Investigation on the Statistical Distribution of $PM_{2.5}$ Concentration in Chiang Mai, Thailand

SUKANYA INTARAPAK[1], THIDAPORN SUPAPAKORN[2,*]
[1]Department of Mathematics, Faculty of Science,
Srinakharinwirot University, Bangkok 10110, THAILAND
[2]Department of Statistics, Faculty of Science,
Kasetsart University, Bangkok 10900, THAILAND

Abstract: - Recently, it is found that Northern Thailand has very high levels of airborne particulates known as $PM_{2.5}$. $PM_{2.5}$ particulates can cause breathing problems and may raise the risks of heart disease and even some cancers. According to AirVisual, Chiang Mai, the capital of Northern Thailand which offers for tourists in both business and cultural center, had the highest levels of smog in the world in March 2018, reaching at least 183 on the $PM_{2.5}$ Air Quality Index scale. The daily average $PM_{2.5}$ concentration data are determined from July 2016 – June 2018 at two stations in Chiang Mai at Yupparaj Wittayalai school and City Hall. The Weibull, Gamma, Lognormal and Inverse Gaussian distributions are considered for finding the most appropriate probability functions of the daily average $PM_{2.5}$ concentration. The results show that, as evaluated with the goodness- of-fit measures; Komolgorov-Smirnov and Anderson-Darling test statistics, the Inverse Gaussian distribution is the most suitable probability density functions of the daily average $PM_{2.5}$ concentration for two stations. Furthermore, the return periods of the $PM_{2.5}$ concentration are predicted by using the Largest Extreme Value distribution, which can be further applied in air quality management and related policy making.

Key-Words: distribution, goodness-of-fit, Inverse Gaussian, Largest Extreme Value, $PM_{2.5}$, return period

Received: May 12, 2021. Revised: October 20, 2021. Accepted: November 17, 2021. Published: December 3, 2021.

1 Introduction

In 1982, Thailand introduced its first air quality standard, but the focus was on particles wider than 10 microns. In 2004, the standard was added for particles 10 microns or smaller. Six years later a Particulate Matter 2.5 ($PM_{2.5}$) microns standard was adopted in light of strong evidence that the tiniest particles could penetrate the bloodstream via the lungs and trigger more critical medical conditions [1]. Many researchers indicate that the health effects of $PM_{2.5}$ are more harmful than PM_{10} [2-4].

According to the 2018 World Air Quality Report, Thailand has been ranked the world's 23rd most polluted country. The report shows that $PM_{2.5}$ dust particles for the whole year in Thailand averaged 26.4 microns. The $PM_{2.5}$ problem in Thailand has many similarities to that in many other cities of the world, but it also has some particularities for the region. Also, as reported by the Air Visual app, which monitors air quality around the world, Chiang Mai ranked sixth worst in the world in terms of air pollution, according by US Air Quality Index. It ranked just Nepal's Kathmandu, India's New Delhi, Pakistan's Lahore, Bangladesh's Dhaka and India's Kolkata. For safety reasons, the amount of $PM_{2.5}$ should not exceed 50 $\mu g/m^3$ of air, according to Thai health authorities. In Chiang Mai's Muang district, air quality tests revealed the amount of $PM_{2.5}$ ranged between 86 $\mu g/m^3$ and 91 $\mu g/m^3$ in March 2018 that made Chiang Mai the worst air pollution in the world [5].

Chiang Mai is the tourist-friendly cultural center in northern Thailand and not an industrial powerhouse populated by millions. In March 2019, forest fires have made the air worse than Beijing's [6]. The air pollution is caused in part by forest fires, notably the practice of the area's farmers of starting fires to clear land for new harvests. And aside from the health hazards to humans and animals alike, there is another smog side effect. Haze has been a seasonal problem in Chiang Mai for over a decade. It usually appears from January to April. Some of the haze has been attributed to neighbouring countries like Laos and Myanmar. Located in the northern region of Thailand, Chiang Mai is especially vulnerable to air pollution. Above all as it is surrounded by mountain ranges that trap pollution [5-8].

