EnvironmentAsia
The international journal by the Thai Society of Higher Education Institutes on Environment

EnvironmentAsia 15(2) 2022 23-33
DOI 10.14456/ea.2022.30
ISSN 1906-1714; ONLINE ISSN: 2586-8861

Assessment of Parameter Estimator of Lognormal Distribution for a Better Long-Term Prediction of PM_{10} Concentrations in Suburban Areas

Muhammad Uthman Omar and Hazrul Abdul Hamid *

School of Distance Education, Universiti Sains Malaysia, 11800 Penang, Malaysia

*Corresponding author: hazrul@usm.my
Received: January 23; Revised: February 28; Accepted: March 16, 2022

Abstract

Lognormal distribution is one of the parent distributions commonly used in air pollution modelling. The nature of lognormal distribution that is skewed to the right makes it suitable for non-extreme air pollution data. Parameter estimation is a critical step in getting the best prediction result with this distribution since the values of parameters might affect the accuracy and errors of the prediction. The main objective of this study is to compare and determine the most appropriate estimator to predict the PM_{10} concentration in suburban area. Using PM_{10} concentrations in Jerantut, Sungai Petani, Muar, and Kuantan, this study assessed the performance of four distinct estimators; method of moments, maximum likelihood estimation, probability weighted moments, and uniformly minimum variance unbiased estimator. The method of moments proved to be the best estimator when five performance indicators were used. It is also worth noting that the method of moments has the lowest scale parameter value and the highest shape parameter value for both Jerantut and Sungai Petani. Not only method of moments show good results, uniformly minimum variance unbiased estimator also shows a good result in terms of accuracy of prediction in Muar and Kuantan.

Keywords: Air pollution; PM_{10}; Statistical distribution; Parameter estimation

1. Introduction

Statistical distribution has been widely used to model air pollution. It is important to understand the characteristics and the distribution of air pollutants for a better representation of data and also to make sure the prediction can be made more precisely. It was reported that air pollution has severed the health of the world population and caused premature deaths of an estimated 4.2 million in 2016 (WHO, 2018). These morbidity and mortality have a close quantitative relationship to the exposure of high concentrations of small particulates.

Predicting major air pollutants such as particulate matter can help relevant parties especially authorities to have a better understanding and draw plans and policies to reduce the impact of air pollution. Among the harmful pollutants is particulate matter with an aerodynamic diameter equal to 10 µm or smaller, which known as PM_{10}. Modelling using statistical distribution can help in prediction of air pollutants (Ul-Saufie *et al.,* 2015). Many studies have been conducted to predict PM_{10} concentration using various parent distributions. Pearson V (Perišić *et al.,* 2015; Todorovic *et al.,* 2015), Weibull (Giavis *et al.,* 2009; Al-Dhurafi *et al.,* 2016; Plocoste *et al.,* 2020), log-logistic (Agarwal and Shiva Nagendra, 2016; Menon and S.M, 2018; Ghavidel *et al.,* 2019), and gamma distributions (Ozel and Cakmakyapan, 2015; Al-Dhurafi *et al.,* 2016; Huerta-Viso *et al.,* 2020) have been used to fit the PM_{10} concentration. Though previous studies have shown that different types of locations such as industrial, urban, and suburban, will have different best- fitting distributions (Karaca *et al.,* 2005; Noor *et al.,* 2011; Qiu *et al.,* 2018), lognormal distribution turned out to be the best fit distribution to fit and predict PM_{10} concentration in most locations in Malaysia (Hamid *et al.,* 2013; Yunus and Hasan, 2017). In general, the lognormal distribution's skewed nature makes it suitable for data that are not too extreme. The extreme value distribution, on the other hand, is more appropriate to utilize if the air pollution values are excessively high, particularly during the high particulate event period.

Fitting lognormal distribution is affected by the value of shape and scale parameters. These two parameters can be estimated by using existing estimators like the method of likelihood (MLE) (Maciejewska *et al.,* 2015; Todorovic *et al.,* 2015) and the method of moments (MOM) (Mijić *et al.,* 2009). These two estimators are popular and widely used though the selection of the estimators is not clear. Some studies use only an estimator while others use several. The selection of estimators is crucial because different estimators can affect the accuracy and errors in predicting PM_{10} concentration. A previous study by Lu (2002) has shown that MLE yielded more accurate results than MOM in predicting PM_{10} concentration in Taiwan. Contrary to a study by Md Yusof *et al.* (2010) that showed MOM gave a better estimation for two-parameter lognormal distribution for the year 2000 – 2002 in Penang Malaysia, as compared to MLE.

