MJFAS MALAYSIAN JOURNAL OF FUNDAMENTAL AND APPLIED SCIENCES
PRINT ISSN: 2289-5981 | ONLINE ISSN: 2289-599X

RESEARCH ARTICLE

Quantile regression for analysing PM_{10} concentrations in Petaling Jaya

Ng Kar Yong* and Norhashidah Awang

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM Penang, Malaysia

* *Corresponding author: karyong92@yahoo.com*

Article history
Received 18 January 2017
Accepted 8 May 2017

Abstract

Particulate matter with diameter less than 10µm (PM_{10}) data usually exhibit different variations as they include normal days and pollution days. This paper applied quantile regression (QR) technique to inspect the changing relationship between predictor variables and PM_{10} concentrations at Petaling Jaya monitoring station in the year 2014 over different PM_{10} distributions. For comparative purpose, multiple linear regression (MLR) using ordinary least squares (OLS) estimation approach was also performed. The QR analysis results showed that the interrelationship between predictor variables and PM_{10} was not consistent across the PM_{10} quantile distributions and hence, proved discordancy with MLR estimates. The lagged PM_{10} concentration was the only important factor throughout the quantile distributions of PM_{10}. It was found that the effects of lagged PM_{10}, temperature, carbon monoxide (CO) increased from low to high quantile distributions, while the effects of lagged humidity, east-west wind component, wind speed and nitrogen monoxide (NO) showed the otherwise patterns. The lagged NO associated significantly with PM_{10} at low quantiles, whereas the lagged temperature and CO associated significantly at high quantiles only. Lagged humidity, east-west wind component and wind speed correlated significantly and negatively with PM_{10} at low and middle quantiles. Ozone (O_3), however, had effect of changing nature from positive association at low PM_{10} distributions to negative association at high levels. Thus, QR is helpful to provide a more complete description of predictor variable effects on PM_{10} at different distributions, and may assist in PM_{10} management especially during haze periods.

Keywords: Multiple linear regression, ordinary least squares, quantile regression, PM_{10}

INTRODUCTION

In Malaysia, especially in Klang Valley region, particulate matter with diameter less than 10µm (PM_{10}) has been recognized as one of the major air pollutants (Liew *et al*., 2011). PM_{10} is an air pollutant which is a mixture of solid particles and liquid droplets exists in atmosphere. It is a larger group of coarse particle pollutant which comprises the fine particles ($PM_{2.5}$) and ultrafine particles ($PM_{0.1}$) (Anderson *et al*., 2012). PM_{10} has long been the subject of researches due to its adverse health impacts. PM_{10} has found to be linked to both short-term and long-term health effects. The coarse particles can deposit in upper respiratory airways, and fine particles can penetrate even deeper into smaller airways and alveoli. This will bring about respiratory, lung and heart diseases (Peng *et al*., 2008). Studies have revealed that short-term exposure to PM_{10} was predominantly correlated to cardiovascular morbidity (Anderson *et al*., 2012). Furthermore, PM_{10} also provoked asthma and lung diseases while increasing hospital admissions. On the other hand, long-term exposure was associated with mortality from cardiopulmonary problems (Anderson *et al*., 2012). In addition, exposure of pregnant women to large amount of PM_{10} from vehicular emissions also increased the risk of premature defects (Vinceti *et al*., 2016). Vulnerable population to PM_{10} pollution includes the elders, children and patients with lung and heart illnesses in which their health effects may be more serious (Anderson *et al*., 2012).

In the awareness of the importance to control the air quality for the public interest, Malaysia Department of Environment (DOE) has promulgated the Air Pollution Index (API) in year 1996 to regulate five types of main air pollutants in which PM_{10} is one among them. There are six categories in API indicating the air quality status as shown in Table 1 (DOE, 2000).

Table 1 Categorization of API.

