Comparing the Performance of Statistical Models for Predicting PM$_{10}$ Concentrations

Arwa S. Sayegh^{1*}, Said Munir^2, Turki M. Habeebullah^2

^{1} *SETS International, Beirut, Lebanon*
^{2} *The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al Qura University, Makkah, Saudi Arabia*

ABSTRACT

The ability to accurately model and predict the ambient concentration of Particulate Matter (PM) is essential for effective air quality management and policies development. Various statistical approaches exist for modelling air pollutant levels. In this paper, several approaches including linear, non-linear, and machine learning methods are evaluated for the prediction of urban PM$_{10}$ concentrations in the City of Makkah, Saudi Arabia. The models employed are Multiple Linear Regression Model (MLRM), Quantile Regression Model (QRM), Generalised Additive Model (GAM), and Boosted Regression Trees1-way (BRT1) and 2-way (BRT2). Several meteorological parameters and chemical species measured during 2012 are used as covariates in the models. Various statistical metrics, including the Mean Bias Error (MBE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), the fraction of prediction within a Factor of Two (FACT2), correlation coefficient (R), and Index of Agreement (IA) are calculated to compare the predictive performance of the models. Results show that both MLRM and QRM captured the mean PM$_{10}$ levels. However, QRM topped the other models in capturing the variations in PM$_{10}$ concentrations. Based on the values of error indices, QRM showed better performance in predicting hourly PM$_{10}$ concentrations. Superiority over the other models is explained by the ability of QRM to model the contribution of covariates at different quantiles of the modelled variable (here PM$_{10}$). In this way QRM provides a better approximation procedure compared to the other modelling approaches, which consider a single central tendency response to a set of independent variables. Numerous recent studies have used these modelling approaches, however this is the first study that compares their performance for predicting PM$_{10}$ concentrations.

Keywords: Performance evaluation; Multiple linear regression; Quantile regression model; Generalised additive model; Boosted regression trees.

INTRODUCTION

The main objectives of modelling air quality are to obtain air quality forecasts, quantify air pollutants temporal trends, increase scientific understanding of the underlying mechanisms for production and destruction of pollutants, and estimate potential air pollution related health effects (e.g., Baur *et al.*, 2004).

Many previous investigations have used statistical modelling techniques and machine learning methods to analyse and predict concentrations of Particulate Matter (PM) with an aerodynamic diameter of up to 10 μm (PM$_{10}$). Aldrin and Haff (2005) used generalised additive modelling to relate traffic volumes, meteorological conditions, and

time-related variables to model five pollutants separately, one of which was PM$_{10}$ concentrations. Aldrin and Haff (2005) concluded that the most important predictors for modelling PM$_{10}$ were related to traffic volume, wind vector and relative humidity. Grivas and Chaloulakou (2006) used both Neural Networks (NN) and Multiple Linear Regression (MLR) models for predicting hourly PM$_{10}$ concentrations. They related PM$_{10}$ concentration to input variables, such as lagged PM$_{10}$, wind conditions, day of the week, and hour of the day. More recently, Munir *et al.* (2013a) modelled PM$_{10}$ concentrations with the aid of meteorological variables and traffic-related air pollutant concentrations, such as Carbon Monoxide (CO), Sulphur Dioxide (SO$_2$), Nitric Oxide (NO), Nitrogen Dioxide (NO$_2$), and lagged PM$_{10}$, employing generalised additive models. They identified that meteorological variables, such as temperature and wind speed largely controlled PM$_{10}$ concentrations. McKendry (2002) employed artificial neural networks relating PM$_{10}$ and PM$_{2.5}$ concentrations to meteorological variables, persistence, and co-pollutant data. Mckendry (2002) found that

---

* Corresponding author.
  Tel.: +961-76-014414; Fax: +961-1-755422
  E-mail address: asayegh@setsintl.net

meteorological variables (e.g., wind speed and temperature), persistence (which is likely to reflect the multiday synoptic time scales that modulate dispersion conditions), and co-pollutants (e.g., Ozone (O_3) and Nitrogen Oxide (NO_x)) were useful for predicting PM.

Furthermore, a number of studies have compared the performance of various modelling approaches to determine the best model for the prediction of PM_{10} in different locations. Kukkonen et al. (2003) compared the performance of NN, linear regression model, and a deterministic modelling system in the prediction of both PM_{10} and NO_2 concentrations in Helsinki. Chaloulakou et al. (2003), Papanastasiou et al. (2007), and Ul-Saufie et al. (2011), comparing MLR with NN models for the prediction of PM_{10}, have concluded that non-linear NN method showed better performance. Alternatively, Pires et al. (2008) investigated the performance of five linear models: MLR, principal component regression, independent component regression, quantile regression, and partial least squares regression. The study showed that the dataset size was a critical parameter in the evaluation of models for predicting daily mean PM_{10} concentrations; and concluded that independent component regression was more efficient when using a smaller dataset, whereas partial least squares regression was more efficient when using a larger dataset. In addition, Westmoreland et al. (2007) assessed the performance of GAM with dispersion modelling approach (ADMS-Urban) and favoured the use of GAM, whereas Baur et al. (2004) compared the performance of Quantile Regression Model (QRM) with MLRM, where QRM significantly outperformed MLRM for predicting O_3 concentrations. Carslaw et al. (2009) suggested the use of Boosted Regression Trees (BRT) model for predicting NO_x concentration at mixed source location.

Despite the extensive research on comparing different modelling techniques (e.g., Kukkonen et al., 2003; Paschalidou et al., 2011; Ul-Saufie et al., 2011), no specific research has been devoted into comparing a number of linear and non-linear models for predicting PM_{10} levels. Authors (e.g., El-Assouli, 2010; Seroji, 2011; Seroji, 2012; Habeebullah, 2013; Munir et al., 2013a, b) have analysed the levels, composition, and toxicity of airborne PM in Makkah, particularly during the peak season of Hajj (Pilgrimage) when millions of Muslims visit Makkah. PM_{10}, with the fine and coarse particle fractions combined, is a highly toxic traffic-related pollutant; in a number of studies (Jaecker-Voirol and Pelt, 2000; Pires et al., 2008; Pisoni and Volta, 2009; Pai et al., 2011; Habeebullah, 2013) both short-term and long-term exposures to PM_{10} have been identified to have negative consequences on public health. Health damages associated with exposure to PM_{10} have urged decision makers to act accordingly, however this requires better understanding of the different factors influencing PM_{10} levels and accurate forecasting of its atmospheric concentrations, employing different modelling approaches. The comparison of the operational performance of different models pertinent to local conditions is required so that the model with better performance can be identified. This paper intends to compare the performance of five

statistical models: MLR, GAM, QRM, and BRT with 1-way (BRT1) and 2-way (BRT2) variable interaction to identify the best approach for modeling the atmospheric PM_{10} concentrations. This is the first study on this topic and the findings would be helpful for better air quality management in Makkah and elsewhere.

METHODOLOGY

Study Site and Data Used

The City of Makkah, one of the large cities in the Kingdom of Saudi Arabia (KSA), exhibits special features in comparison to other cities in the world in terms of its location, its environment, and its significance to a significant number of the world population. Makkah is located to the West of Riyadh, the capital of KSA, and is surrounded by the Red Sea to the west. Characterised by its mountainous terrain and harsh topography, Makkah suffers from a major challenge limiting its expansion and leading to dense urban areas within. Nevertheless, due to the existence of Masjid Al Haram (the Holy Mosque), the City of Makkah with a population of 7,026,805 in 2010 (Central Department of Statistics and Information, 2010) holds large-scale religious events such as Hajj and Umrah which attract millions of visitors each year; in particular it attracted approximately 3 million people in October 2012 (Central Department of Statistics and Information, 2012). This, inevitably, leads to extremely high traffic demand and unusual demand patterns in comparison to other normal cities and consequently exerts higher pressure on the environment and on air quality, in particular.