Though MLE and MOM are commonly and widely used, there are other estimators as well that are not commonly used like probability weighted moments (PWM) (Hosking, 1990), uniformly minimum variance unbiased estimator (UMVUE) (Shen, 1998) and method of fractiles (Georgopoulos and Seinfeld, 1982). A study by Hamid *et al.* (2013) showed that PWM performs better than MOM in predicting PM_{10} concentration using lognormal distribution in Nilai, Negeri Sembilan. Wan Deraman *et al.* (2017) on flood frequency analyses using lognormal distribution showed that MOM performed better than PWM. Mage and Ott (1984) on 100 simulated years of lognormally distributed air pollution showed that MLE performs better than MOM and the method of fractiles.

Despite having different best estimators for lognormal distribution, parameter estimation using MLE can give the smallest sampling variance of the estimated parameters and the estimated quantiles, thus considered efficient, yet may perform poorly if the number of parameters is large and the sample length is small. Sometimes PWM is more accurate than MLE if the estimation is from small samples (Tosunoğlu, 2018).

Since there has not been the best estimator to estimate the two-parameter lognormal distribution, this study focused on finding the best estimator to fit lognormal distribution in predicting PM_{10} concentration and see how different estimators could affect the values of scale and shape parameters. This study would also see how the estimation values influenced the accuracy and errors in prediction.

2. Materials and Methods

2.1 Data and study area

Data sets of daily average of PM_{10} concentrations for the year assessment of 2016 were obtained from the Department of Environment Malaysia (DOE). With a total of 8760 secondary data, a prediction will be done with the same quantity using the estimated parameters before being fitted to the distribution.

This study focused on suburban areas since this area usually more into residential area. Four monitoring stations in Peninsular Malaysia were selected; Sungai Petani, Muar, Kuantan and Jerantut as reference station. Selected suburban areas are based on the similarities in term of surrounding condition such as the density, preferred as residential area and the distance from industrial area. The information for these stations is given in Table 1. The classification of background and suburban is determined by DOE. The geographical location of the stations on a map is shown in Figure 1.

Table 1. Information for four monitoring stations

Station	Location / Coordinates	Description	Classification
Jerantut	Central 3.9713°N 102.3484°E	- Surrounded by forest - 1km away from River Som Forest - 4km away from town	Background
Sungai Petani	North 5.6284°N 100.4717°E	- Surrounded by fast developing residential areas - Has small industrial areas thatare 7.21 km away from the monitoring station	Suburban
Muar	South 2.0603°N 102.5952°E	- Surrounded by residential areas - 6km away from the seashore - Has a small industrial area, about 3.38km from the station	Suburban
Kuantan	East 3.8178°N 103.2979°E	- Fast-developing city surrounded by residential areasand many tourist attractions - Has many industrial areas andthe closest one is Kawasan Perindustrian Semambu, whichis 5.1km away from the station	Suburban

Figure 1. The location of the monitoring stations in Peninsular Malaysia

2.2 Lognormal Distribution

A statistical distribution of logarithmic values from a related normal distribution is known as a log-normal distribution. Using associated logarithmic calculations, a log-normal distribution can be converted to a normal distribution and vice versa. While most people are familiar with the normal distribution, the log-normal distribution may be unfamiliar. Logarithmic mathematics can be used to convert a normal distribution to a lognormal distribution. Lognormal distribution can only emerge from a normally distributed set of random variables, this is the primary foundation. There are several reasons to use log-normal distributions alongside normal distributions. In general, most lognormal distributions are obtained by using the natural log with e = 2.718 as the base.

The lognormal distribution, on the other hand, can be scaled using a different basis, which changes the shape of the distribution. In the field of air quality, lognormal distributions are often used because of their nature which is skewed to the right which is suitable for the non-extreme data. Several previous studies show that lognormal distribution fit the PM_{10} concentration in Malaysia (Sansuddin *et al.,* 2011; Hamid *et al.,* 2013; Yunus and Hasan, 2017).