API	Air Quality Status
0-50	Good
51-100	Moderate
101-200	Unhealthy
201-300	Very Unhealthy
>300	Hazardous
>500	Emergency

In order to monitor the concentrations of air pollutants, 52 continuous air monitoring stations were built at strategic locations throughout Peninsular Malaysia, Sabah and Sarawak (DOE, 2015). Referring to the Recommended Malaysia Air Quality Guidelines (RMAQG), the 24-hour and one-year averages of PM_{10} concentrations should be maintained below 150µg m^{-3} and 50µg m^{-3}, respectively. When the daily average PM_{10} concentration exceeds 150µg m^{-3}, it reaches the unhealthy stage in API (DOE, 2000).

PM_{10} is well known for its uniqueness and complexity for which it is not only emitted directly from an emission source, but is also formed through a series of chemical reactions between pollutant gases or precursor gases such as nitrogen oxides (NO_x) and sulphur dioxide (SO_2). The natural sources of PM_{10} include volcanoes and forest fires, while the man-made sources involve industries, traffic, agriculture, construction and combustion (Bhattacharjee *et al*., 1999). Moreover, PM_{10} concentration is also influenced by meteorological parameters such as temperature, humidity, wind speed and direction (Liew *et al*., 2011; Dominick *et al*., 2012). Hence, understanding of the parts played by different gaseous pollutants and meteorological variables in PM_{10} variation is crucial for directive countermeasures.

A vast numbers of studies have been conducted to analyse the interrelationship between PM_{10} and its explanatory variables in Malaysia. For example, the factors correlated to daily mean PM_{10} during summer (May to August) in Klang Valley area has been examined by using multiple linear regression (MLR) (Liew *et al*., 2011). The local meteorological parameters including temperature, humidity and wind speed were found to be correlated significantly to PM_{10} variations. Furthermore, synoptic meteorological factors and foreign hotspot counts were also important factors correlated to PM_{10}, but local hotspot counts had little impact. MLR and Pearson correlation coefficient were also employed to analyse the PM_{10} concentrations at Kuching, Shah Alam and Johor Bahru monitoring stations (Dominick *et al*., 2012). The results showed that PM_{10} correlated negatively with humidity and wind speed, while correlated positively with temperature. In Negeri Sembilan, the PM_{10} concentrations were modelled by combining principal component analysis (PCA) with MLR and feed-forward back-propagation (FFBP) neural network models (Ul-Saufie *et al*., 2013). The variables considered in modelling included lagged PM_{10}, meteorology and air pollutants. The daily PM_{10} concentrations in Klang Valley were studied by performing anomaly detection (Shaadan *et al*., 2015). The findings demonstrated monsoon and weekend effects. There were more extreme PM_{10} concentrations occurred during Southwest and Northeast monsoons and during weekdays. In addition, wind speed was shown to be positively correlated to extreme PM_{10}.

Most of these studies modelled and analysed the PM_{10} using mean distribution. However, it is more meaningful to investigate the effects of explanatory variables at the high quantile distributions of PM_{10} which portraying the high anomalies, considering the health implications (Yu *et al*., 2003).

There have been expanding literatures regarding pollution research showing the usefulness of quantile regression (QR) in describing a more thorough picture of varying effects of explanatory variables on PM_{10} or other pollutants' distributions as well as modelling the nonlinear relationships. For instances, QR was used to study the ozone (O_3) distribution in Athens (Baur *et al*., 2004). It was found that the effects of explanatory variables differ over the O_3 quantile distributions and that QR was capable to delineate the nonlinear relationship between O_3 and the explanatory variables. Furthermore, the prediction performance of QR was compared to MLR, and it was confirmed that QR was better for predicting the future prediction of PM_{10} concentrations in Seberang Perai, Malaysia (Ul-Saufie *et al*., 2012). The impacts of lagged PM_{10}, meteorological and pollutants' variables on PM_{10} concentrations in Makkah were also investigated by using QR (Munir, 2016). The various impacts of meteorological variables on O_3 levels in Hong Kong were evaluated by applying QR and MLR techniques, and the ability of QR dealing with changing effects of meteorology at various percentiles was proven (Zhao *et al*., 2016).