The Hajj Research Institute (HRI) at Umm Al-Qura University runs several air quality and meteorology monitoring stations in Makkah, where several air pollutants and meteorological parameters are measured continuously. Hourly concentrations of CO, SO_2, NO, NO_2, and PM_{10} are amongst the pollutants' parameters measured at the monitoring site, selected for this study. In addition, wind speed, wind direction, temperature, relative humidity, rainfall, and pressure are also continuously monitored at the Presidency of Meteorology and Environment (PME) monitoring site.

The PME site is situated near the Holy mosque to its East. Al Masjid Al Haram road, Ar-Raqubah minor road, and King Abdul Aziz road are located 0.2 km east, 0.25 km northwest, and 0.45 km south of the monitoring site, respectively. In addition to the local road traffics, potential sources of particle emissions include the construction site to the northwest, and the two bus stations along Al Masjid Al Haram road. More detailed information about the site can be found in Munir et al. (2013a, b). Furthermore, Makkah being a part of a tropical arid country, the pollution levels are also affected by the harsh meteorological conditions with high temperatures, frequent sand storms, and very low rainfall.

PM_{10} concentration levels are monitored using a Dust Beta-Attenuation Monitor (BAM 1020, supplied by Horiba), which provides each 30 minute concentration levels in the units of $\mu g/m^3$. BAM consists of a paper band filter located

between a source of beta rays and a radiation detector, where a pump draws ambient air through the filter and the reduction in intensity of beta-radiation measured at the detector is proportional to the mass of particulate deposited on the filter. The sample flow is 16.7 L/m (adjustable within the range of 0 to 20 L/m) and not heated, no correction was thus applied. The temperature range is 0 to +50°C. Also, the filter tape is changed every two months; the inlet filter is cleaned every two months or after a rainfall event. The lower detection limit is about 1.4 µg/m^3 with a range of 1000 µg/m^3. As for the calibration checks of the continuous monitoring system, automatic hourly zero/span adjustments are performed.

A summary of the dataset (January to December, 2012) used in this paper is shown in Table 1.

Selection of Predictor Variables

Road transport is considered a major source of primary PM. In this study, data on traffic flow and speed is not available; therefore, CO and NO$_x$ are considered as surrogates for the traffic conditions variables (Pont and Fontan, 2000). On the other hand, monitored SO$_2$ and NO$_x$ are included in the model as a source of secondary PM$_{10}$ (Kim *et al.*, 2000; Sawant *et al.*, 2004).

Since pressure is rather static and rainfall is zero throughout the study period (2012), therefore only four meteorological variables were included in the model, which are wind speed ($\bar{U}$), wind direction (∅), Temperature (T), and Relative Humidity (RH). Wind speed and wind direction play a significant role in the transport, dilution, and re-suspension of soil particles as indicated in a number of studies such as Harrison *et al.* (1997) and Godish (1997), while temperature and relative humidity are reported to have a strong impact on PM$_{10}$ concentration (Branis and Vetvicka, 2010; Barmpadimos *et al.*, 2011).

In addition to the above predictors, lag_PM$_{10}$ (PM$_{10}$ concentration of the previous day) is also added as a covariate in the model, since fine and ultrafine particles stay in the atmosphere for long time and contribute to the measured concentrations hours or even days later (Baur *et al.*, 2004; AQEG, 2005; Munir *et al.*, 2013a).

Data Processing and Model Evaluation

The one year (2012) data are split into two subsets: training dataset and testing dataset. The training dataset is used for model development, whereas the testing dataset is used to validate the model. The testing dataset is thus an independent set not used in the model development process. Specifically, data corresponding to one month (June) are used as the independent testing dataset, whereas the remaining 11 months are used as the training dataset. All the results presented accordingly are obtained through the analysis of an independent data to the model development process and this ultimately provides the real forecasting ability of the models.

The five models are developed using the training dataset and their predictive performance is assessed using the testing dataset. To conduct a thorough and insightful evaluation and comparison of the five different models, a range of statistic indicators is necessary. Hence, the predictive performance of the MLRM, GAM, QRM, BRT1, and BRT2 models are compared through the calculation of a selected number of statistical parameters, as suggested by Martins *et al.* (2009), Derwent *et al.* (2010), and Carslaw (2011). These are the Mean Bias Error (MBE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), the fraction of predictions within a Factor of two (FACT2), the Pearson correlation coefficient (R) and the Index of Agreement (IA). Equations for each of the metrics are provided in Table 2. While the MBE provides an indication of whether the predictions are over or under estimated, MAE, RMSE, and FACT2 provide a good indication of how close the modelled and observed values are (Martins *et al.*, 2009; Carslaw, 2011). A negative MBE value indicates underestimation, whereas a positive MBE indicates an overestimation of the predicted PM$_{10}$ concentrations. Higher MAE and RMSE values indicate higher error, which shows poorer agreement of the modelled and observed values. FACT2 values closer to one indicate closer match between observed and modelled values and thus indicate better model performance. The correlation coefficient provides a measure of deviation between modelled and observed values. Furthermore, a dimensionless index to illustrate the error amount is also needed and this is addressed by the IA parameter as suggested by different researchers, such as Chaloulakou *et al.* (2003) and Paschalidou *et al.* (2011). In addition to the abovementioned parameters, the Standard Deviation (SD) has also been used as a measure to check difference in the ability of models in capturing the variability of original data; therefore SD calculated for both measured and predicted data are compared.

Furthermore, a non-parametric Wilcoxon Signed Rank Test is performed to test whether the difference in the mean values of the modelled and observed PM$_{10}$ concentrations

Table 1. Data summary for key monitored variables, Makkah – 2012.

Variables	Mean	Median	Percentile			Data Capture (%)
			25th	75th	95th	
PM$_{10}$ (µg/m^3)	157.1	123.0	79.0	195.0	402.0	92.8%
CO (mg/m^3)	1.1	1.0	0.8	1.3	2.2	95.4%
SO$_2$ (µg/m^3)	11.0	8.0	5.0	16.0	30.0	87.5%
NO$_x$ (µg/m^3)	42.6	33.0	21.0	52.0	107.0	98.9%
Relative Humidity (%)	32.9	31.1	18.5	45.3	61.8	99.8%
Temperature (°C)	31.6	32.0	27.1	35.9	40.4	99.8%
Wind Speed (m/s)	1.2	1.1	0.8	1.5	2.0	99.8%

is significant. This test is normally used to avoid the assumption of normality of the data. PM_{10} concentration is not normally distributed and rather is positively (right) skewed as clearly indicated by the fact that mean (157.1 μg/m^3) is significantly greater than the median (123.0 μg/m^3) as shown in Table 1. The test allows us to study the differences in the mean values of the predictions in more details and thus allows us to make comparison between the predictive performances of the models, which is useful for determining the best model out of these five for predicting PM_{10} concentrations in Makkah.

Model Development

Multiple Linear Regression Model (MLRM)

MLRM is probably the most widely used technique for the modelling of air pollution levels; however, the method has several limitations. Hao and Naiman (2007) and Ul-Saufie *et al.* (2012) have explicitly explained its limitations in terms of its inability to extend the response to non-central locations of explanatory variables as well as its inability to meet the model assumptions, such as homoscedasticity. On the other hand, several authors have compared MLRM with other models and have concluded (e.g., Baur *et al.*, 2004) that non-linear and learning machine methods outmatched the linear regression methods. Nevertheless, the method is still in use due to its simplicity and easiness. For modelling PM_{10}, several authors have used linear statistical models and compared them with other potential models (e.g., Mckendry, 2002; Chaloulakou *et al.*, 2003; Grivas and Chaloulakou, 2006). The dependent variable (here PM_{10}) is obtained by additive associations of a number of explanatory/predictor variables. The measured dependent variable (Eq. (1)) and the predicted variable (Eq. (2)) can be expressed as given below:

$$Y = P_o + \sum_{i=1}^{k} P_i X_i + \varepsilon \tag{1}$$

$$\hat{Y} = P_o + \sum_{i=1}^{k} P_i X_i \tag{2}$$

where Y is the dependent variable, X_i is the ith independent variable. P_i s are the regression parameters which are calculated by minimising the sum of square errors through Eq. (3), k is the number of independent variables, and ε is the associated regression error (Martins *et al.*, 2009).

$$P_i = \arg\min \sum_{i=1}^{k} \left(Y_i - \hat{Y}_i \right)^2 \tag{3}$$

Based on the above, the MLRM model used in this study is expressed as follows:

$$Y = P_o + P_1 CO + P_2 SO_2 + P_3 NO_x + P_4 Lag\_PM_{10} + P_5 \overline{U} \tag{4}$$
$$+ P_6 \varnothing + P_7 T + P_8 RH + \varepsilon$$

Generalised Additive Model

GAM is a statistical modelling technique developed by Wahba (1990), and Hastie and Tibshirani (1990) as an extension of the Generalised Linear Models (GLM). The method has become quite popular since then and has been effectively applied in different research areas, such as ecology, aquatic sciences, medical sciences, and environmental sciences (Guisan *et al.*, 2002). The method allows for non-parametric adjustments of the non-linear confounding effects of variables (Dominici *et al.*, 2002). While Carslaw *et al.* (2007) and Westmoreland *et al.* (2007) used GAM for modelling the concentrations of various air pollutants, such as NO_2 and NO_x, Aldrin and Haff (2005), Barmpadimos *et al.* (2011), and Munir *et al.* (2013a) used GAM for modelling the concentration of PM_{10}.