The probability density function (pdf) of lognormal distribution is given by Forbes *et al.* (2010) as follows:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right)} \quad (1)$$

where μ is the scale parameter and σ is the shape parameter.

Lognormal distributions are well-known for its suitability for non-extreme data modelling. Previous research has demonstrated, however, that extreme data that occur only infrequently have little effect on the performance of the lognormal distribution (Jaffar *et al.,* 2018). When there is a lot of extreme data or outliers, the extreme value distribution (EVD) is generally utilized. Because the study area consists of suburban areas rather than large industrial areas, the lognormal distribution was chosen carefully based on past research suggestions.

2.3 Parameter Estimation

To plot a probability density function (pdf) graphs of lognormal distribution, the scale and shape parameters need to be estimated. Scale and shape parameter will affect the shape of lognormal distribution. Basically, scale parameter will determine the shrinks or stretches of the curve while the value of shape parameter will affect the general shape of the distribution. The process of parameter estimation is crucial since the estimated values of the parameters can affect the accuracy and errors of prediction. Thus, it was best to use and compare several parameter estimators. In this study, four estimators were used to estimate the two-parameter lognormal distribution, namely method of moments (MOM), method of likelihood estimation (MLE), probability weighted moments (PWM) and uniformly minimum variance unbiased estimator (UMVUE).

2.4 Performance Indicators

To determine the best estimator, five performance indicators were used. Two indicators will evaluate the error while the other three will assess the accuracy of prediction. The indicators are shown in the following table. Estimator performance is evaluated using two error measurements and three accuracy measures. The lowest value for error measures implies that the estimator performs better than another, whilst the value for accuracy measures must approach the value one for a good estimator.

3. Results and Discussion

Table 4 shows the descriptive statistics for PM_{10} concentrations in Jerantut, Sungai Petani, Muar, and Kuantan in 2016. Each station's mean is greater than its median, indicating that the data is skewed to the right. There were no daily exceedances at any of the stations (the threshold limit is 150 µg/m³). However, with a mean of 64.93 µg/m³, Sungai Petani exceeds the yearly limit of 50 µg/m³ specified by the Malaysian Ambient Air Quality Guidelines (MAAQG).

Table 2. Parameter Estimators

Parameter Estimator	Scale Parameter, μ	Shape Parameter, σ	
MOM	$\mu = \ln(\bar{x}) - \frac{\sigma^2}{2}$ $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\sigma = \sqrt{\ln\left[s^2 + (\bar{x})^2\right] - 2\ln(\bar{x})}$ $s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$	Md Yusof *et al*. (2010)
MLE	$\mu = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i)$	$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[\ln x_i - \mu]^2}$	Wang *et al*. (2013)
PWM	$\mu = \log l_1 - \frac{\sigma^2}{2}$ $l_1 = b_0$ $b_0 = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\sigma = 2erf^{-1}\left(\frac{l_2}{l_1}\right)$ $\mathrm{erf}^{-1}(x)$: inverse error function $l_2 = 2b_1 - b_0$ $b_1 = \frac{1}{n}\sum_{i=2}^{n}\frac{(i-1)}{(n-1)} x_i$	Hosking (1990) Chen *et al*. (2004) Wan Deraman *et al*. (2017)
UMVUE	$\mu = \ln(M) - \frac{\sigma^2}{2}$ $M = \exp(\bar{y}) \cdot g\left(\frac{S_F^2}{2}\right)$ $\bar{y} = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i)$ $g(t) = \exp(t) \cdot \left[1 - \frac{t(t+1)}{n} + \frac{t^2\left(3t^2 + 22t + 21\right)^3}{6n^2} + \ldots\right]$ $S_F^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\ln x_i - \bar{y})^2$	$\sigma = \sqrt{\ln\left[V + M^2\right] - 2\ln(M)}$ $V = \exp(2\bar{y}) \cdot \left[g\left(2S_F^2\right) - g\left(\frac{S_F^2(n-2)}{(n-1)}\right)\right]$	Finney (1941), Johnson *et al*. (1994) Ginos *et al*. (2009)