This study aims to analyse, in a detailed way, the relationship between PM_{10} and predictor variables in Petaling Jaya which is one of the locations experiencing high PM_{10} levels, using data in year 2014 by applying QR. However, the prediction is not in the scope of this paper. This paper is arranged as follows. The subsequent section presents the data used in this study. MLR using ordinary least squares (OLS) estimation method and QR models will be briefly explained in the following section. The next section presents the results and discusses their discrepancy between MLR and QR approaches. This paper is then ended with conclusion.

MATERIALS AND METHODS

The data

Petaling Jaya air monitoring station is one of the monitoring stations in highly populated Klang Valley area, situated in an industrial area. The dataset of daily average concentrations of air pollutants (PM_{10}, nitrogen dioxide (NO_2), SO_2, carbon monoxide (CO), O_3 and nitrogen monoxide (NO)) and meteorological variables (temperature, humidity, wind speed and direction) from 1 January 2014 to 17 December 2014 was taken from Malaysia DOE. The dataset had small percentages of missing values ranging from 0.29% to 1.71%. These missing data were estimated by linear interpolation. Inspired by previous works (Baur *et al*., 2004; Siwek & Osowski, 2012), the wind direction was represented by two perpendicular wind components, namely east-west (wx) and north-south (wy) wind components.

As shown in the histogram in Fig. 1, the PM_{10} data is skewed to the right. Hence, the logarithmic PM_{10} ($\ln PM_{10}$) series was used in this study since this transformation produced series reasonably close to normal distribution.

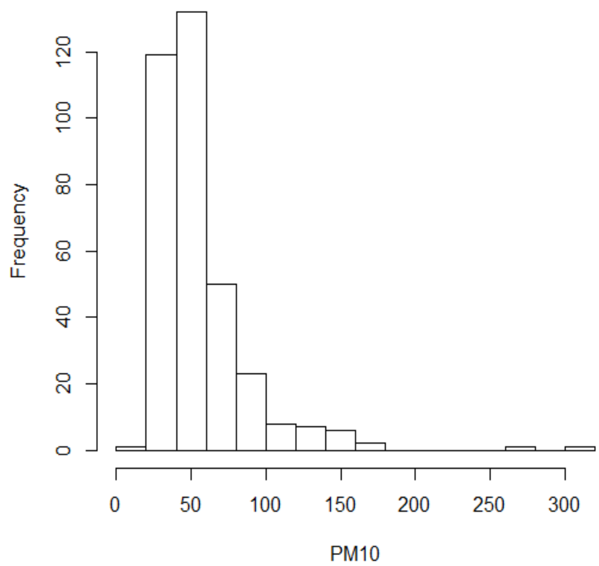

Fig. 1 Histogram of PM_{10} concentrations at Petaling Jaya monitoring station.

Besides studying the relationship between PM_{10} and the predictor variables, this study also hopes to provide some useful information to PM_{10} forecasting model development. Thus, the predictor variables for modelling the MLR and QR would, thereafter, be the lagged (one-previous-day) pollutants and meteorology mentioned above. Furthermore, the lagged $\ln PM_{10}$ concentration was also considered to account for the high autocorrelation (0.71) of $\ln PM_{10}$ data (Baur *et al*., 2004; Munir, 2016).

Multiple linear regression (MLR) with ordinary least squares (OLS) estimation method

MLR model with OLS estimation method, also known as OLS model, is a popular statistical tool for analysing the relationship between response and predictor variables in various fields. It is a linear model expressing the response variable as a function of predictor variables:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \ldots + \beta_p X_{ip} + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{1}$$

where Y is the response variable (i.e. $\ln PM_{10}$ in our case), { $X_k, k = 1, \ldots, p$ } are p predictor variables, and { $\beta_k, k = 0, 1, \ldots, p$ } are $(p + 1)$ regression coefficients, in which β_k measures the change in mean of Y in one unit change of predictor X_k while other predictor variables are held constant. The error term ε is assumed to be independent and identically distributed (i.i.d) with Gaussian distribution. The estimation of conditional expectation function

$$E(Y_i | X_{i1},\ldots,X_{ip}) = \beta_0 + \sum_{k=1}^{p} \beta_k X_{ik} = \mathbf{X}_i^T \boldsymbol{\beta}, \tag{2}$$

or equivalently, the regression coefficients is obtained by minimizing the summation of squared differences between the observed response variable and the fitted values, i.e.