GAMs extend traditional GLMs by replacing linear explanatory variables of the form $\sum_{i=1}^{k} P_i X_i$ with $\sum_{i=1}^{k} f_i(x_i)$ where $f_i(x_i)$ are unspecified nonparametric functions. The estimation procedure for a GAM requires iterative approximation to deduce the optimal estimates, unlike linear regression models, which are fitted by using the weighted least squares method (Wood, 2001). The additive model in a general form can be described as follows:

$$Y = s_1(X_1) + s_2(X_2) + \ldots + s_n(X_n) \tag{5}$$

where Y is the response variable (PM_{10}) and s_i is the unspecified smoothing function which corresponds to an associated explanatory variable (X_i). Using this model, its simplest form has been considered where interaction between variables has not been accounted for.

Based on the above, the GAM model used in this study is as follows:

$$Y = s_1(CO) + s_2(SO_2) + s_3(NO_x) + s_4(LagPM_{10})$$
$$+ s_5(\overline{U}) + s_6(\varnothing) + s_7(T) + s_8(RH) \tag{6}$$

Quantile Regression Model

QRM allows the covariates to have different contribution at different quantiles of the modelled variable distribution and is robust (insensitive) to departures from normality and to skewed tails. Readers are referred to Cade and Noon (2003) Koenker (2005), and Hao and Naiman (2007) for further details on QRM; and to Baur *et al.* (2004), Sousa *et al.* (2008), and Munir *et al.* (2012) for the applicability of QRM to ground-level O_3 concentrations. QRM has also been used by Carslaw *et al.* (2013) on emission data for the first time in order to explain the emission characteristics of petrol and diesel cars. Carslaw *et al.* (2013) explained the advantage of using this method in comparison to other regression methods which consider the mean response only. However, Ul-Saufie *et al.* (2012) have used QRM for modelling PM_{10} concentration and compared it performance with MLRM. Modelling a dependent variable (Y) against a set of independent variables, MLRM regression specifies the conditional mean function, whereas QRM specifies the conditional quantile function.

$$Y = \beta_0^{(p)} + \beta_1^{(p)}(X_1) + \beta_2^{(p)}(X_2) + \ldots + \beta_n^{(p)}(X_n) + \varepsilon_i^{(p)} \tag{7}$$

In Eq. (7) β_0 represents the intercept, β_1 to β_n the slopes (gradients) of the covariates and ε_i the error term. The p shows the pth quantile and its value lies between 0 and 1. Eq. (7) can have numerous quantiles and will require a separate equation for each quantile and therefore will produce numerous coefficients for each variable. This study adopts 10 quantiles and therefore 10 equations will generate the same number of quantile regression coefficients ($\beta_1^{0.1}$, $\beta_1^{0.2}$, …, $\beta_1^{0.9}\beta_1^{0.99}$) for each covariate. The equidistant quantiles make them easier to interpret, however they do not have to be at equal intervals (Hao and Naiman, 2007).

Based on the above, the QRM model used in this study is as follows:

$$Y = \beta_0^{(p)} + \beta_1^{(p)}CO + \beta_2^{(p)}SO_2 + \beta_3^{(p)}NO_x + \beta_4^{(p)}LagPM_{10}$$
$$+ \beta_5^{(p)}\overline{U} + \beta_6^{(p)}\varnothing + \beta_7^{(p)}T + \beta_8^{(p)}RH_2 + \varepsilon_i^{(p)} \qquad (8)$$

QRM makes several predictions, one for each quantile and therefore the metrics used for assessing the model performance can be calculated for each quantile. These metrics are called local metrics, e.g., local RMSE, and local MBE etc. The local metrics cannot be compared with the metrics estimated for the other models, as they have different nature and have different methods of calculation. Therefore, global metrics need to be estimated for QRM to take account of all quantiles and make them comparable with other models. To estimate global metrics for QRM, this study adopts the Amalgamated Quantile Regression Model (AQRM) technique suggested by Baur *et al.* (2004). However, Baur *et al.* (2004) study is limited to only coefficient of determination, whereas this study has extended the concept to several statistical metrics. The first step is to run QRM and determine quantile regression coefficients for all the quantiles used in the model. QRM will normally give numerous predictions according to the number of quantiles. To turn that into one global prediction, the dataset is divided into the same number of subsets as the number of quantiles and then the model for that respective quantile is used to predict PM_{10} concentration which is then re-integrated in such a way that it corresponds to the observed concentrations in the exact order. This ultimately produces a global prediction which takes into account all quantiles.

Boosted Regression Tree Model

Classification and regression trees have offered new methods for analysis and prediction in a number of fields. BRT is one of them and is explored and applied mainly by ecologists (Cappo *et al.*, 2005; Moisen *et al.*, 2006; De'ath, 2007; Elith *et al.*, 2008). The method combines the strengths of two algorithms: regression trees models that relate a response to its predictors by recursive binary splits; and boosting method that combines many simple models to give improved predictive performance (Elith *et al.*, 2008). Carslaw and Taylor (2009) have developed a BRT model for hourly concentrations of NO_x close to the international Heathrow airport in order to understand the influence of different covariates and distinguish the complex interactions between different sources. BRT approach consists of many simpler models, which describe the relationship between the

dependent and independent variables, while the boosting algorithm uses an iterative method for developing a final model, progressively adding trees to the method, and re-weighting the data to address poor prediction cases by previous trees (Leathwick *et al.*, 2006; Carslaw and Taylor, 2009). Learning machine methods, especially the BRT can be time consuming, however what distinguishes BRT from previously explained models is the way it handles non-numerical variables and the way it deals with interactions between variables; unlike the other models, interactions are not predetermined for the BRT method.

Unlike the other models (described above), a number of parameters need to be specified before determining the BRT model that best reduces the error. These are the learning rate (*lr*), interaction depth (*tc*), and trees number (*nt*), as defined by Ridgeway (2012). The learning rate is a shrinkage parameter applied to each tree during the expansion process to shrink the contribution of each tree as it is added to the model, while the interaction depth is the maximum depth of variable interactions controlling the size of the trees. Generally, an interaction depth of 1 implies an additive model where each tree consists of a single node, 2 implies a model with up to 2-way interactions indicating the usage of two nodes in each tree and so on. Based on these two parameters, the tree number required for optimal prediction is determined (Elith *et al.*, 2008) using one of the three known methods: Independent Test set (test); Out-of-Bag (OOB) estimation; and *v*-fold cross validation (CV). In order to determine the optimal number of trees, model fit statistics, such as squared error are calculated using the CV method. The CV method has the advantage of using all the data for both training and validation by repeating the process *v*-times on different combinations of subsamples and calculating the mean performance of the *v*-models. It partitions the data into *v*-subsets where *v*-models are built based on the *'v–1'* subsets and model performance is tested based on the last remaining subset (Ridgeway, 2007). This is repeated *nt* times until the input parameter for the number of trees is reached and the optimum number of trees with the minimal error is subsequently determined. BRT also allows the use of a random component to improve the prediction performance; this is achieved using a bag fraction of the training dataset, randomly selected to fit each consequent tree (Friedman, 2002). A Gaussian distribution is assumed and a bag fraction of 0.5 is used for both models. A 1-way BRT model uses an interaction depth value of 1, while a 2-way BRT model uses 2 as an interaction depth to account for interactions between variables. As for the selection of *lr* and *nt* parameters, many simulations with different combination of *lr* (0.1 to 0.001) and *nt* (1,000 and 50,000) have been tested, especially for their least square errors, while considering that an *lr* between 0.01 and 0.001, and *nt* above 3,000 are recommended (Ridgeway, 2007). Using 10-fold cross validation, the two models with the optimal parameters' values are developed using the training data only (11 months data) and then applied on the testing dataset for measuring its predictive performance and comparing them with the other models.