Table 3. Performance Indicators

Indicators	Equations	
The Root Mean Squared Error (RMSE)	$RMSE = \left[\frac{1}{N} \sum_{i=1}^{N} (P_i - O_i)^2 \right]^{\frac{1}{2}}$	(Norazian *et al.* 2008)
The Normalized Absolute Error (NAE)	$NAE = \frac{\sum_{i=1}^{N} \lvert P_i - O_i \rvert}{\sum_{i=1}^{N} O_i}$	Mubin Zahari *et al.* (2019)
The prediction accuracy (PA)	$PA = \frac{\sum_{i=1}^{N} (P_i - \bar{P})(O_i - \bar{O})}{(N-1)\sigma_P \sigma_O}$	Norazian *et al.* (2008)
The coefficient of determination (R^2)	$R^2 = \left(\frac{\sum_{i=1}^{N} (P_i - \bar{P})(O_i - \bar{O})}{(N)\sigma_P \sigma_O} \right)^2$	Norazian *et al.* (2008)
The index of agreement (IA)	$IA = 1 - \frac{\sum_{i=1}^{N} (P_i - O_i)^2}{\sum_{i=1}^{N} \left(\lvert P_i - \bar{O} \rvert + \lvert O_i - \bar{O} \rvert \right)^2}$	Junninen *et al.* (2004)
Notations: N = number of imputations, P_i = the predicted values, O_i = the observed values, $\bar{P}$ = mean of predicted values, $\bar{O}$ = mean of observed values, σ_P = standard deviation of the predicted values, σ_O = standard deviation of the observed values		

Table 4. Descriptive statistics of PM_{10} concentration in Jerantut, Sungai Petani, Muar and Kuantan for the year assessment of 2016

Station Name	**Jerantut**	**Sungai Petani**	**Muar**	**Kuantan**
Mean	28.52	64.93	34.54	39.32
Standard Deviation	11.84	16.31	12.10	12.26
Median	25.58	62.58	32.54	38.70
Mode	17.92	49.79	29.79	25.04
Minimum	3	37	1	3
Maximum	88	149	98	98
Kurtosis	6.88	8.45	7.59	6.03
Skewness	1.55	1.89	1.47	1.03

Despite the fact that PM_{10} concentrations in Sungai Petani did not reach the daily limit of 150 μg/m³, there were many instances of PM_{10} concentrations surpassing 50 μg/m³ throughout the year, affecting the yearly average and exceeding the yearly limit. This is most likely due to the activities of recycling facilities, including illegal factories that have been suspected of burning plastic waste in their compounds (Yasina Yusuf *et al.,* 2020).

The selection of the best estimator was based on goodness-of-fit of five performance indicators namely: the root mean squared error (RMSE), the normalized absolute error (NAE), the prediction accuracy (PA), the coefficient of determination (R2) and the index of agreement (IA). All the parameter estimates, as well as the performance indicators of the estimates of lognormal distribution for the four stations, are given in Table 5.

Table 5. Parameter estimates and performance indicators of the four stations.

Stations	Mean	Standard Deviation	Values of estimators			Performance Indicators (PI)					Estimators	Est. PI Count	Best Estimator
			Estimators	Scale Value	Shape Value	RMSE	NAE	PA	R^2	IA			
Jerantut	28.5228	11.8372	**MOM**	3.2712	0.39864	1.4148	0.028124	0.99299	0.98065	0.99632	**MOM**	5	MOM
			MLE	3.2749	0.38811	1.5478	0.029067	0.99243	0.97954	0.99546	**MLE**	0	
			PWM	3.2747	0.39000	1.5175	0.028901	0.99253	0.97975	0.99566	**PWM**	0	
			UMVUE	3.2738	0.39100	1.5056	0.028623	0.99259	0.97986	0.99574	**UMVUE**	0	
Sungai Petani	64.9257	16.3109	**MOM**	4.1426	0.24739	4.0440	0.038934	0.96880	0.93345	0.98408	**MOM**	4	MOM
			MLE	4.1469	0.22181	4.5032	0.035285	0.96502	0.92618	0.97794	**MLE**	0	
			PWM	4.1477	0.22619	4.3757	0.036135	0.96568	0.92745	0.97960	**PWM**	0	
			UMVUE	4.1468	0.22232	4.4886	0.035275	0.96510	0.92633	0.97814	**UMVUE**	1	
Muar	34.5393	12.1005	**MOM**	3.4842	0.34025	2.0901	0.035512	0.98496	0.96485	0.99230	**MOM**	2	MOM & UMVUE
			MLE	3.4781	0.39859	2.9729	0.068876	0.98753	0.96989	0.98710	**MLE**	0	
			PWM	3.4896	0.32391	2.2489	0.030444	0.98403	0.96303	0.99063	**PWM**	1	
			UMVUE	3.4768	0.40183	3.0606	0.071087	0.98764	0.97010	0.98646	**UMVUE**	2	
Kuantan	39.3170	12.2619	**MOM**	3.6252	0.30466	1.7385	0.023691	0.98992	0.97460	0.99484	**MOM**	2	MOM & UMVUE
			MLE	3.6218	0.33090	2.0003	0.038797	0.99012	0.97498	0.99376	**MLE**	1	
			PWM	3.6272	0.29816	1.7826	0.021716	0.98984	0.97443	0.99445	**PWM**	1	
			UMVUE	3.6213	0.33247	2.0339	0.039845	0.99012	0.97499	0.99358	**UMVUE**	2	