$$\sum_{i=1}^{n} (Y_i - \mathbf{X}_i^T \boldsymbol{\beta})^2. \tag{3}$$

The goodness of fit of the MLR model can be assessed by the coefficient of multiple determination R^2:

$$R^2 = 1 - \frac{\sum_i (Y_i - \hat{Y}_i)^2}{\sum_i (Y_i - \bar{Y})^2} \tag{4}$$

where Y_i is the observed response variable Y, $\hat{Y}_i$ is the fitted value of Y and $\bar{Y}$ is the sample mean of Y. R^2 spans from zero to one. Higher value means better goodness of fit of the model.

Quantile regression (QR) model

QR, as the name suggested, estimates the conditional quantile function, instead of the conditional mean function (Koenker & Hallock, 2001). This method was first advanced by Koenker and Bassett in 1978 (Koenker & Bassett, 1978). A linear QR model takes the form of

$$Y_i = \beta_0(\tau) + \beta_1(\tau) X_{i1} + \ldots + \beta_p(\tau) X_{ip} + \varepsilon_i(\tau), \quad \tau \in (0,1), \quad i = 1,\ldots,n \tag{5}$$

where Y is the response variable, $\{X_k, k = 1,\ldots,p\}$ are p predictor variables, and $\{\beta_k(\tau), k = 0,1,\ldots,p\}$ are $(p + 1)$ quantile coefficients at the τ-th quantile. Here, the $\beta_k(\tau)$ quantifies the change in the τ-th quantile of Y due to the one unit marginal change in predictor X_k. In this way, QR generates a set of coefficients and equations at different quantiles. Thus, QR is able to provide a more holistic picture of the effects of predictors at various PM_{10} distributions. The error terms are not imposed by any distributional assumption or homoscedasticity condition. As such, QR is more robust comparing to MLR. The τ-th conditional quantile function is then given as

$$Q_Y(\tau | X_1,\ldots,X_p) = \beta_0(\tau) + \sum_{k=1}^{p} \beta_k(\tau) X_{ik} = \mathbf{X}_i^T \boldsymbol{\beta}(\tau). \tag{6}$$

Differing from the OLS estimation method, QR minimizes a sum of asymmetrically weighted absolute differences by simplex method to estimate the quantile coefficients:

$$\begin{aligned} \hat{\boldsymbol{\beta}}(\tau) &= \arg\min_{\boldsymbol{\beta}(\tau)\in\mathbb{R}^p} \sum_{Y_i \geq \mathbf{X}_i^T \boldsymbol{\beta}(\tau)} \tau \left| Y_i - \mathbf{X}_i^T \boldsymbol{\beta}(\tau) \right| \\ &\quad + \sum_{Y_i < \mathbf{X}_i^T \boldsymbol{\beta}(\tau)} (1-\tau) \left| Y_i - \mathbf{X}_i^T \boldsymbol{\beta}(\tau) \right| \\ &= \arg\min_{\boldsymbol{\beta}(\tau)\in\mathbb{R}^p} \sum_i \rho_\tau \left(Y_i - \mathbf{X}_i^T \boldsymbol{\beta}(\tau) \right) \end{aligned} \tag{7}$$

where $\rho_\tau(z) = z(\tau - I(z<0))$ and I is the indicator function. The *xy*-pair or design matrix bootstrap method was used to estimate the standard errors since it holds under more general heteroscedasticity situation (Koenker, 2005).