MLRM, GAM, QRM and 1-way and 2-way BRT models

are fitted in statistical software R programming language (R Development Core Team, 2012) and its packages 'gbm' or 'generalised boosted machine' package (Ridgeway, 2012), 'mgcv' package (Wood, 2011) and the 'quantreg' package (Koenker, 2013) for BRT, GAM and QRM, respectively. No specialised package is required for fitting MLRM and rather is fitted by the core installation packages in R.

RESULTS AND DISCUSSION

Models Evaluation

Using the testing dataset for each of the developed models PM_{10} concentrations were predicted, a summary of the observed and predicted PM_{10} concentrations for the month of June is presented in Fig. 1. The Boxplot summarises PM_{10} model predictions. While the bottom and top of the box show the 25th and 75th percentiles, respectively, the whiskers present the 1.5 times the inter-quartile range of the data. Generally, the Boxplot shows lowest variation between results from the QRM model and the observed PM_{10} concentrations.

The mean observed PM_{10} concentration for the testing dataset is 224.3 µg/m³. All five models approach this value in a numerical range of 3.2 to 43.9 µg/m³. The difference between average predicted and observed values is referred to by MBE, indicating whether a model under-predicts or over-predicts the observations. QRM showed lowest difference between the mean predicted and observed values. QRM under-predicted the mean value by only 3.2 µg/m³, while the other models largely under predicted the mean observed values: MLRM by 31.1 µg/m³; 2-way BRT by 41.1 µg/m³; GAM by 41.7 µg/m³; 1-way BRT by 43.9 µg/m³.

The non-parametric Wilcoxon test has been performed to check if there is a significant difference between the observed

and the modelled mean PM_{10} concentrations of each model. P-values lower than 0.01 are obtained for GAM, BRT1, and BRT2 models. This shows that these three models show a significant difference between mean values of the observed and modelled mean PM_{10} concentrations at 1% significant level (p-value = 0.01). However, the test on both MLRM and QRM yielded p-values greater than 0.01, showing an insignificant difference between the observed and predicted PM_{10} concentration levels.

SD (as illustrated in both the Boxplot and Table 2) for the observed PM_{10} concentrations (141.3 µg/m³) is greater than that for the predicted PM_{10} concentrations by all the models, except QRM (160.8 µg/m³), which has an SD higher than that of the observed values. Nevertheless, the absolute difference in SD is lowest for QRM (19 µg/m³), for the other models it ranges from 60 to 73 µg/m³.

The scatter plots of modelled versus observed values are displayed in Fig. 2. A 1:1 line is added on each graph to facilitate the comparison to the ideal model, and a factor of two scatter is indicated by the dashed 1:2 and 2:1 lines (Chaloulakou *et al.*, 2003; Barmpadimos *et al.*, 2011; Paschalidou *et al.*, 2011). The Pearson correlation coefficients are also added on the graphs to measure the strength of the linear relationship between the observed and modelled values (Carslaw, 2011). The scatter plots give some insight into the model performance during high PM_{10} concentration levels. As shown in Fig. 2, all the models, except QRM seem to under-predict higher values. Fig. 3 plots both observed and modeled PM_{10} concentrations of each model versus time and depicts the under-prediction of PM_{10} at higher values.

Table 2 summarises quantitatively the performance of the models in terms of MBE, MAE, RMSE, FACT2, and IA. As indicated by the error indices, QRM model seems to outclass the other linear and non-linear models. The MAE
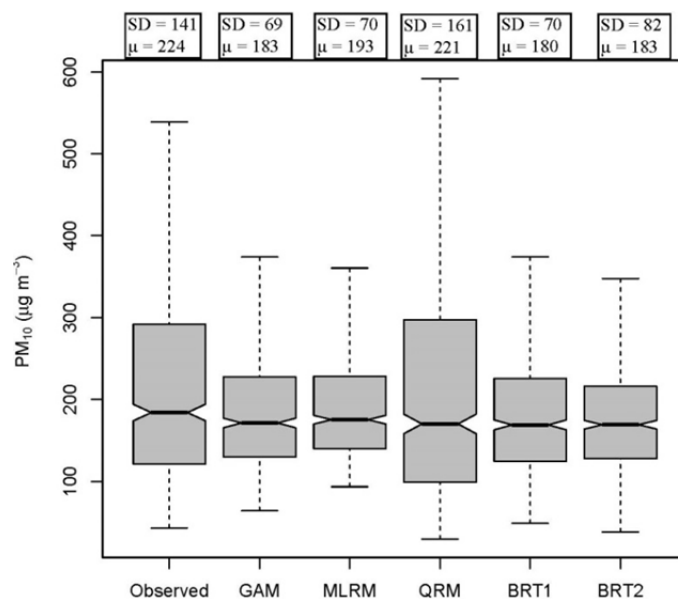
Fig. 1. Boxplot summarizing PM_{10} model predictions on the testing dataset; the bottom and top of the box show the 25th and 75th percentiles respectively, whiskers present the 1.5 times the interquartile range of the data, and notches on either side of the median gives an estimate of the 95% confidence interval of the median. SD is the standard deviation in µg/m³ and µ is the mean in µg/m³.

Table 2. Model evaluation parameters for each of the five models under study.

Parameter	GAM	MLRM	QRM	BRT1	BRT2
Mean Bias Error (MBE)	–39.9	–29.3	–1.4	–43.9	–41.1
Mean Absolute Error (MAE)	74.3	80.0	61.0	75.6	80.4
Mean Absolute Percentage Error (MAPE - %)	33.1	35.7	27.2	33.7	35.8
Root Mean Squared Error (RMSE)	120.1	123.8	95.6	121.1	125.6
Relative RMSE - %	53.5	55.2	42.6	54.0	56.0
FACT2	0.89	0.87	0.96	0.88	0.87
Correlation Coefficient (R)	0.60	0.51	0.81	0.61	0.54
Index of Agreement (IA)	0.65	0.61	0.89	0.66	0.66
Standard Deviation (SD)	68.8	70.2	160.8	69.8	81.5

* Parameter calculation equations:

$$MBE = \frac{1}{n}\sum_{i=1}^{N} M_i - Q_i \; ;$$

$$MAE = \frac{1}{n}\sum_{i=1}^{N} |M_i - Q_i| \; ;$$

$$MAPE = 100 \times \left(\frac{MAE}{\mu_{observed}}\right)$$

$$RMBE = \sqrt{\frac{1}{n}\sum_{i=1}^{N} \left(M_i - Q_i\right)^2} \; ;$$

$$Relative\ RMSE = 100 \times \left(\frac{RMSE}{\mu_{observed}}\right)$$

FACT2 is the fraction of model predictions that satisfy: $0.5 \le M_i/Q_i \le 2.0$;

$$r = \frac{1}{n-1}\sum_{i=1}^{n} \left(\frac{M_i - \bar{M}}{\sigma_M}\right)\left(\frac{O_i - \bar{O}}{\sigma_O}\right);$$

$$IA = 1 - \frac{\sum_{i=1}^{N}\left(M_i - O_i\right)^2}{\sum_{i=1}^{N}\left(|M_i - \bar{O}| + |O_i - \bar{O}|\right)^2}$$

for the QRM model is 27.2% of the arithmetic mean (224.3 µg/m³) of the observed concentration, whereas that of the MLRM, GAM, 1-way BRT, and 2-way BRT models are 35.6%, 33.1%, 33.7%, and 35.8%, respectively. This is referred to in Table 2 as the Mean Absolute Percentage Error (MAPE). This indicates that QRM has the lowest deviation relative to the mean value in comparison with other models.