It can be seen from Table 5 that the best estimator can be identified based on the five performance indicators. The least values of RMSE and NAE indicate that the estimators will have the lowest error while the highest values of PA, R2 and IA indicate that the estimators will have the highest accuracy.

For Jerantut and Sungai Petani stations, the best estimator is MOM. MOM performs the best for all goodness-fit tests in which it has the highest accuracy and lowest error, except for NAE at Sungai Petani where UMVUE is the lowest. It can be seen that the scale value of MOM is the lowest as compared to other estimators while the shape value is the highest for both stations. These characters somehow affect the accuracy and errors tests for the two stations. Though findings by Md Yusof *et al.* (2010) showed that MOM performs the best when the scale value of MOM is higher than the other estimator and the shape value is lower for all three years assessment of 2000, 2001 and 2002. However, the contradictory findings are likely based on the location of sampling where the previous study was focusing on one industrial location only.

It is also worth mentioning that MLE performs the worst for Jerantut and Sungai Petani. Contrary to MOM, MLE has the highest value of scale parameter and the lowest value of shape parameter. MLE has the worst accuracy having the lowest value on PA, R2 and IA as compared to others, and the highest value on RMSE and NAE except for NAE at Sungai Petani. Study by Sansuddin et al. (2011) also found that MLE performed poorly when compared to MOM in fitting lognormal distribution for PM10 concentration.

For Muar and Kuantan, MOM and UMVUE are the best estimators. While MOM has the lowest value of RMSE and highest value of IA, UMVUE has the highest value of accuracy for PA and R2, but performs the worst in terms of errors on both RMSE and NAE. It can be seen that the scale value of UMVUE is the lowest as compared to other estimators while the shape value is the highest.

From the values of scale and shape parameters obtained from Table 5, the probability density function (pdf) graphs of PM10 concentration were then plotted and compared by stations as shown in Figure 2.

