 percent of Indonesia's 250 million people live in areas where average particulate pollution levels exceeded WHO guidelines in 2016.

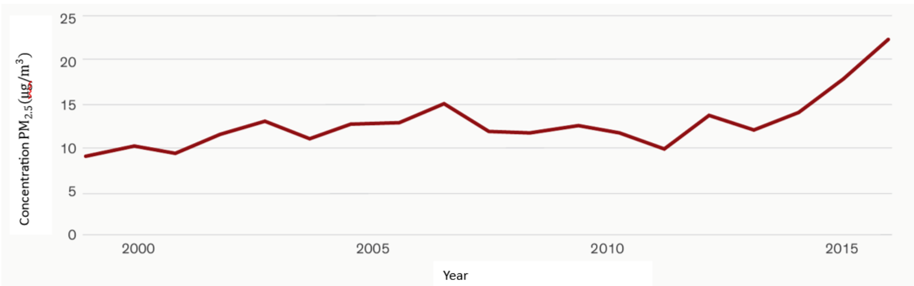

Fig. 1 Average Annual PM Concentration of Indonesia 1998-2016 [6]

Unhealthy air pollution index (API) levels have serious harmful impacts on human health and the environment [7] [8]. Therefore, proper preparation for risk events can be achieved with the help of accurate estimation of unhealthy API and its return period [9]. API indicates the health level of environmental air quality which can be stated that if the quality level exceeds the threshold of 100 for API and its pollutants it is an unhealthy level [10]. Healthy API levels have been modeled using parametric distributions, such as lognormal, while generalized Pareto distribution (GPD) has been used to characterize unhealthy API. This study will identify provincial capitals in Indonesia that have high levels of air pollution using Generalized Pareto Distribution.

1.2 Purpose of Research

Based on the background and formulation of the problems that have been put forward, the aims of this research are:

1. Identifying cities in Indonesia that are in the unhealthy air pollution index category.
2. Estimating Generalized Pareto Distribution parameters to estimate the frequency and severity of unhealthy air pollution in provincial capitals in Indonesia?

2 Method

2.1 Research Methodology

This research uses non-experimental quantitative methods with a special descriptive design. Quantitative research is a research procedure that produces data in the form of numbers and is generally analyzed using descriptive or inferential statistics [11]. This means that the numbers obtained are processed and sought to find out their effect on the formulation of research problems that have been determined. Specific description aims to describe a phenomenon with clear calculation analysis. In this study, the calculation of air pollution models for cities in Indonesia will be carried out. The research framework is in Figure 2.

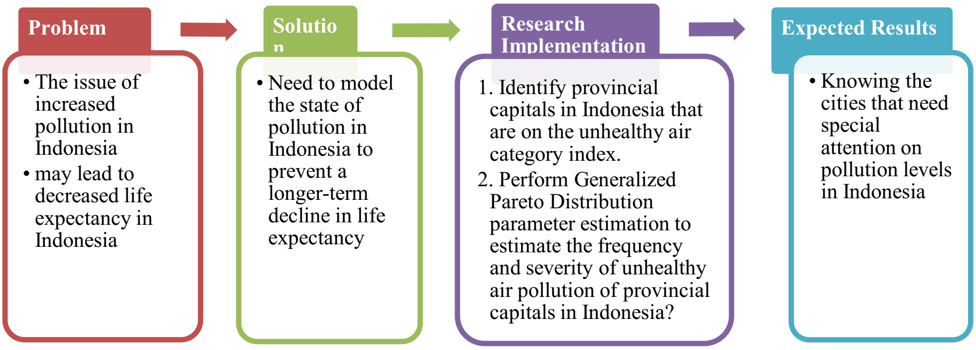

Fig.2 Research Framework

2.2 Research Data Sources

The data used in the study are air pollution index data in all provincial capitals in Indonesia, namely Mamuju, Palu, Tanjung Selor, Ambon, Sofifi, Kupang, Banda Aceh, Tanjung Pinang, Makassar, Sorong, Kendari, Manado, Denpasar, Manokwari, Pontianak, Gorontalo, Pekanbaru, Padang, Mataram, Bengkulu, Samarinda, Pangkal Pinang, Yogyakarta, Palangkaraya, Bandar Lampung, Jambi, Semarang, Jayapura, Banjarbaru, Medan, Surabaya, Bandung, Palembang, Jakarta and Serang. The air pollution index was observed in the past year starting in September 2023 until August 2024 every day. The observed data is the median Air Quality Index data which can be accessed at https://aqicn.org/ [12].