As a result of the power term, RMSE is more appropriate to illustrate the presence of significant under or over-predictions. Similar to MAE results, QRM performs better in terms of RMSE. RMSE values (µg/m³) for MLRM, GAM, QRM, 1-way BRT, and 2-way BRT models are 124, 120, 96, 121, and 126, respectively. The percentages of the RMSE over the mean value of the observations for the test set, which is referred to as Relative RMSE, are 55% for MLRM, 54% for GAM, 43% for QRM, 54% for the 1-way BRT, and 56% for the 2-way BRT model. In addition, percent difference between RMSE for QRM and the other models is calculated. RMSE of QRM is 23%, 20%, 21%, and 24% lower than that of MLRM, GAM, 1-way BRT model, and 2-way BRT model, respectively. This shows that

QRM prediction is significantly better that the predictions of the other models.

Scatter plots in Fig. 2 clearly show that most of the QRM predictions lie within a factor of two of the observed values. Calculation of FACT2 reveals that 96% of the QRM predictions are within a factor of two of the observed PM$_{10}$ concentrations, whereas for the rest of the models FACT2 values ranged from 87% to 89%.

Based on the Pearson correlation coefficient (R), QRM with R value of 0.81 topped the other models, exhibiting R values less than 0.61. Furthermore, in terms of the dimensionless index (IA) QRM has shown the same patterns of superiority over the other models. QRM has an IA value of 0.89, whereas MLRM, GAM, BRT1, and BRT2 have IA values of 0.61, 0.65, 0.66, and 0.66, respectively.

It is important to mention that in 2012 the Hajj event took place during the last week in October, which was used in the training data and not in the testing data. However, due to effective traffic management, which include blocking of several roads near Al-Haram, and banning of outside-traffic to enter the Makkah City during the Hajj (Park and Ride services were available), the PM levels in the month
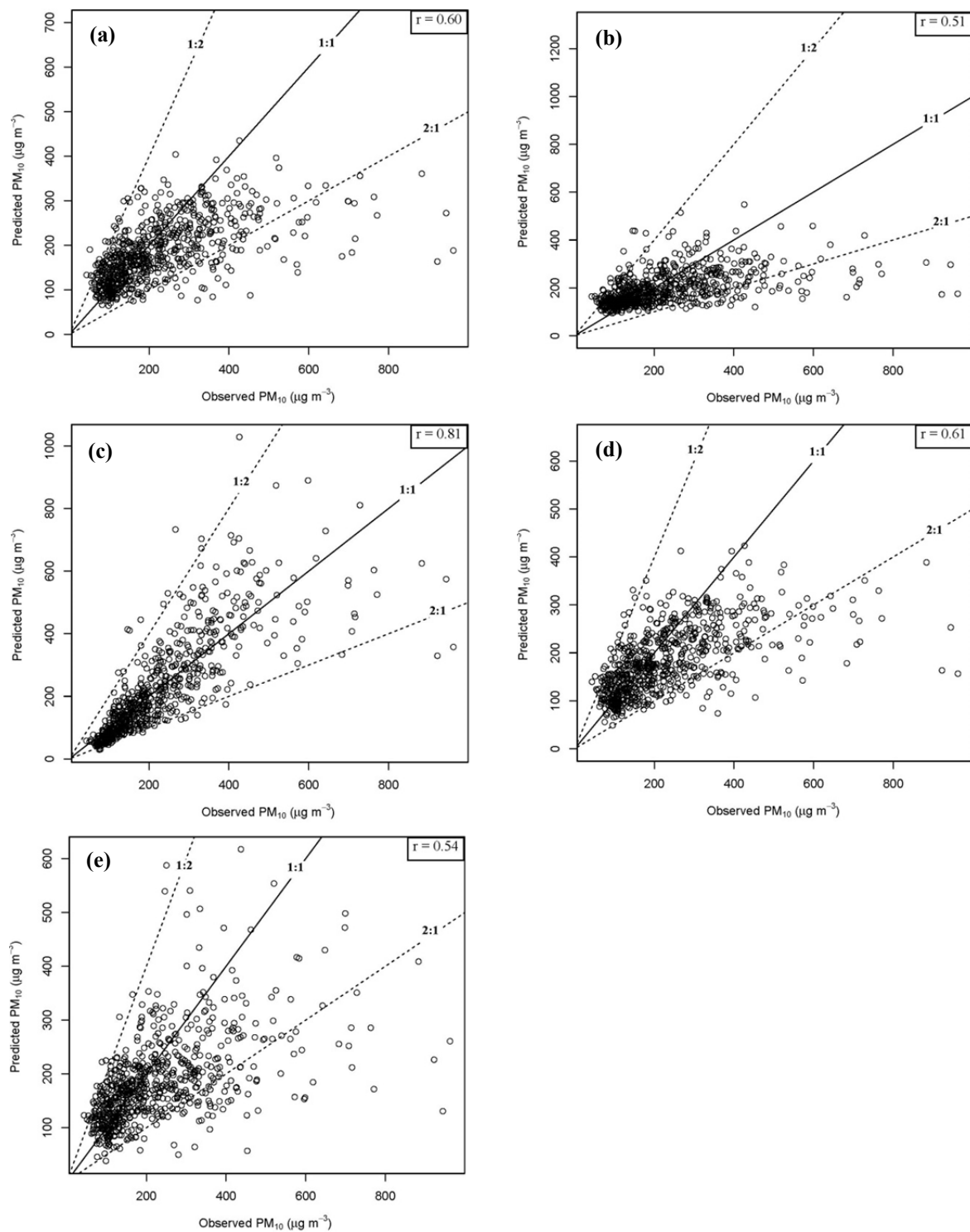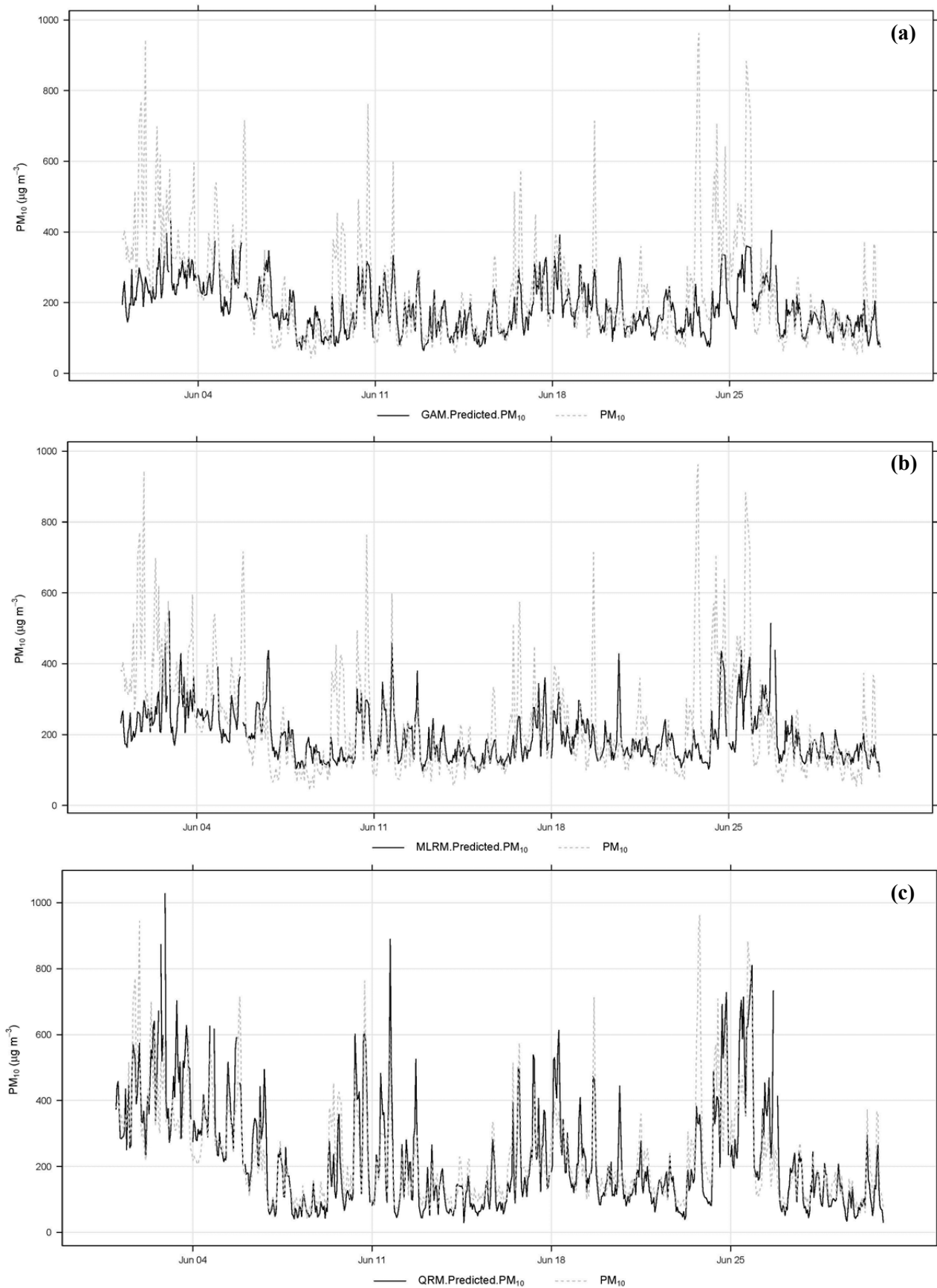
Fig. 2. Observed versus modelled PM$_{10}$ concentration for the validation period, month of June, by each of the models (a) GAM, (b) MLRM, (c) QRM, (d) 1-way BRT model, (e) 2-way BRT model.