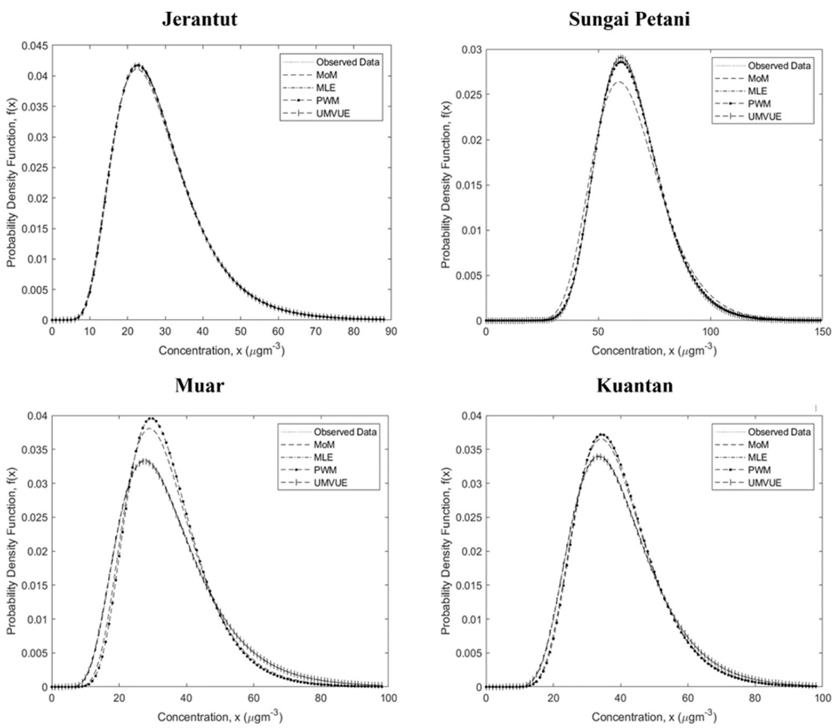

Figure 2. The probability density function plot of PM_{10} concentration for Jerantut, Sungai Petani, Muar and Kuantan

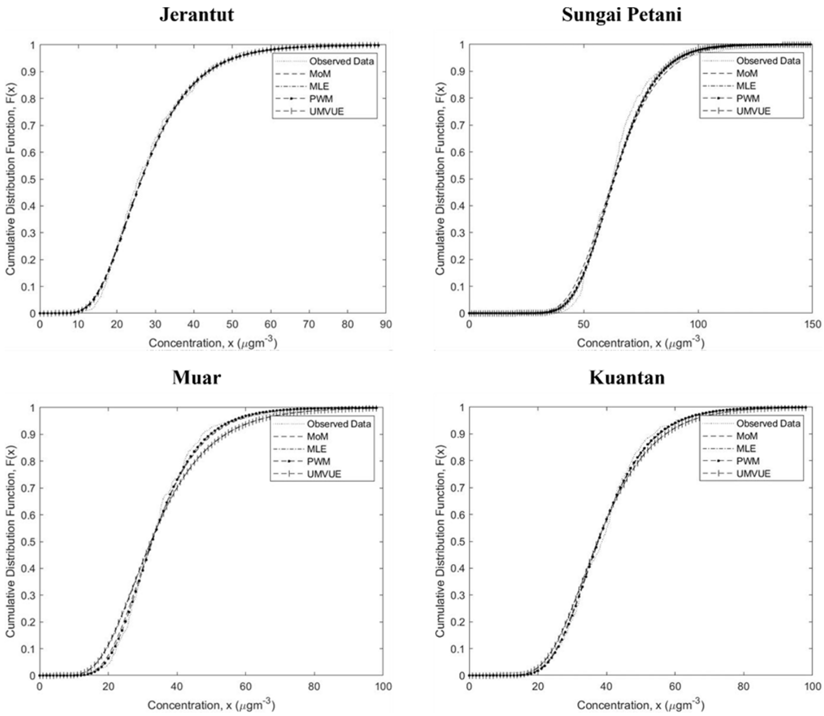

Figure 3. The cumulative distribution function plot of PM_{10} concentration for Jerantut, Sungai Petani, Muar and Kuantan

The pdf plots in Figure 2 show that the distribution for each station is positively skewed. From the figures, it can be clearly seen that the data fitting using all estimators are very close to the observed data in Jerantut. Only Muar and Kuantan, the pdf plot for less efficient estimators can be clearly seen. Figure 3 represents graphs for the cumulative distribution function (cdf) of PM_{10} concentration by stations. This plot can be used for prediction of exceeded and also the return period. From both pdf and cdf plots, it can be seen that no stations exceeded the daily limit set by MAAQG.

4. Conclusion

This study used and compared four different estimators to estimate the scale and shape parameters of the lognormal distribution to predict PM_{10} concentration. The estimated parameters' values are then used for prediction, then compared by stations and tested for goodness of fit. Based on the results, the best estimator to estimate the parameters of lognormal distribution is the method of moments. Method of moments is consistent in terms of high accuracy and low errors for all stations. It is followed by uniformly minimum variance unbiased estimator for Muar and Kuantan stations. Based on descriptive statistics, the PM_{10} readings at the Muar and Kuantan stations had almost similar descriptive statistical values. When the estimated parameter values are compared, the method of moments has the lowest value of the scale parameter and the highest value of the shape parameter for both Jerantut and Sungai Petani. For both Muar and Kuantan, uniformly minimum variance unbiased estimator has the lowest scale value and the highest shape value. This study's findings also indicate that UMVUE can be a good estimator if the value of its shape parameter is greater than that of other estimators.