Moreover, the air pollution in Chiang Mai is caused by lots of constructions and vehicles. There are open burning in forest areas, community areas and agricultural areas. The landscape of northern Thailand is mountain and plain, in which has lots of living people. The surrounding mountains are barriers to ventilation, which cause the occurrence of $PM_{2.5}$. The $PM_{2.5}$ has affected health in both short and long term. The levels of $PM_{2.5}$ in Chiang Mai is the highest in the world from 11-14 March 2019, according to Air Visual (https://www.iqair.com/th/), referring to the ranking of air quality of the world.

The probability distribution can be used to predict the number of days when the ambient air quality standard (AQS) is exceeded. It is necessary to use an appropriate type of statistical distribution to compute the exceeding probabilities and percentiles for setting regulatory targets and issuing environmental alerts for public health [9]. In 2006, Gavril et al. [10] studied the concentration distribution of PM in Athens, Greece. Eight probability distribution functions were fitted to measure concentration of PM_{10} and $PM_{2.5}$ in order to determine the shape of the concentration distribution. The best-fit probability density functions were selected based on the combined results of goodness of fit statistics including Kolmogorov-Smirnov, Anderson-Darling and Chi-Square tests. The results indicated that the Pearson type VI probability density function provided a better fit to the measured data. Other functions exhibiting high accuracy of fit were the inverse Gaussian, the lognormal and Pearson type V.

In 2013, Xi et al. [9] described the statistical distribution characteristics of daily average PM_{10} concentration in Beijing, Guangzhou, Shanghai, Wuhan, and Xi'an. The daily PM_{10} average concentration in the 5 cities was measured from 1 January 2004 to 31 December 2008. The PM_{10} concentration distribution was simulated by using the lognormal, Weibull and Gamma distributions and the best statistical distribution of PM_{10} concentration in the 5 cities was detected using to the maximum likelihood method. The results showed that the best fit distribution for daily PM_{10} concentration in the 5 cities of China was the lognormal distribution. In the same year, Hamid et al. [11] compared the performance of parameter estimator for two-parameter and three-parameter lognormal distribution by using PM_{10} concentration in Nilai, Negeri Sembilan, Malaysia. Two methods were used to estimate the parameters which were the method of moments and the method of probability weighted moments. Five performance indicators were used to determine the best estimator and the best distribution to represent the PM_{10} concentration in Nilai, Negeri Sembilan from 2003 to 2009. The results showed that three-parameter lognormal distribution performed better compared to two-parameter lognormal distribution.

Therefore, the researchers are interested in the probability distribution of daily average concentration of $PM_{2.5}$ in Chiang Mai. The data are collected from July 2016 to June 2018. The purposes of this study are to find out: 1) the most appropriate probability distribution of daily average concentration of $PM_{2.5}$ 2) the probability of daily average concentration of $PM_{2.5}$ exceeding air quality standard and 3) the return periods, period for daily average concentration of $PM_{2.5}$ exceeding air quality standard.

2 Materials and Methods

2.1 Data

In this research, the daily average concentration of $PM_{2.5}$ are collected at Yupparaj Wittayalai school station and City Hall station in Chiang Mai's Muang district from 1st July 2016 to 30th June 2018 (https://www.iqair.com/th/).

2.2 Distributions

The daily average concentration of $PM_{2.5}$ at Yupparaj Wittayalai school station and City Hall station are used to find the appropriate distribution. Four distributions of daily average concentration of $PM_{2.5}$: Weibull, Gamma, Lognormal, and Inverse Gaussian distributions are examined and the parameters are estimated by the maximum likelihood method. The probability density function and cumulative distribution function of Weibull, Gamma, Lognormal, and Inverse Gaussian distributions are represented in Table 1 [12].