of October were not significantly higher than the other months. Therefore, the models, particularly QRM was able to capture the variation in PM$_{10}$ caused by the Hajj season. Moreover, despite the fact that the 2-way BRT model accounts for interactions between independent variables, the predictive performance of the model did not show any improvement in comparison to the 1-way BRT model as

shown by calculated statistical metrics. This shows that the main effects' modelling of PM$_{10}$ concentrations performs better compared to modelling with interaction, predetermined in this case.

The three models: MLRM, GAM, and BRT (both 1-way and 2-way) have yielded very similar results in terms of statistical metrics that indicate their predictive performance.

Fig. 3. Time series comparison of observed versus modelled PM$_{10}$ concentration levels of (a) GAM, (b) MLRM, (c) QRM, (d) 1-way BRT, and (e) 2-way BRT on the testing dataset.
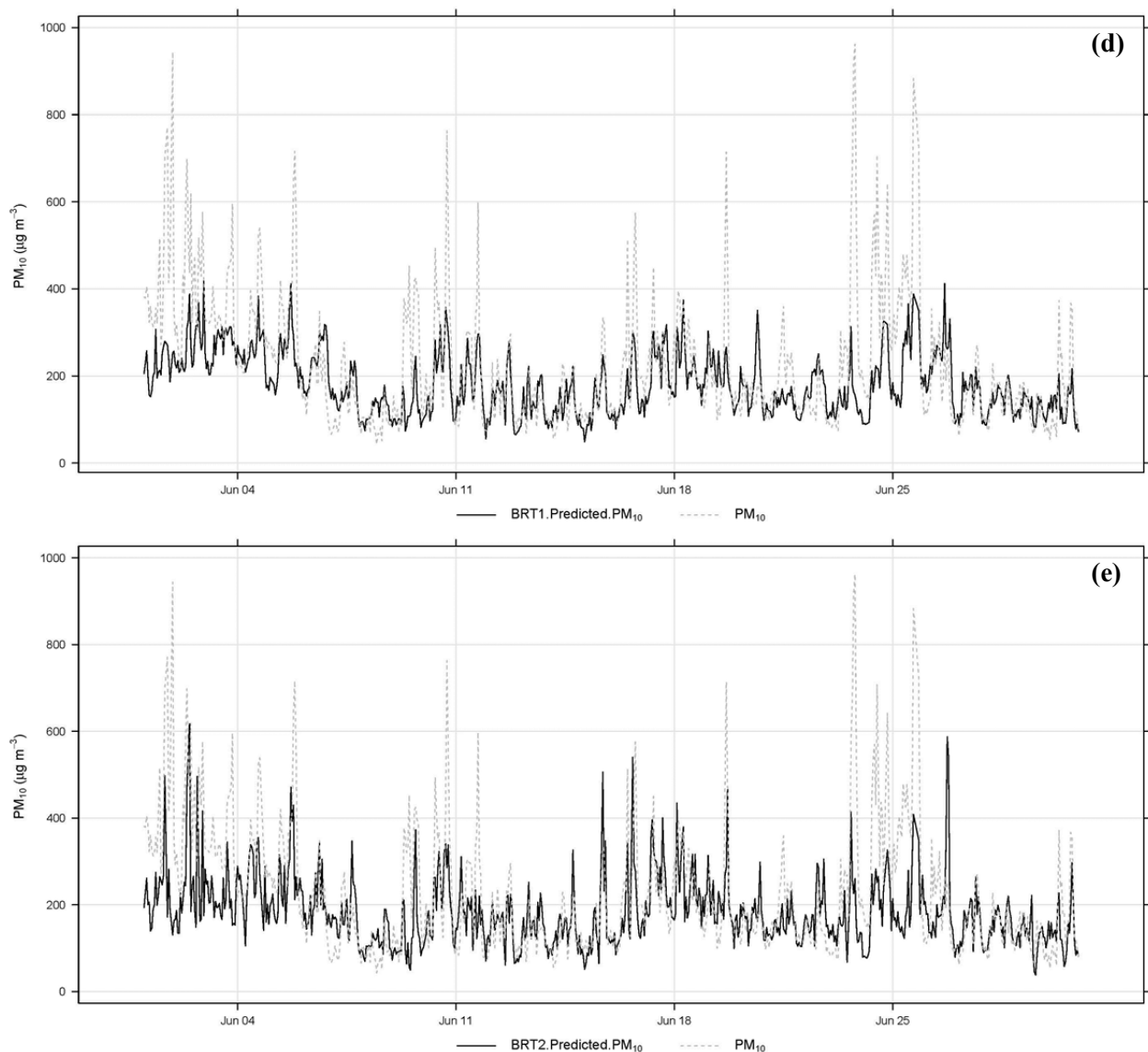
Fig. 3. *(continued).*

The above results indicate that QRM, which estimates the conditional quantiles of the PM_{10} distribution, has outclassed the rest of the three models. First, while MLRM captures the mean value of PM_{10} concentration levels, only QRM captures both the mean value and variability of PM_{10} concentration levels. Chaloulakou *et al.* (2003) using NN and multiple regression models identified the capturing of the mean value as a typical observation for the MLRM model, since it attempts to approximate an average behavior and thus fails to capture the variations in the response variable. Also, BRT which applies boosting to regression problems attempts to estimate the mean of the response. According to Zheng (2012), this is considered a weakness compared to QRM since BRT does not use quantiles and this is why Quantile Boosted Regressions have been proposed by researchers which allows the application of boosting to estimate the response at different quantiles.

Hao and Naiman (2007) and Ul-Saufie *et al.* (2012) explained that the ability of QRM to capture both the mean and variation of the response variable is related to its ability of examining the entire distribution of the variable rather than a single measure of the central tendency of its distribution. Cade and Noon (2003) have also stated that QRM estimates multiple rates of change or slopes from the minimum to the maximum response and thus provides a more complete picture of the relationships between response and explanatory variables, unlike other regression methods, such as MLRM, GAM, and BRT that consider only the mean response and ultimately yield a poor predictive performance.