2.3 Data Analysis

This research uses a quantitative method by mathematically modeling the Provincial Capital Air Quality Index in Indonesia. The data analysis steps are:

1. Identify capitals that have an unhealthy API by looking at the distribution on the scatterplot of the air quality index distribution on 1 year data. The limits of the capital city that has an unhealthy API are those that exceed the 100 threshold identified based on Table 1.

Table 1. Level of Quality Index [12]

Quality Index	Air Pollution Level	Health Implications	Cautionary Statement (for PM2.5)
0 - 50	Good	Air quality is considered satisfactory, and air pollution poses little or no risk	None
51 -100	Moderate	Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution.	Active children and adults, and people with respiratory disease, such as asthma, should limit prolonged outdoor exertion.
101-150	Unhealthy for Sensitive Groups	Members of sensitive groups may experience health effects. The general public is not likely to be affected.	Active children and adults, and people with respiratory disease, such as asthma, should limit prolonged outdoor exertion.
151-200	Unhealthy	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects	Active children and adults, and people with respiratory disease, such as asthma, should avoid prolonged outdoor exertion; everyone else, especially children, should limit prolonged outdoor exertion

Quality Index	Air Pollution Level	Health Implications	Cautionary Statement (for PM2.5)
201-300	Very Unhealthy	Health warnings of emergency conditions. The entire population is more likely to be affected.	Active children and adults, and people with respiratory disease, such as asthma, should avoid all outdoor exertion; everyone else, especially children, should limit outdoor exertion.
300+	Hazardous	Health alert: everyone may experience more serious health effects	Everyone should avoid all outdoor exertion

2. Ensure heavytail testing in each province that has a distribution of unhealthy air conditions. This test looks at the kurtosis, skewness and histogram of the data. Data that have heavytail are those that have kurtosis greater than 3 and have skewness values not equal to zero [13]. This heavytail distribution test uses Rstudio. On the histogram, a heavytail distribution will be seen to form a curve that drops or rises slowly and forms a chart that tends to the right or left [14].
3. Identifying extreme values in unhealthy air quality index data. The threshold value in this study is u=100. To find out this, the percentage method can be used, which is carried out:
   a. Sort the air pollution index from the smallest to the largest.
   b. Count the number of extreme data above the threshold with:
   $$n = \frac{k}{N}$$
   Where k is extreme data above threshold, N the number of all data.
4. Estimating the parameter Generalized Pareto Distribution in extreme data that have unhealthy air pollution index. Parameter on Generalized Pareto Distribution is ξ. The probability density function is found in equation 2 [15].
   $$f_\xi(x) = \begin{cases} \left(\frac{1}{\sigma} - (1 + \xi\frac{x}{\sigma}\right)^{-\frac{\xi+1}{\xi}} & untuk\ \xi \neq 0 \\ \frac{1}{\sigma} - e^{-\frac{x}{\sigma}} & untuk\ \xi = 0 \end{cases} \quad (1)$$
   Where x is a random variable, σ is the extreme value scale parameter of the air quality index data and ξ is the extreme value tail parameter of the air quality index. In this study, we used the maximum likelihood method to estimate the model parameters assisted by RStudio application.
5. Interpret the parameter results of each provincial capital indicated to have unhealthy air quality.

3 Results and Discussion

3.1 Identification of Unhealthy Provincial Capital

There is one city that is in the very unhealthy group with an air pollution index distribution of 201-300, namely Palembang. The distribution of the air quality index can be seen in Figure 3. Then Palembang city will be checked for kurtosis, skewness and histogram.

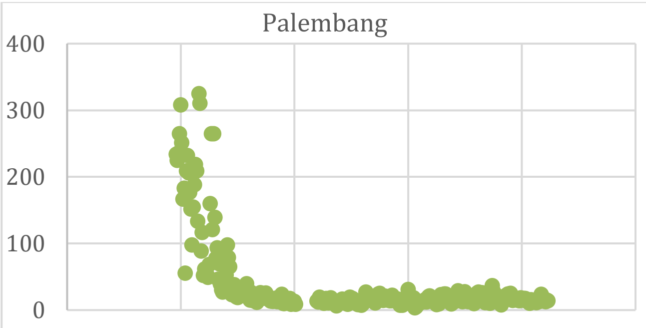

Fig.3 Distribution of Palembang Air Quality Index

There are 3 cities categorized in the unhealthy group for several groups, namely Jambi, Palangkaraya and Banjarbaru. The distribution of the air quality index is shown in Figure 4. Then these three cities will be checked for kurtosis, skewness and histograms.