The maximum likelihood estimation is the method that determines values for the parameters of the model. The parameter values are found such that they maximize the likelihood that the process described by the model produced the data that were actually observed. Suppose that a random sample is $X_1, X_2, \ldots, X_n$ for which the probability density (or mass) function of each X_i is $f(x_i;\theta)$. Then, the joint probability density (or mass) function of $X_1, X_2, \ldots, X_n$, is

$$L(\theta)=P(X_1=x_1, X_2=x_2, \ldots, X_n=x_n) = f(x_1;\theta)\cdot f(x_2;\theta)\cdots f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta) \quad (1)$$

The first equality is of course just the definition of the joint probability mass function. The second equality comes from the fact that we have a random sample, which implies by definition that the X_i are independent. And, the last equality just uses the shorthand mathematical notation of a product of indexed terms. Now, in light of the basic idea of maximum likelihood estimation, one reasonable way to proceed is to treat the "likelihood function" $L(\theta)$ as a function of θ, and find the value of θ that maximizes it [13].

2.3 Goodness-of-fit Tests

To investigate the performance of the probability functions, the goodness-of-fit tests indicate whether or not it is reasonable to assume that a random sample comes from a specific distribution. The null and alternative hypotheses of Goodness-of-fit tests are that the sample data come from the stated distribution and sample data do not come from the stated distribution, respectively. In this research, Kolmogorov-Smirnov and Anderson-Darling tests are used for comparing distributions to find the appropriate distributions.

The first measure of goodness of fit for general distributions was derived by Kolmogorov (1933). The Kolmogorov statistic or the Kolmogorov-Smirnov (*KS*) test statistic is

$$KS = \sup_x |F^*(x) - S(x)| \qquad (2)$$

where $F^*(x)$ be some completely specified distribution function, the hypothesized distribution function, $S(x)$ be the empirical distribution function (EDF) based on the random sample X_1, X_2, ..., X_n and "sup" is supremum [14].

Anderson and Darling (1954) looked to improve upon the Kolmogorov-Smirnov statistic by modifying it for distributions of interest. The Anderson-Darling test is used to verify if a sample of data came from a population with a specific distribution. It is a modification of the *KS* test that accounts for the distribution and test and gives more attention to the tails. As mentioned before, the *KS* test is distribution free, in the sense that the critical values do not depend on the specific distribution being tested. The Anderson-Darling test makes use of the specific distribution in calculating the critical values. The advantage is that this sharpens the test, but the disadvantage is that critical values must be calculated for each hypothesized distribution.

The Anderson–Darling statistic (A_n^2) is

$$A_n^2 = -\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln \hat{F}(X_{(i)}) + \ln\left(1 - \hat{F}(X_{(n-i+1)})\right)\right] - n \qquad (3)$$

where $\hat{F}(x)$ is the cumulative distribution function associated with the null hypothesis and n is sample size [14].

2.4 Return Period

The general probability distribution is suitable for the daily average concentration of $PM_{2.5}$, but not suitable for high concentration data. To reduce error for predicting, the Largest Extreme Value distribution is used to predict return periods for exceeding the air quality standard of the daily average $PM_{2.5}$ concentration.

In this study, if the daily average concentration of $PM_{2.5}$ at Yupparaj Wittayalai school station and City Hall station are greater than 50 μg/m^3 according to Thailand's safe standard [7], the parameters of the Largest Extreme Value distribution are estimated by the maximum likelihood method.

The probability density function and the cumulative distribution function of the Largest Extreme Value distribution are shown as, respectively,

$$f(x;\mu,\sigma) = \frac{1}{\sigma}exp\left[-\left(\frac{x-\mu}{\sigma}\right) - exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\}\right]; -\infty < x < \infty \qquad (4)$$

and

$$F(x;\mu,\sigma) = exp\left[-exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\}\right] \qquad (5)$$

where μ is a location parameter, $\mu > 0$and σ is a scale parameter, $\sigma > 0$ [12].