In order to evaluate the goodness of fit of QR model, pseudo R^2 measure (R_τ^1) which is analogous to R^2 in MLR model was used (Koenker & Machado, 1999). However, dislike R^2 which assesses the global goodness of fit, R_τ^1 measures the local goodness of fit of QR model at the τ-th quantile. It is computed by

$$R_\tau^1 = 1 - \frac{\sum_i \rho_\tau (Y_i - \hat{Y}_i)}{\sum_i \rho_\tau (Y_i - Y_q)} \tag{8}$$

where $\hat{Y}_i$ is the fitted value of Y for the τ-th QR model and Y_q is the sample quantile of Y estimated by the restricted QR model including intercept only. R_τ^1 is also ranged [0,1], and the value closer to one indicates better goodness of fit of the QR model.

The MLR and QR methods were implemented by R software.

RESULTS AND DISCUSSION

Initially, MLR and QR were estimated by using all predictor variables. For QR model, the quantile coefficients were estimated at quantiles from 0.05 to 0.95 with increment of 0.05. Next, the predictors which were insignificant at 10% significance level for both regressions were removed to fit a simpler model. The bootstrap method resampled randomly. By a few times of experiments, it was observed that the significance of variables varied within 5% and 10% significance level. Hence, 10% significance level was chosen to obtain a more consistent result. For the same reason, 200 replicates were selected for bootstrapping. After removal of insignificant variables, the regression models were left with lagged $\ln PM_{10}$, temperature, humidity, east-west wind component, wind speed, CO, O_3 and NO. The variance inflation factor (VIF) values for the included predictors are all below 10 (ranged from 1.30 to 4.50), indicating no serious multicollinearity among them (Zhao *et al.*, 2016). Fig. 2 summarizes the estimated quantile coefficients as well as the MLR estimates. The dots with grey-shaded area are the estimated quantile coefficients with the 90% bootstrap confidence interval. The red horizontal line is the MLR estimate, with two parallel dashed lines representing the 90% confidence interval. The black horizontal line indicates zero value.

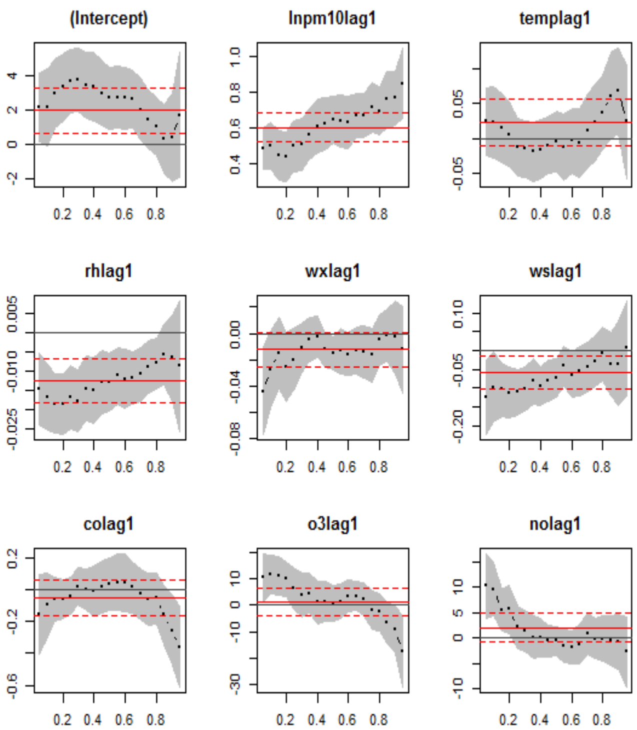

Fig. 2 Quantile regression from 0.05 to 0.95 quantiles.