Comparison with Other Studies

As stated earlier, although the models under considerations have been used for modelling different air pollutants, for modelling PM_{10} researchers have focused mainly on GAM (e.g., Aldrin and Haff, 2005; Barmpadimos *et al.*, 2011; Munir *et al.*, 2013a) and MLRM. Studies on the application of QRM for modelling PM_{10} concentrations are very rare (Ul-Saufie *et al.*, 2012), while the application of BRT to

PM_{10} concentration does not exist yet. However, other models, such as NN have been applied for analysing PM_{10} levels and their results are compared with different models, especially MLRM (Chaloulakou *et al.,* 2003; Grivas and Chaloulakou, 2006; Ul-Saufie *et al.*, 2011) and statistical parameters such correlation coefficients, bias errors, and indices of agreement were calculated.

Although the above authors attempted to predict PM_{10} concentrations, they had different forecast targets and, therefore, direct comparisons with this paper cannot be made. However, they are briefly stated here as a preliminary tool for comparison and in order to provide a comprehensive analysis of results. Barmpadimos *et al.* (2011) used GAM to study the influence of meteorology on PM_{10} and calculated FACT2 and R for each season of the year. Average value of 0.93 with a minimum of 0.8 was observed for the FACT2 and an average value of 0.7 was observed for the R coefficient. Aldrin and Haff (2005) estimated a GAM model on log-scale on PM_{10} concentrations for four different stations. The reported R coefficient varied between 0.69 and 0.84. Munir *et al.* (2013a) have also developed a GAM model and compared the predicted and the observed PM_{10} concentrations on an independent testing dataset. They have reported 0.72, 84, and 0.88 as the values of R coefficient, RMSE, and FACT2, respectively. The value of FACT2 was similar to this study, however the values of R and RMSE were different.

Alternatively, Chaloulakou *et al.* (2003) have developed an MLRM and calculated various statistical metrics using an independent testing set of data. The reported values of MBE, MAE, RMSE, and R were –0.39, 14.07, 18.37, and 0.77, respectively. They used lag_PM_{10} as one of the explanatory variables.

CONCLUSION

In this study, five different models were used for the prediction of hourly PM_{10} concentration levels in the City of Makkah, KSA. The predictors used in each of the five models are meteorological parameters, surrogate variables to traffic volumes, and lagged PM_{10} concentration. QRM clearly outmatched the rest of the models: MLRM, GAM, and the 1-way and 2-way BRTs. By investigating the comparative performance of the models, significant differences were observed between the mean observed and predicted PM_{10} concentrations for all the models, except MLRM and QRM. However, QRM topped the other models, including MLRM, in capturing the variability of PM_{10}. Various statistical parameters and error indices, calculated to assess the performance of the models, showed that QRM was able to predict hourly PM_{10} concentrations with minimal errors. The difference between QRM and the other models is due to the approximation behavior. While the other models attempt to approximate an average behavior, QRM approximate is based on the number of quantiles adopted in the model. This suggests that the ability of QRM to capture the contributions of covariates at different quantiles produces better prediction, compared to the procedures where a single central tendency is considered.

It should be noted that the models developed for PM_{10} did not directly use traffic characteristics (traffic flow, speed and composition) as predictors, rather surrogate variables were used. Also, the comparison of models was limited to only one monitoring site in Makkah and for a short period of time. This can limit the performance and comparison of the models. Accordingly, further investigation using data from different monitoring sites over a longer period of time is suggested. Furthermore, traffic characteristics data (e.g., traffic flow, vehicles speed and fleet composition), which can be collected using inductive loops, are also recommended for future investigations to be directly included in the models. Traffic characteristics are required to estimate air pollutant emissions and may help in more accurate prediction of PM_{10} concentrations and therefore may further improve the model comparison in the City of Makkah. Alternatively, when using the best model for predicting future PM_{10} concentrations, it might require data on forecasted independent variables, such as weather conditions. Currently, weather forecasts are limited to international weather channels. This might add uncertainty to model the models' outcomes. This is also highlighted in Paschalidou *et al.* (2011), who reported that the use of accurate weather forecast stations is important when the models are set to be used for future air quality predictions.

ACKNOWLEDGEMENT

REFERENCES

Aldrin, M. and Haff, I. (2005). Generalised Additive Modelling of Air Pollution, Traffic Volume and Meteorology. *Atmos. Environ.* 39: 2145–2155.

AQEG (2005). Air Quality Expert Group, Particulate Matter in the United Kingdom, Published by the Department for the Environment, Food and Rural Affairs of the United Kingdom.

Barmpadimos, I., Hueglin, C., Keller, J., Henne, S. and Prevot, A. (2011). Influence of Meteorology on PM_{10} Trends and Variability in Switzerland from 1991 to 2008. *Atmos. Chem. Phys*. 11: 1813–1835.

Baur, D., Saisana, M. and Schulze, N. (2004). Modelling the Effects of Meteorological Variables on Ozone Concentration–A Quantile Regression Approach. *Atmos. Environ*. 38: 4689–4699.

Braniš, M. and Větvička, J. (2010). PM_{10}, Ambient Temperature and Relative Humidity during the XXIX Summer Olympic Games in Beijing: Were the Athletes at Risk? *Aerosol Air Qual. Res*. 10: 102–110.

Cade, B.S. and Noon, B.R. (2003). A Gentle Introduction to Quantile Regression for Ecologists. *Front. Ecol. Environ.* 1: 412–420.

Cappo, M., De'ath, G., Boyle, S., Aumend, J., Olbrich, R., Hoedt, F., Perna, C. and Brunskill, G. (2005). Development of a Robust Classifier of Freshwater Residence in Barramundi (Latescalcarifer) Life Histories Using Elemental Ratios in Scales and Boosted Regression Trees. *Mar. Freshwater Res.* 56: 713–723.

Carslaw, D. (2011). DEFRA Urban Model Evaluation Analysis – Phase 1.

Carslaw, D.C., Beevers, S.D. and Tate, J.E. (2007). Modelling and Assessing Trends in Traffic-related Emissions Using a Generalized Additive Modelling Approach. *Atmos. Environ.* 41: 5289–5299.

Carslaw, D.C. and Taylor, P.J. (2009). Analysis of Air Pollution Data at a Mixed Source Location Using Boosted Regression Trees. *Atmos. Environ.* 43: 3563–3570.

Carslaw, D.C., Williams, M.L., Tate, J.E. and Beevers, S.D. (2013). The Importance of High Vehicle Power for Passenger Car Emissions. *Atmos. Environ.* 68: 8–16.

Central Department of Statistics and Information (2010). Results of Hajj Statistics.

Chaloulakou, A., Grivas, G. and Spyrellis, N. (2003). Neural Network and Multiple Regression Models for PM_{10} Prediction in Athens: A Comparative Assessment. *J. Air Waste Manage. Assoc.* 53: 1183–1190.

De'ath, G. (2007). Boosted Trees for Ecological Modelling and Prediction. *Ecology* 88: 243–251.

Derwent, D., Fraser, A., Abbott, J., Jenkin, M., Willis, P., and Murrells, T. (2010). Evaluating the Performance of Air Quality Models. DEFRA Report Issue 3. 7: 81.

Dominici, F., McDermott, A., Zeger, S. and Samet, J. (2002). On the Use of Generalized Additive Models in Time-Series Studies of Air Pollution and Health. *Am. J. Epidemiol.* 156: 193–203.

ElAssouli, S. (2011). Airborne Particulate Matter (PM_{10}) Composition and its Genotoxicity at Two Pilgrimage Sites in Makkah, Saudi Arabia. *J. Environ. Chem. Ecotoxicol.* 3: 93–102.

Elith, J., Leathwick, J.R. and Hastie, T. (2008). A Working Guide to Boosted Regression Trees. *J. Anim. Ecol.* 77: 802–813.

Friedman, J.H. (2002). Stochastic Gradient Boosting. *Comput. Stat. Data Anal.* 38: 367–378.

Godish, T. (1997). *Air Quality* (Third Edition), Lewis Publishers, New York.

Grivas G. and Chaloulakou A. (2006). Artificial Neural Network Models for Prediction of PM_{10} Hourly Concentrations, in the Greater Area of Athens, Greece. *Atmos. Environ.* 40: 1216–1229.