In the extreme value analysis, the return period is defined as the average length of time between events of the same magnitude or greater. From the Largest Extreme Value distribution, the return value is defined as a value that is expected to be equal or exceeded on average once every interval of time. Therefore, the return period $(R(x_l))$ is obtained by the following equation,

$$R(x_l) = \frac{1}{1 - F(x;\mu,\sigma)} \qquad (6)$$

where $F(x;\mu,\sigma)$ is the cumulative distribution function of the Largest Extreme Value distribution and x_l is the air quality standard of this research that equals 50 μg/m^3 [15].

3 Results

The descriptive statistics of daily average $PM_{2.5}$ concentrations at Yupparaj Wittayalai school station and City Hall station from 1st July 2016 to 30th June 2018 are shown in Table 2. The daily average concentration of $PM_{2.5}$ exceeded air quality standard at Yupparaj Wittayalai school station and City Hall station are 11.71% and 16.89% of the data, respectively. In one year, there are about 43 days and 62 days of the daily average concentration of $PM_{2.5}$ exceeded air quality standard at Yupparaj Wittayalai school station and City Hall station, respectively. However, the mean of daily average $PM_{2.5}$ concentration from Yupparaj Wittayalai school station and City Hall station are not exceeded air quality standard. The daily average $PM_{2.5}$ concentrations of two stations are the positively skewed distribution.

The parameters of the distributions of daily average concentration of $PM_{2.5}$ are estimated by using the maximum likelihood method. The estimated parameters for four distributions; Weibull, Gamma, Lognormal and Inverse Gaussian distributions, of daily average concentration of $PM_{2.5}$, are presented in Table 3 and the histograms for four distributions of daily average concentration of $PM_{2.5}$ are performed in Fig. 1. for both two stations in Chiang Mai.

The Kolmogorov-Smirnov and Anderson-Darling values from the following equations (2) and (3), respectively for four distributions of daily average concentration of $PM_{2.5}$ are presented in Table 4. The smaller values indicated a better fit with the actual data. Considering the Kolmogorov-Smirnov and Anderson-Darling values, the Inverse Gaussian distribution of daily average $PM_{2.5}$ concentration is most suitable for City Hall station. At Yupparaj Wittayalai school station, the Inverse Gaussian distribution of daily average $PM_{2.5}$ concentration is most suitable by considering the Anderson-Darling values. Whereas the Kolmogorov-Smirnov values are considered, the Lognormal distribution of daily average $PM_{2.5}$ concentration is the best fit, but the Kolmogorov-Smirnov values of Lognormal and Inverse Gaussian distributions are similar. Therefore, the Inverse Gaussian is carried out for Yupparai Wittayalai school and City Hall stations.

The estimated parameters and the return periods obtained using equation (6) from the two stations of the Largest Extreme Value distribution estimated from the following equations (4) and (5) are shown in Table 5. According to this research, the return period for Yupparaj Wittayalai school station is 20.97 days, whenever $PM_{2.5}$ concentration exceed the air quality standard. Thus $PM_{2.5}$ concentration are more likely to exceed the air quality standard again for about 21 days. Similarly, the return period for City Hall station is 22.61 days, whenever $PM_{2.5}$ concentration exceed the air quality standard, also $PM_{2.5}$ concentration are more likely to exceed the air quality standard again for about 23 days.

4 Discussion

The main pollutant in Chiang Mai area are fine particulates grouped as $PM_{2.5}$. The air pollution is combined with other minor pollutants but $PM_{2.5}$ is the one considered to provide the Air Quality Index (AQI). The toxicity of particulate matter regarding short- and long-term exposure is quite established as increasing morbidity and mortality. PM is detrimental to the respiratory and cardiovascular systems and high levels are classically associated with an increase incidence of stroke and myocardial infarctions, increase in ER admission for asthma, chronic obstructive pulmonary disease (COPD), respiratory infection and increase incidence of lung cancer [16-17].