From Fig. 2, it can be seen that the MLR estimates show that there are only three significant predictors, namely, the lagged PM_{10} concentration, humidity and wind speed, correlated to PM_{10} concentration as their confidence intervals do not include zero value. The lagged PM_{10} concentration shows positive correlation, as expected, since the previous-day concentration can linger in atmosphere and contribute to the next-day concentration (Munir, 2016). On the other hand, the lagged humidity and wind speed correlate negatively to PM_{10} concentration. This is because high humidity is usually related to high rainfall amounts which dilute the PM_{10} concentrations in the atmosphere (Dominick *et al*., 2012). Negative coefficient of wind speed indicates that wind speed acted as dispersing agent of PM_{10} concentrations in the region. However, there have been mixing results with respect to the nature of association between wind speed and PM_{10} (Shaharuddin *et al*., 2008). Therefore, its effect is worth investigates more deeply by using QR.

On the other side, QR offers richer information regarding the effects of predictors. Similar to the MLR estimate, the lagged PM_{10} is positively and significantly associated with PM_{10} at all quantiles, suggesting the persistence effect of previous-day PM_{10} concentration. Moreover, the positive persistence effect increases as the quantile increases. This explains that the previous-day PM_{10} concentration had larger effect on the next-day concentration at extreme PM_{10} distribution compared to low level distributions.

The lagged temperature is significant only at 0.90 quantile with positive association. This can be explained from two aspects. The elevation in temperature promoted the formation of secondary PM_{10} (Munir, 2016). Furthermore, it is also well known that PM_{10} pollution in Malaysia was linked closely to the peatland burning or forest fires from neighbouring country (Shaharuddin *et al*., 2008). Hence, at the extreme PM_{10} values, temperature tended to show its strong effect where it encouraged the burning activities in wide scale.

The effect of lagged humidity is negative which is same as that in the MLR model. Nonetheless, the magnitudes of the effect are not constant but decreasing across quantiles and are significant from 0.05 to 0.80 quantiles. This might tell that humidity did not play a crucial role during PM_{10} peaks.

The lagged east-west wind component is significant at 0.05 and 0.50 quantiles with negative association with PM_{10} concentrations. The negative coefficients may imply the eastern wind that brought about the PM_{10} concentrations.

The lagged wind speed shows significant impact from low to middle quantiles. Like humidity, it also has declining effect on PM_{10} when passing through towards high concentration distribution. The wind speed is negatively associated with PM_{10} except at 0.95 quantile. Hence, it can be said that wind speed had the clearing effect on PM_{10} concentrations in normal days. At high levels, the positive though insignificant correlation means that the wind accumulated PM_{10} concentrations. This may due to the fact that the wind blew the PM_{10} from outside the country, suggesting that the PM_{10} pollution was regional during peak period. This is in agreement with the findings in previous studies (Shaharuddin *et al*., 2008; Shaadan *et al*., 2015).

Although the air pollutants are anticipated to have positive impacts on PM_{10}, lagged CO demonstrates significant negative impact at high PM_{10} distributions (τ = 0.90, 0.95). The lagged O_3 exhibits complex relationship with PM_{10} since its effect changes from positive to negative when moving across quantiles. Similar pattern also occurs on lagged NO. The lagged O_3 shows significant effect at two extreme quantile distributions with opposite signs. The positive coefficients of O_3 at low quantiles may be due to the common precursor gases such as NO_x and volatile organic compounds (VOCs) (Bhattacharjee *et al*., 1999; Rahman *et al*., 2015). The lagged NO shows positive and significant effect at low quantile distributions. This is logical as NO is one of the constituents of NO_x. The increase in previous-day NO concentration tends to raise the next-day PM_{10} concentration. On the other hand, the negative coefficients of the lagged CO and O_3 at high PM_{10} distribution may suggest the different emission origins with PM_{10}. CO and O_3 were predominantly attributed to vehicular emissions (Rahman *et al*., 2015), while high PM_{10} concentrations were mostly because of the biomass burning from outside the country.