Acknowledgements

This research was supported by Fundamental Research Grant Scheme (FRGSPJJAUH/6711750), Ministry of Higher Education Malaysia. The authors would like to thank Universiti Sains Malaysia for the support, and the Department of Environment Malaysia (DOE) for generously providing the air quality data used in this study.

References

Agarwal N, Shiva Nagendra SM. Modelling of particulate matters distribution inside the multilevel urban classrooms in tropical climate for exposure assessment. Building and Environment 2016; 102:73–82.

Al-Dhurafi NA, Razali AM, Masseran N, Zamzuri ZH. The probability distribution model of air pollution index and its dominants in Kuala Lumpur. AIP Conference Proceedings 2016; 1784.

Chen Y, Xu S, Sha Z, van Gelder P, Gu SH. Study on L-moment estimations for log-normal distribution with historical flood data. IAHS-AISH Publication 2004; 289:107–113.

Finney DJ. On the Distribution of a Variate Whose Logarithm is Normally Distributed. Supplement to the Journal of the Royal Statistical Society 1941; 7(2):155.

Forbes C, Evans M, Hastings N, Peacock B. Statistical Distributions. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2010.

Georgopoulos, Seinfeld. Statistical distributions of air pollutant concentrations. Environmental Science & Technology 1982; 16(7):401A-416A.

Ghavidel Y, Khorshiddoust AM, Farajzadeh M, Pourshahbaz H. Identification of the thresholds of extreme values and synoptic analysis of PM10 pollution in the atmosphere of Ahvaz. Pollution 2019; 5(3):611–621.

Giavis GM, Kambezidis HD, Lykoudis SP. Frequency distribution of particulate matter (PM10) in urban environments. International Journal of Environment and Pollution 2009; 36(1–3):99–109.

Ginos BF, Grimshaw SD, Engler D, Schaalje GB. Parameter Estimation for the Lognormal Distribution. Statistics. 2009.

Hamid HA, Yahaya AS, Ramli NA, Ul-Saufie AZ. Finding the Best Statistical Distribution Model in PM_{10} Concentration Modeling by using Lognormal Distribution. Journal of Applied Sciences 2013; 13(2):294–300.

Hosking JRM. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal of the Royal Statistical Society: Series B (Methodological) 1990; 52(1):105–124.

Huerta-Viso A, Crespo J, Galindo N, Yubero E, Nicolás JF. Saharan Dust Events over the Valencian Community (Eastern Iberian Peninsula): Synoptic Circulation Patterns and Contribution to PM_{10} Levels. Aerosol and Air Quality Research 2020; 20(11):2519–2528.

Jaffar MI, Hamid HA, Yunus R, Raffee AF. Fitting statistical distribution on air pollution : An overview. International Journal of Engineering and Technology 2018; 7 : 15-21.

Johnson NL, Kotz S, Balakrishnan N. Continuous Univariate Distributions. Wiley & Sons, Michigan USA. 1994.

Junninen H, Niska H, Tuppurainen K, Ruuskanen J, Kolehmainen M. Methods for imputation of missing values in air quality data sets. Atmospheric Environment 2004; 38(18):2895–2907.

Karaca F, Alagha O, Ertürk F. Statistical characterization of atmospheric PM_{10} and $PM_{2.5}$ concentrations at a non-impacted suburban site of Istanbul, Turkey. Chemosphere 2005; 59(8):1183–1190.

Lu HC. The statistical characters of PM_{10} concentration in Taiwan area. Atmospheric Environment 2002; 36(3): 491–502.

Maciejewska K, Juda-Rezler K, Reizer M, Klejnowski K. Modelling of black carbon statistical distribution and return periods of extreme concentrations. Environmental Modelling & Software 2015; 74: 212–226.

Mage DT, Ott WR. An evaluation of the methods of fractiles, moments and maximum likelihood for estimating parameters when sampling air quality data from a stationary lognormal distribution. Atmospheric Environment (1967) 1984; 18(1):163–171.

Md Yusof NFF, Ramli NA, Yahaya AS, Sansuddin N, Ghazali NA, al Madhoun W. Monsoonal differences and probability distribution of PM_{10} concentration. Environmental Monitoring and Assessment 2010; 163(1–4): 655–667.

Menon JS, S.M SN. Statistical distribution and particle dosimetry models to estimate personal exposure at urban sidewalks of tropical climate. Sustainable Cities and Society 2018; 40: 254–265.