Based on the daily average concentration of $PM_{2.5}$ during July 2016 to June 2019 and Thailand's safe standard (50 $\mu g/m^3$), around twice the threshold levels recommended by the World Health Organization (WHO), the daily average concentration of $PM_{2.5}$ exceeded air quality standard are at least 1-2 months a year during the dry season (January – April) in Chiangmai, Thailand as it can be seen clearly with red colour in Fig. 2. The air pollution in Chiang Mai is usually caused by regional forest fires and farmers burning waste to clear land for the next harvest season [6-8]. While the World Health Organization's recommended level is at 25 micrograms per cubic meter [7], the daily average concentration of $PM_{2.5}$ exceeded air quality standard are about 33% and 49% for Yupparaj Wittayalai school and City Hall stations, respectively, as the yellow and red colours in Fig. 2. Thus, the daily average concentration of $PM_{2.5}$ in Chiang Mai exceeded WHO's safe standard (25 $\mu g/m^3$) are about 4-6 months a year.

5 Conclusion

A general conclusion is that the distribution of daily average $PM_{2.5}$ concentration are approximately right skewed continuous probability density functions for Yupparai Wittayalai school and City Hall stations in Chiang Mai. Four different distributions: Weibull, Gamma, Lognormal, and Inverse Gaussian distributions are fitted of the actual data and the parameters of each distribution are estimated by using the maximum likelihood method. The appropriate distributions of daily average $PM_{2.5}$ concentration for Yupparai Wittayalai school and City Hall stations are the Inverse Gaussian distribution considered by the goodness-of-fit tests; the Kolmogorov-Smirnov and the Anderson-Darling tests. From the Largest Extreme Value distribution, the $PM_{2.5}$ concentration for Yupparaj Wittayalai school and City Hall stations are more likely to exceed the air quality standard again for about 21 and 23 days, respectively. Hence, the $PM_{2.5}$ concentration occur recurrent year over year. These estimates can help policy makers to create initiatives to solving health and environment problems. However, this research uses the air quality standard that equals 50 $\mu g/m^3$ based on Thailand's safe standard. In the future, the research should be considered to improve the air quality standard that equals 25 $\mu g/m^3$ according to the World Health Organization's level. In addition, it may also expand the scope of the research in Bangkok, where is the capital city of Thailand in the aspect of $PM_{2.5}$ dust particles as well.

Acknowledgements:
This work was supported by Strategic Wisdom and Research Institute and Faculty of Science, Srinakharinwirot University, Thailand.

References:

[1] Wongruang, P. (2019). REPORT: Diesel combustion drives this $PM_{2.5}$ mess, said PCD latest study. Retrieved August 15, 2019, from https://www.nationthailand.com/national/30362542

[2] Chen, C. C., Tsai, S. S., & Yang, C. Y. (2015). Association between Fine Particulate Air Pollution and Daily Clinic Visits for Migraine in a Subtropical City: Taipei, Taiwan. *International journal of environmental research and public health*, 12(5), pp. 4697-4708.

[3] Kloog, I., Ridgway, B., Koutrakis, P., Coull, B. A., & Schwartz, J. D. (2013). Long- and Short-Term Exposure to $PM_{2.5}$. and Mortality: Using Novel Exposure Models. *Epidemiology*, 24(4), pp. 555-561.

[4] Xing, Y. F., Xu, Y.H., Shi, M. H., & Lian, Y. X. (2016). The impact of $PM_{2.5}$ on the human respiratory system. *Journal of Thoracic Disease*, 8(1), pp. E69-E74.

[5] The Nation. (2018). Chiang Mai ranked sixth worst for air pollution globally. Retrieved June 3, 2020, from https://www.nationthailand.com/around_thailand/30341186

[6] Tamkin, E. (2019). Northern Thailand was once a paradise. Now forest fires have made the air worse than Beijing's. Retrieved June 3, 2020, from https://www.washingtonpost.com/world/2019/05/07/northern-thailand-was-once-paradise-now-forest-fires-have-made-air-worse-than-beijings/

[7] Wipatayotin, A. (2019). Chiang Mai air pollution worst in the world. Retrieved August 15, 2019, from https://www.bangkokpost.com/thailand/general/1643388.