To conclude, the lagged PM_{10} concentration is the only predictor which exhibits significant effect at all PM_{10} levels. The lagged humidity, east-west wind component, wind speed and NO are significant at low and/or middle quantiles, whereas lagged temperature and CO demonstrates their dominant effects merely at high PM_{10} levels. The lagged O_3 shows significant impact at both tails of PM_{10} distributions. The QR has shown its ability to uncover the heterogeneous effects of various predictors throughout the quantiles. Particularly, the MLR approach fails to detect the significant impacts of lagged temperature, east-west wind component and air pollutants on PM_{10} concentrations since they are significant at either ends of the PM_{10} distributions only. Therefore, QR is competent to contribute deeper insights on the effects of predictors.

Furthermore, R_τ^1 also proved that the goodness of fit of the model varies over different quantiles. Fig. 3 presents the R_τ^1 of QR model at different quantiles.

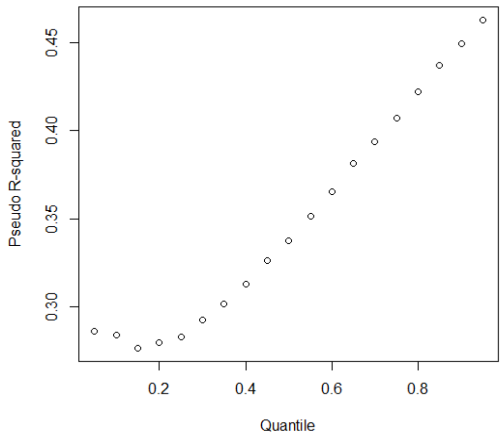

Fig. 3 R_τ^1 of the QR model for quantiles from 0.05 to 0.95

The values of R_τ^1 drop from 0.29 at 0.05 quantile to 0.28 at 0.15 quantile, and then steadily grow up to 0.46 at 0.95 quantile. This suggests that the PM_{10} distributions at high levels are better explained by the model compared to the lower quantiles. This might suggest that the lagged air pollutants and meteorology played larger role in PM_{10} variation during haze period than any other time.

CONCLUSION

This paper studied the effects of lagged pollutants and meteorology on the concentration of PM_{10}. The QR was able to unveil the heterogeneous effects of predictors at different PM_{10} distribution levels which were otherwise hidden by the MLR method. For example, the lagged PM_{10} was not only having positive and significant association with the next-day PM_{10} concentration, but its strength rose as the quantile increases. Furthermore, MLR estimates indicated that the lagged temperature, east-west wind component and air pollutants were not significant, but QR showed significant effects at either high or low ends of distributions which the central tendency MLR estimates unable to detect. Moreover, the goodness of fit showed that the model explained the PM_{10} distribution better at high quantiles than at low quantiles.

In conclusion, QR is a more favourable comprehensive tool than MLR as it renders deeper understanding of the effects of predictors at various distribution levels of PM_{10}. It is also more flexible to be applied to non-Gaussian or heteroscedastic data. In the meantime, the results are readily to be interpreted. Finally, it is hoped that the analysis from this study can help in air quality control during haze episodes as well as developing forecasting model.

ACKNOWLEDGEMENT

We would like to acknowledge Malaysia DOE for providing data. The first author also appreciates the financial support from Universiti Sains Malaysia (Fellowship Scheme) and Ministry of Higher Education.

REFERENCES

Anderson, J. O., Thundiyil, J. G., Stolbach, A. 2012. Clearing the air: A review of the effects of particulate matter air pollution on human health. *Journal of Medical Toxicology* 8, 166-175.

Baur, D., Saisana, M., Schulze, N. 2004. Modelling the effects of meteorological variables on ozone concentration - A quantile regression approach. *Atmospheric Environment* 38, 4689-4699.

Bhattacharjee, H., Drescher, M., Good, T., Hartley, Z., Leza, J-D., Lin, B., Moss, J., Massey, R., Nishino, T., Ryder, S., Sachs, N., Tozan, Y., Taylor, C., Wu, D. 1999. *Particulate Matter in New Jersey*, Introduction, Section 1. Princeton University.

DOE (Department of Environment). 2000. *A guide to Air Pollution Index (API) in Malaysia*. Kuala Lumpur: Ministry of Science, Technology and the Environment.