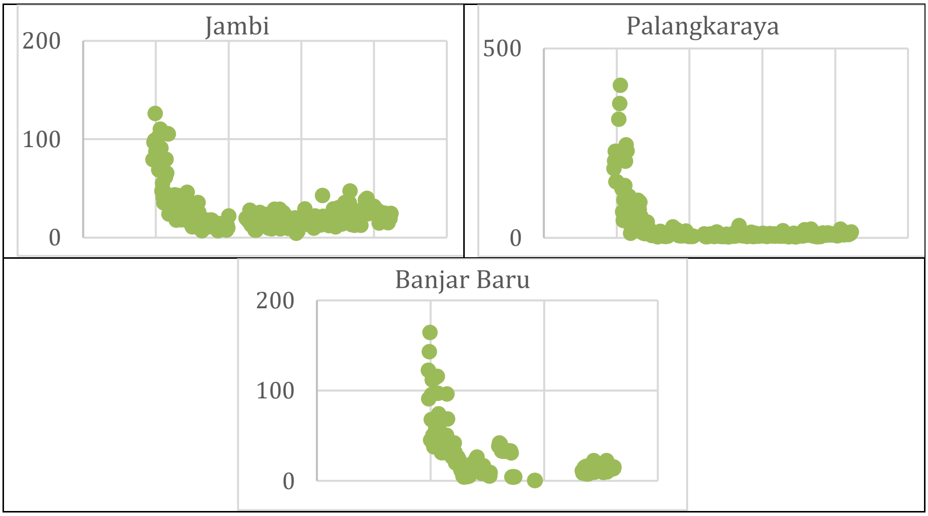

Fig.4 Distribution of Jambi, Palangkaraya and Banjarbaru Air Quality Index

There are 13 cities categorized in the moderate category, namely Medan, Padang, Serang, Jakarta, Bandung, Sorong, Bengkulu, Tanjung Pinang, Lampung, Pontianak, Samarinda, Yogyakarta and Surabaya. The distribution of several cities of moderate air quality index can be seen in Figure 5. All cities have an air index below 100, but some time in DKI Jakarta and Sorong passes above 100, so in these two cities will be checked kurtosis, skewness and histogram.

Fig.5 Distribution of Provincial Capital in Moderate Category

There are 18 provincial capitals in Indonesia that are in the good category, namely Banda Aceh, Bengkulu, Pekan Baru, Pangkal Pinang, Tanjung Selor, Semarang, Denpasar, Kupang, Mataram, Gorontalo, Mamuju, Palu, Manado, Kendari, Makassar, Sofifi, Ambon, Manokwari and Jayapura. All of these cities are at a good air quality index and are always below 50. Therefore these cities are not subjected to further testing.

3.2 Kurtosis, Skewness and Histogram Unhealthy Provincial Capitals

After seeing the capitals that have an Air Quality Index value above 100, then checking the kurtosis, skewness and histogram values of provincial capitals as assumptions that must be met in the Generalized Pareto Distributon (GDP) distribution. We will check 6 capitals that are indicated to be unhealthy, namely Palembang, Jambi, Palangkaraya, Banjar Baru, Jakarta and Sorong. The Kutosis, Skewness values are listed in Table 2.

Table 2 Kurtosis and Skewness 6 Capitals with Unhealthy Air Index

Provincial Capital	**Kurtosis**	**Skewness**	Conclusion
Palembang	13.06132	3.235442	Heavytail
Jambi	13.73005	3.130353	Heavytail
Palangkaraya	24.95821	4.482096	Heavytail
Banjar Baru	11.46368	2.616551	Heavytail
Jakarta	3.551022	0.613896	Heavytail
Sorong	9.437644	1.938354	Heavytail

Based on Table 2, all the unhealthy air index data are heavytailed as they have kurtosis values above 3 and skewness values that are not equal to zero. However, it should be noted that the skewness of the city of Jakarta is not too large because the kurtosis value is close to 3 and the skewness is close to 0. To further confirm, a histogram of the data of the 6 capitals is determined and all are seen to form a tail to the left. The results of the histograms of the 6 capitals are in Figure 6.