[8] CTN News. (2020). Northern Thailand's Chiang Mai Once Again Has World's Worst Air-Quality. Retrieved June 3, 2020, from https://www.chiangraitimes.com/thailand-national-news/northen-thailand/chiang-mai-air-quality/

[9] Xi, W., Jie, C.R., Heng, C. B., & Dong, K. H. (2013). Application of Statistical Distribution of PM_{10} Concentration in Air Quality Management in 5 Representative Cities of China. *Biomed Environ Sci*, 26(8), pp. 638-646.

[10] Gavrill, I., Grivas, G., Kassomenos, P., Chaloulakou, A. & Spyrellis, N. (2006). An Application of Theoretical Probability Distributions, to the Study of PM_{10} and $PM_{2.5}$ Time Series in Athens. Greece. *Global NEST Journal*, 8(3), pp. 241-251.

[11] Hamid H. A., Yahaya A. S., Ramli N. A., & Ul-Saufie A. Z. (2013). Finding the Best Statistical Distribution Model in PM_{10} Concentration Modeling by using Lognormal Distribution. *Journal of Applied Sciences*, 13(2), pp. 294-300.

[12] Walck, C. (2007). *Hand-book on statistical distributions for experimentalists*. International report No.SUF–PFY/96–01, University of Stockholm, Stockholm, Sweden.

[13] Strickland, J. (2017). *Logistic Regression Inside and Out*. Lulu Inc., the United States of America.

[14] Kvam, P. H. & Vidakovic, B. (2007). *Nonparametric Statistics With Applications to Science and Engineering*. Wiley-Interscience, New York.

[15] Bedient P.B., Huber W.C. & Vieux B.E. (2013). *Hydrology and floodplain analysis*. Upper Saddle River, NJ: Prentice Hall. New Jersey.

[16] Aungkulanon, S., Tangcharoensathien, V, Shibuya, K., Bundhamcharoen, K. & Chongsuvivatwong, V. (2016). Post universal health coverage trends and geographical inequalities of mortality in Thailand. *International Journal for equity in health*, 15(1):190, pp. 1-12. doi: 10.1186/s12939-016-0479-5.

[17] Cassian, P. (2019). Some facts & tips about Chiang Mai burning season. Retrieved June 3, 2020, from https://www.chiangmaidoctor.com/burning-season

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Sukanya Intarapak and Thidaporn Supapakorn investigated the statistical distribution of $PM_{2.5}$ concentration. Thidaporn Supapakorn set up the scope, analysed the data and then carried out the implementation. Sukanya Intarapak wrote the manuscript with input from Thidaporn Supapakorn who conceived the study and was in charge of overall direction and planning. All authors discussed the results and contributed to the final manuscript.

Appendix

Table 1. Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of Weibull, Gamma, Lognormal, and Inverse Gaussian Distributions0652401564