Mijić Z, Tasić M, Rajšić S, Novaković V. The statistical characters of PM_{10} in Belgrade area. Atmospheric Research 2009; 92(4): 420–426.

Mubin Zahari N, Ezzah Shamimi R, Hafiz Zawawi M, Zia Ul-Saufie A, Mohamad D. Prediction of Future Ozone Concentration for Next Three Days Using Linear Regression and Nonlinear Regression Models. IOP Conference Series: Materials Science and Engineering 2019; 551: 012006.

Noor NM, Tan C-Y, Ramli NA, Yahaya AS, Yusof NFFM. Assessment of various probability distributions to model PM_{10} concentration for industrialized area in peninsula Malaysia: A case study in shah alam and nilai. Australian Journal of Basic and Applied Sciences 2011; 5(12): 2796–2811.

Norazian MN, Shukri YA, Azam RN, al Bakri AMM. Estimation of missing values in air pollution data using single imputation techniques. ScienceAsia 2008; 34(3): 341–345.

Ozel G, Cakmakyapan S. A new approach to the prediction of PM_{10} concentrations in central Anatolia region, Turkey. Atmospheric Pollution Research 2015; 6(5): 735–741.

Perišić M, Stojić A, Stanišić Stojić S, Šoštarić A, Mijić Z, Rajšić S. Estimation of required PM10 emission source reduction on the basis of a 10-year period data. Air Quality, Atmosphere and Health 2015; 8(4): 379–389.

Plocoste T, Calif R, Euphrasie-Clotilde L, Brute F-N. The statistical behavior of PM_{10} events over guadeloupean archipelago: Stationarity, modelling and extreme events. Atmospheric Research 2020; 241:104956.

Qiu Z, Song J, Hao C, Li X, Gao HO. Investigating traffic-related PM exposure on and under pedestrian bridges: A case study in Xi'an, China. Atmospheric Pollution Research 2018; 9(5): 877–886.

Sansuddin N, Ramli NA, Yahaya AS, Yusof NFFM, Ghazali NA, Madhoun WA al. Statistical analysis of PM_{10} concentrations at different locations in Malaysia. Environmental Monitoring and Assessment 2011; 180(1–4): 573–588.

Shen W-H. Estimation of Parameters of A Lognormal Distribution. Taiwanese Journal of Mathematics 1998; 2(2): 243–250.

Todorovic MN, Perisic MD, Kuzmanoski MM, Stojic AM, Sostarić AI, Mijic ZR, Rajsic SF. Assessment of PM10 pollution level and required source emission reduction in Belgrade area. Journal of Environmental Science and Health - Part A Toxic/ Hazardous Substances and Environmental Engineering 2015; 50(13):1351–1359.

Tosunoğlu F. Accurate estimation of T year extreme wind speeds by considering different model selection criterions and different parameter estimation methods. Energy 2018; 162: 813–824.

Ul-Saufie AZ, Yahaya AS. Ramli NA, Hamid HA. PM_{10} concentrations short term prediction using feedforward backpropagation and general regression neural network in a sub-urban area. Journal of Environmental Science and Technology 2015; 8(2): 59-73.

Wan Deraman WHA, Abd Mutalib NJ, Mukhtar NZ. Determination of return period for flood frequency analysis using normal and related distributions. Journal of Physics: Conference Series 2017; 890(1): 012162.

Wang X, Chen RJ, Chen BH, Kan HD. Application of statistical distribution of PM_{10} concentration in air quality management in 5 representative cities of China. Biomedical and Environmental Sciences 2013; 26(8): 638–646.

WHO. Ambient (outdoor) air pollution. 2018. https://www.who.int/news-room/fact-sheets/detail/ambient-.

Yasina Yusuf S, Rahim NAAA, Zulaikha Wazam W, Naidu D, Jamian R. Indoor and Outdoor Particulate Matter Concentrations in the Vicinity of Plastic Waste Processing Industries. IOP Conference Series: Earth and Environmental Science 2020; 616: 012015.

Yunus RM, Hasan MM. Predicting Hourly PM_{10} Concentration in Seberang Perai and Petaling Jaya Using Log-Normal Linear Data. Proceedings of IASTEM International Conference 2017; (January): 6–10.