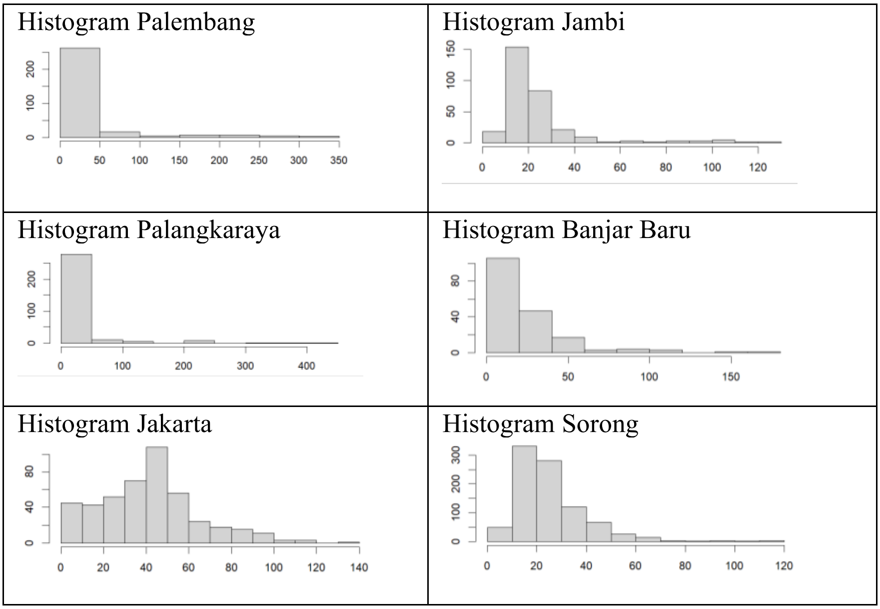

Fig.6 Histogram of six provincial capitals with unhealthy air quality indexes

Based on Figure 6, the histograms of Palembang, Jambi, Palangkaraya, Banjar Baru and Sorong form a heavytail so that the assumptions on GDP are met, although for the city of Jakarta the slope value still tends to be tilted to the left although it is still quite mild. However, over all the assumptions are met.

3.3 Parameter Estimation dan Interpretation

Generalized Pareto Distribution is the distribution used to model the tail of the distribution in the air quality index data of six provincial capitals during extreme events. The results are estimated using Maximum Likelihood Estimation so as to obtain the shape parameter (ξ) and scale parameter (α) of the GDP distribution by substituting as many observed values of air quality of six provincial capitals that are indicated to have unhealthy air quality by limiting the threshold value u = 100. The estimated shape parameter (ξ) and scale parameter (α) are listed in Table 3.

The ξ parameter signifies the severity level that indicates the potential for worsening air quality events to occur. Based on Table 3, in provincial capitals in Indonesia that are indicated to have poor air quality, the ξ parameter in the capitals of Palembang, Jambi, Palangkaraya, Jakarta and Sorong is negative. This parameter indicates the severity of the extreme value of unhealthy air quality is not too large. Meanwhile, in Banjar Baru city, the value of $\xi = 0.026443$. Although positive, the value of ξ tends to be small, indicating that the potential for unhealthy air that may occur is still small. This means that air quality in Indonesia based on 2023-2024 data can still be controlled and there is no need to worry that there will be extreme events of air quality increase in the near future.

In addition, the scale parameter α estimates the frequency and severity of extreme events in the data distribution, in this case the occurrence of unhealthy air quality in each city, Based on the value of α, the city that needs to be a concern because it will allow the occurrence of unhealthy air quality levels more often is Palembang and Palangkaraya, which touches the value of α in Palembang is 203.464 and in Palangkaraya is 171.3727. In addition, the other four cities still have low values.

Tabel 3 Parameter GDP Estimation

Capital Province	ξ	α
Palembang	-0.88917	203.464
Jambi	-0.04836	9.027511
Palangkaraya	-0.49762	171.3727
Banjar Baru	0.026443	22.32232
Jakarta	-0.21834	13.97766
Sorong	-0.13425	20.48764

4 Conclusion

Air quality in various cities in Indonesia shows a significant level of variability. Cities that have had an unhealthy air quality index with air quality indicators above 100 are Palembang, Jambi, Palangkaraya and Banjarbaru. Therefore, to model unhealthy air quality based on the API index in these cities can use the Generalized Pareto Model method. After modeling, the cities of Palembang and Palangkaraya are the cities that need attention because they allow the occurrence of unhealthy air quality levels more frequently. While other cities are in the Moderate category namely Medan, Padang, Serang, Jakarta, Bandung, Sorong, Bengkulu, Tanjung Pinang, Lampung, Pontianak, Samarinda, Yogyakarta and Surabaya as well as the Good category namely Banda Aceh, Bengkulu, Pekan Baru, Tanjung Pinang, Lampung, Pangkal Pinang, Pontianak, Samarinda, Tanjung Selor, Semarang, Yogyakarta, Surabaya, Denpasar, Kupang, Mataram, Gorontalo, Mamuju, Palu, Manado, Kendari, Makassar, Sofifi, Ambon, Manokwari, Jayapura have scatterplots spread evenly so that GDP estimation and parameter search in this city are not suitable.