Guisan, A., Edwards, T.C. Jr. and Hastie, T. (2002). Generalized Linear and Generalized Additive Models in Studies of Species Distributions: Setting the Scene. *Ecol. Modell.* 157: 89–100.

Habeebullah, T. (2013). Health Impacts of PM_{10} Using AirQ2.2.3 Model in Makkah. *J. Basic Appl. Sci.* 9: 259–268.

Hao, L. and Naiman, D.Q. (2007). *Quantile Regression: Series-Quantitative Applications in the Social Sciences*, Sage Publications, 7: 149.

Harrison, R., Deacon, A., Jones, M. and Appleby, R. (1997). Sources and Processes Affecting Concentrations of PM_{10} and $PM_{2.5}$ Particulate Matter in Birmingham (UK). *Atmos. Environ.* 31: 4103–4117.

Hastie, T. and Tibshirani, R. (1990). *Generalised Additive Models*, Chapman and Hall.

Jaecker-Voirol, A. and Pelt, P. (2000). PM_{10} Emission Inventory in Ile de France for Transport and Industrial Sources: PM_{10} Re-suspension, a Key Factor for Air Quality. *Environ. Modell. Softw.* 15: 575–581.

Kim, B., Teffera, S. and Zeldin, M. (2000). Characterization of $PM_{2.5}$ and PM_{10} in the South Coast Air Basin of Southern California: Part 1-Spatial Variation. *J. Air Waste Manage. Assoc.* 50: 2034–2044.

Koenker, R. (2005). *Quantile Regression*, Economic Society Monographs No. 38, Cambridge University Press, ISBN 0-521-608227-9.

Koenker, R. (2013). Quantreg: Quantile Regression. R Package Version 4.97. http://CRAN.R-project.org/package=quantreg.

Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R. and Cawley, G. (2003). Extensive Evaluation of Neural Network Models for the Prediction of NO_2 and PM_{10} Concentrations, Compared with a Deterministic Modelling System and Measurements in Central Helsinki. *Atmos. Environ.* 37: 4539–4550.

Leathwick, J.R., Elith, J., Francis, M.P., Hastie, T. and Taylor, P. (2006). Variation in Demersal Fish Species Richness in the Oceans Surrounding New Zealand: an Analysis Using Boosted Regression Trees. *Mar. Ecol.* 321: 267–281.

Martins, F., Pires, J. and Sousa, S. (2009). Statistical Models for Predicting Ozone and PM_{10} Concentrations, In *Modelling of Pollutants in Complex Environmental Systems*, Hanrahan, G. (Eds.), p. 277–290.

McKendry, I. (2002). Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM_{10} and $PM_{2.5}$) Forecasting. *J. Air Waste Manage. Assoc.* 52: 1096–1101.

Moisen, G.G., Freeman, E.A., Blackard, J.A., Frescino, T.S., Zimmermann, N.E. and Edwards, T.C. (2006). Predicting Tree Species Presence and Basal area in Utah: A Comparison of Stochastic Gradient Boosting, Generalized Additive Models, and Tree-Based Methods. *Ecol. Modell.* 199: 176–187.

Munir, S., Chen, H. and Ropkins, K. (2012). Modelling the Impact of Road Traffic on Ground Level Ozone Concentration using a Quantile Regression Approach. *Atmos. Environ.* 60: 283–291.

Munir, S., Habeebullah, T., Seroji, A., Morsy, E., Mohammed, A., Saud, W., Abdou, A. and Awad, A. (2013a). Modeling Particulate Matter Concentrations in Makkah, Applying a Statistical Modeling Approach. *Aerosol Air Qual. Res.* 13: 901–910.

Munir, S., Habeebullah, T., Seroji, A.R., Gabr, S.S., Mohammed, A.M.F. and Morsy, E.A. (2013b). Quantifying Temporal Trends of Atmospheric Pollutants in Makkah

(1997-2012). *Atmos. Environ.* 77: 647–655.

Pai, T., Ho, C., Chen, S., Lo, H., Sung, P., Lin, S., Lai, W., Tseng, S., Ciou, S., Kuo, J. and Kao, J. (2011). Using Seven Types of GM (1, 1) Model to Forecast Hourly Particulate Matter Concentration in Banciao City of Taiwan. *Water Air Soil Pollut.* 217: 25–33.

Papanastasiou, D., Melas, D. and Kioutsioukis, I. (2007). Development and Assessment of Neural Network and Multiple Regression Models in Order to Predict PM_{10} Levels in a Medium-sized Mediterranean City. *Water Air Soil Pollut.* 182: 325–334.

Paschalidou, A., Karakitsios, S., Kleanthous, S. and Kassomenos, P. (2011). Forecasting Hourly PM_{10} Concentration in Cyprus through Artificial Neural Networks and Multiple Regression Models: Implications to Local Environmental Management. *Environ. Sci. Pollut. Res.* 18: 316–327.

Pires, J., Martins, F., Sousa, S., Alvim-Ferraz, M. and Pereira, M. (2008). Prediction of the Daily Mean PM_{10} Concentrations Using Linear Models. *Am. J. Environ. Sci.* 4: 445–453.

Pisoni, E. and Volta, M. (2009). Modeling Pareto Efficient PM_{10} Control Policies in Northern Italy to Reduce Health Effects. *Atmos. Environ.* 43: 3243–3248.

Pont, V. and Fontan, J. (2000). Comparison between Weekend and Weekday Ozone Concentration in Large Cities in France. *Atmos. Environ.* 35: 1527–1535.

R Development Core Team (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Ridgeway, G. (2007). Generalized Boosted Models: A Guide to the gbm Package.

Ridgeway, G. (2012). Gbm: Generalized Boosted Regression Models. R Package Version 1.6-3.2. http://CRAN.R-project.org/package=gbm.

Sawant A., Na, K., Zhu X. and Cocker D. (2004). Chemical Characterization of Outdoor $PM_{2.5}$ and Gas-Phase Compounds in Mira Loma, California. *Atmos. Environ.* 38: 5517–5528.

Seroji, A. (2011). Particulates in the Atmosphere of Makkah and Mina Valley during the Ramadan and Hajj Seasons of 2004 and 2005, In *Air Pollution XIX*, Brebbia, C.A., Longhurst, J.W.S. and Popov, V. (Eds.), Wessex Institute of Technology, UK.

Seroji, A. (2012). Assessment of PM_{10}, CO and Noise Levels at the Central Area of Makkah during Hajj Season 1429H (2008), In *Environmental Impact*, Brebbia, C.A. and Chon, T.S. (Eds.), Wessex Institute of Technology, UK.

Sousa, S.I.V., Pires, J.C.M., Mrtins, F.G., Pereira, M.C. and Alvim-Ferraz, M.C.M. (2008). Potentialities of Quantile Regression to Predict Ozone Concentrations. *Environmetrics* 20: 147–158.

Ul-Saufie, A., Yahya, A., Ramli, N. and Hamid, H. (2011). Comparison between Multiple Linear Regression and Feed Forward Back propagation Neural Network Models for Predicting PM_{10} Concentration Level Based on Gaseous and Meteorological Parameters. *Int. J. Appl. Sci. Technol.* 1: 42–49.

Ul-Saufie, A., Yahya, A., Ramli, N. and Hamid, H. (2012). Future PM_{10} Concentration Prediction Using Quantile Regression Models, International Conference on Environmental and Agriculture Engineering, IACSIT Press, Singapore, Vol. 37.

Wahba, G. (1990). *Spline Models for Observational Data*, SIAM, Philadelphia.

Westmoreland, E.M., Carslaw, N., Carslaw, D.C., Gillah, A. and Bates, E. (2007). Analysis of Air Quality within a Street Canyon using Statistical and Dispersion Modelling Techniques. *Atmos. Environ.* 41: 9195–9205.

Wood, S. (2001). Mgcv: GAMs and Generalised Ridge Regression for R. *R News* 1: 20–25.

Wood, S. (2011). Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semi Parametric Generalized Linear Models. *J. R. Stat. Soc. B*. 73: 3–36.

Zheng, S. (2003). QBoost: Predicting Quantiles with Boosting for Regression and Binary Classification. *Expert Syst. Appl.* 39: 1687–1697.