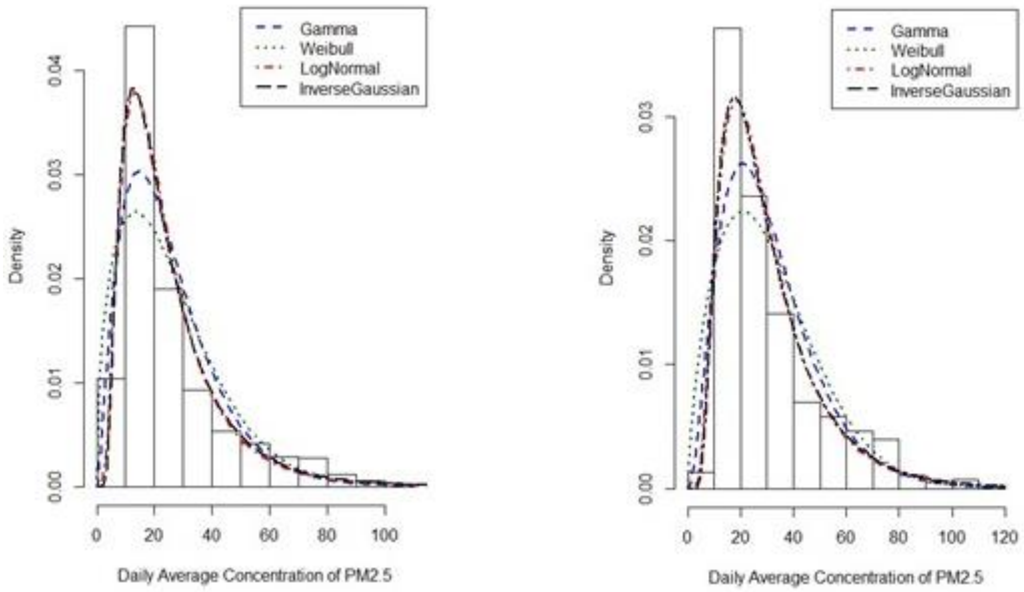
Distribution	PDF	CDF	Parameter
Weibull	$f(x)=\frac{\lambda}{\beta}\left(\frac{x}{\beta}\right)^{\lambda-1}e^{-\left(\frac{x}{\beta}\right)^{\lambda}};x>0$	$F(x)=1-e^{-\left(\frac{x}{\beta}\right)^{\lambda}}$	β is a location parameter, $\beta>0$. λ is a shape parameter, $\lambda>0$
Gamma	$f(x)=\frac{1}{\Gamma(\lambda)\beta^{\lambda}}x^{\lambda-1}e^{-\frac{x}{\beta}};x\geq 0$	$F(x)=\frac{1}{\Gamma(\lambda)}\int_0^x\frac{1}{\beta^{\lambda}}x^{\lambda-1}e^{-\frac{x}{\beta}}dx$	β is a location parameter, $\beta>0$. λ is a shape parameter, $\lambda>0$
Lognormal	$f(x)=\frac{1}{x\sqrt{2\pi\sigma^2}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}};x>0$	$F(x)=\int_0^x\frac{1}{x\sqrt{2\pi\sigma^2}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}dx$	μ is a location parameter, $-\infty<\mu<\infty$. σ is a shape parameter, $\sigma>0$
Inverse Gaussian	$f(x)=\sqrt{\frac{\sigma}{2\pi x^3}}e^{-\frac{\sigma(x-\mu)^2}{2\mu^2x}};x>0$	$F(x)=\int_0^x\sqrt{\frac{\sigma}{2\pi x^3}}e^{-\frac{\sigma(x-\mu)^2}{2\mu^2x}}dx$	μ is a location parameter, $-\infty<\mu<\infty$. σ is a shape parameter, $\sigma>0$

Table 2. Descriptive statistics of daily average concentrations of $PM_{2.5}$ ($\mu g/m^3$) at two stations in Chiang Mai

	Yupparaj Wittayalai school station	City Hall station
Dataset	700	687
Number of data exceed air quality standard of $PM_{2.5}$	82 (11.71%)	116 (16.89%)
Mean	25.18	31.00
Standard deviation	18.84	19.82
Minimum	4.46	4.79
Maximum	108.13	114.00
Skewness	1.70	1.44
Kurtosis	2.65	1.78

Table 3. Estimated parameters for the distributions of daily average concentrations of $PM_{2.5}$