5 Suggestion

This study focuses on analyzing the air pollution index only in the capital city, which provides useful insights but also has limitations. To get a more comprehensive picture of air pollution conditions in Indonesia, it is recommended to conduct research that covers various cities across the country, not just the capital city.

References

1. P. K. M. Soesilo and F. Rahman, "The Pillars of Survival in the COVID-19 Pandemic: The Case of Indonesia BT - Community, Economy and COVID-19: Lessons from Multi-Country Analyses of a Global Pandemic," I. I. Shultz Clifford J., D. R. Rahtz, and M. J. Sirgy, Eds. Cham: Springer International Publishing, 2022, pp. 267–289.
2. C. Huang, X. Zhao, W. Cheng, Q. Ji, Q. Duan, and Y. Han, "Statistical Inference of Dynamic Conditional Generalized Pareto Distribution with Weather and Air Quality Factors," Mathematics, vol. 10, no. 9, pp. 1–25, 2022, doi: 10.3390/math10091433.
3. A. Budiyono, "Indeks Kualitas Udara," Pencemaran Udar. Dampak Pencemaran Udar. pada Lingkung., vol. 02, no. 01, pp. 1–8, 2001.
4. N. Limayani and E. Tanur, "Happiness and air quality: microdata analysis in Indonesia," J. Heal. Popul. Nutr., vol. 43, no. 1, p. 25, 2024, doi: 10.1186/s41043-024-00517-3.
5. M. Greenstone and Q. (Claire) Fan, "Kualitas udara Indonesia yang memburuk dan dampaknya terhadap harapan hidup," Air Qual. Life Index, pp. 1–10, 2019, [Online]. Available: https://aqli.epic.uchicago.edu/wp-content/uploads/2019/03/Indonesia.Indonesian.pdf.

6. N. G. Safana, Y. Astor, and F. Yulistiani, " Correlation Model of Horizontal Geometric and Road Grade on CO 2 and PM 2.5 Vehicle Emission In Bandung City ," no. January, 2020, doi: 10.2991/aer.k.201221.041.
7. W. C. Leong, R. O. Kelani, and Z. Ahmad, "Prediction of air pollution index (API) using support vector machine (SVM)," J. Environ. Chem. Eng., vol. 8, no. 3, p. 103208, 2020, doi: 10.1016/j.jece.2019.103208.
8. N. S. M. Fauzi, M. M. Kasim, and N. H. M. Desa, "Construction of Air Pollution Index with the Inclusion of Aggregated Weights of the Pollutants," J. Inf. Commun. Technol., vol. 21, no. 4, pp. 495–512, 2022, doi: 10.32890/jict2022.21.4.2.
9. N. Masseran and M. A. Mohd Safari, "Intensity–duration–frequency approach for risk assessment of air pollution events," J. Environ. Manage., vol. 264, no. March, p. 110429, 2020, doi: 10.1016/j.jenvman.2020.110429.
10. N. A. AL-Dhurafi, N. Masseran, and Z. H. Zamzuri, "Hierarchical-Generalized Pareto Model for Estimation of Unhealthy Air Pollution Index," Environ. Model. Assess., vol. 25, no. 4, pp. 555–564, 2020, doi: 10.1007/s10666-020-09696-9.
11. N. Nasir and S. Sukmawati, "Analysis of Research Data Quantitative and Qualitative," Edumaspul J. Pendidik., vol. 7, no. 1, pp. 368–373, 2023.
12. The World Air Quality Index Project, "Air Quality and Pollution Measurement," 2024. https://aqicn.org/.
13. R. M. Vogel, S. M. Papalexiou, J. R. Lamontagne, and F. Dolan, When Heavy Tails Disrupt Statistical Inference. 2024.
14. J. Vecchio and D. Raggi, "Dependence in Extreme Events An Analysis Based on Copula Functions," Ca Foscari University of Venice, 2023.
15. K. Klinjan, T. Sottiwan, and S. Aryuyuen, "Extreme Value Analysis with New Generalized Extreme Value Distributions: A Case Study for Risk Analysis on PM2.5 And PM10 In Pathum Thani, Thailand," Commun. Math. Biol. Neurosci., pp. 1–22, 2024.