DOE (Department of the Environment). 2015. *Malaysia Environmental Quality Report 2014*. Kuala Lumpur: Ministry of Natural Resources and Environment Malaysia.

Dominick, D., Latif, M. T., Juahir, H., Aris, A. Z., Zain, S. M. 2012. An assessment of influence of meteorological factors on PM_{10} and NO_2 at selected stations in Malaysia. *Sustainable Environment Research* 22, 305-315.

Koenker, R. 2005. *Quantile Regression*. New York: Cambridge University Press.

Koenker, R., Bassett, G. 1978. Regression quantiles. *Econometrica* 46, 33-50.

Koenker, R., Hallock, K. F. 2001. Quantile regression. *Journal of Economic Perspectives* 15(4), 143-156.

Koenker, R., Machado, J. A. F. 1999. Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association* 94, 1296-1310.

Liew, J., Latif, M. T., Tangang, F. T. 2011. Factors influencing the variations of PM10 aerosol dust in Klang Valley, Malaysia during the summer. *Atmospheric Environment* 45, 4370-4378.

Munir, S. 2016. Modelling the non-linear association of particulate matter (PM_{10}) with meteorological parameters and other air pollutants—a case study in Makkah. *Arabian Journal of Geosciences* 9: 64.

Peng, R. D., Chang, H. H., Bell, M. L., McDermott, A., Zeger, S. L., Samet, J. M., Dominici, F. 2008. Coarse particulate matter air pollution and hospital admissions for cardiovascular and respiratory diseases among medicare patients. *JAMA* 299, 2172-2179.

Rahman, S. R. A., Ismail, S. N. S., Ramli, M. F., Latif, M. T., Abidin, E. Z., Praveena, S. M. 2015. The assessment of ambient air pollution trend in Klang Valley, Malaysia. *World Environment* 5, 1-11.

Shaadan, N., Jemain, A. A., Latif, M. T., Mohd Deni, S. 2015. Anomaly detection and assessment of PM_{10} functional data at several locations in the Klang Valley, Malaysia. *Atmospheric Pollution Research* 6, 365-375.

Shaharuddin, M., Zaharim, A., Mohd. Nor, M. J., Karim, O. A., Sopian, K. 2008. Application of wavelet transform on airborne suspended particulate matter and meteorological temporal variations. *WSEAS Transactions on Environment and Development* 4, 89-98.

Siwek, K., Osowski, S. 2012. Improving the accuracy of prediction of PM_{10} pollution by the wavelet transformation and an ensemble of neural predictors. *Engineering Applications of Artificial Intelligence* 25, 1246-1258.

Ul-Saufie, A. Z., Yahaya, A. S., Ramli, N. A., Abdul Hamid, H. 2012. Future PM_{10} concentration prediction using quantile regression models. *The proceedings of 2012 2nd International Conference on Environmental and Agriculture Engineering IPCBEE, IACSIT Press, Singapore* 37, 15-19.

Ul-Saufie, A. Z., Yahaya, A. S., Ramli, N. A., Rosaida, N., Abdul Hamid, H. 2013. Future daily PM_{10} concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). *Atmospheric Environment* 77, 621-630.

Vinceti, M., Malagoli, C., Malavolti, M., Cherubini, A., Maffeis, G., Rodolfi, R., Heck, J. E., Astolfi, G., Calzolari, E., Nicolini, F. 2016. Does maternal exposure to benzene and PM_{10} during pregnancy increase the risk of congenital anomalies? A population-based case-control study. *Science of the Total Environment* 541, 444-450.

Yu, K., Lu, Z., Stander, J. 2003. Quantile regression: applications and current research areas. *The Statistician* 52(3), 331-350.

Zhao, W., Fan, S., Guo, H., Gao, B., Sun, J., Chen, L. 2016. Assessing the impact of local meteorological variables on surface ozone in Hong Kong during 2000-2015 using quantile and multiple line regression models. *Atmospheric Environment* 144, 182-193.