Distribution	Estimated parameters	
	Yupparaj Wittayalai school station	City Hall station
Weibull	$\hat{\beta}$ = 28.19, $\hat{\lambda}$= 1.50	$\hat{\beta}$ = 35.06, $\hat{\lambda}$ = 1.71
Gamma	$\hat{\beta}$ = 10.43, $\hat{\lambda}$= 2.41	$\hat{\beta}$ = 10.29, $\hat{\lambda}$= 3.01
Lognormal	$\hat{\mu}$ = 3.00, $\hat{\sigma}$ = 0.64	$\hat{\mu}$ = 3.26, $\hat{\sigma}$ = 0.58
Inverse Gaussian	$\hat{\mu}$ = 25.18, $\hat{\sigma}$ = 50.97	$\hat{\mu}$ = 31.00, $\hat{\sigma}$= 78.62



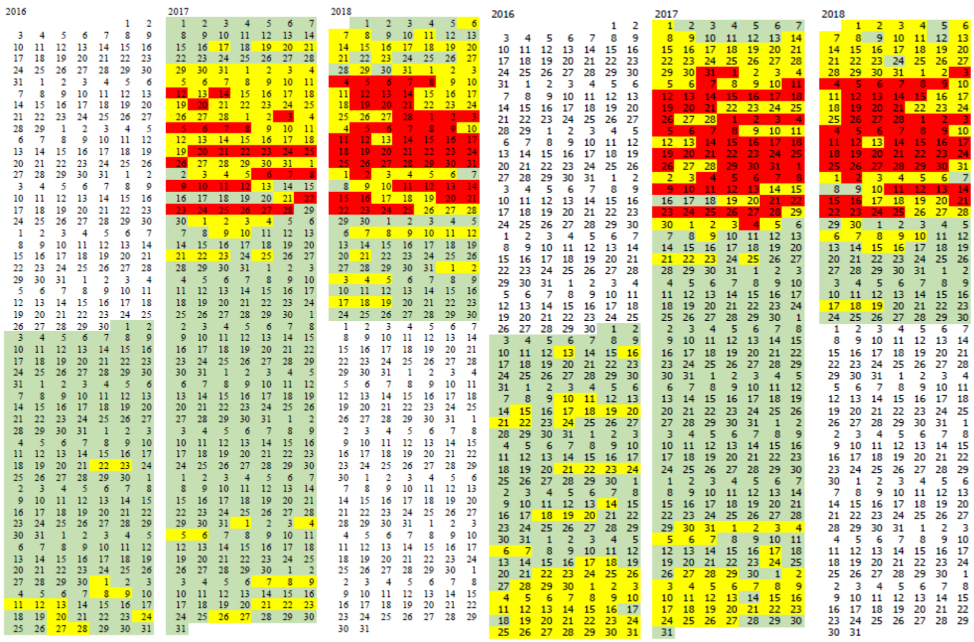
(a) Yupparaj Wittayalai school station (b) City Hall station
Fig. 1: Histograms of daily average concentration of $PM_{2.5}$ by station in Chiangmai.

Table 4. Goodness of fit statistics for the distribution of daily average $PM_{2.5}$ concentration

Distribution	Yupparaj Wittayalai school station		City Hall station	
	Kolmogorov-Smirnov	Anderson-Darling	Kolmogorov-Smirnov	Anderson-Darling
Weibull	0.1097	21.8089	0.2405	15.3122
Gamma	0.1180	18.4528	0.0929	11.0677
Lognormal	**0.0875**	9.0850	0.0609	4.8151
Inverse Gaussian	0.0880	**8.2948**	**0.0575**	**3.9943**

Table 5. Estimated parameters of the Largest Extreme Value distribution and return periods for $PM_{2.5}$ concentration

Station	Estimated parameters	Return period (days)
Yupparaj Wittayalai school	$\hat{\mu} = 61.42$, $\hat{\sigma} = 10.26$	20.97
City Hall	$\hat{\mu} = 61.64$, $\hat{\sigma} = 10.24$	22.61



(a) Yupparaj Wittayalai school station (b) City Hall Station

Fig. 2: Daily average concentration of $PM_{2.5}$ from 1st July 2016 to 30th June 2018. (green = less than 25 $\mu g/m^3$, yellow = 25 - 50 $\mu g/m^3$, red = more than 50 $\mu g/m^3$)