Journal of Sustainability Science and Management *eISSN: 2672-7226*
Volume 14 *Number* 4, August 2019: 152-190

AIRBORNE PARTICULATE MATTER RESEARCH: A REVIEW OF FORECASTING METHODS

HERRINI MOHD PAUZI AND LAZIM ABDULLAH*

School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia.

**Corresponding author: lazim_m@umt.edu.my*

Abstract: Previous studies indicate that the main cause of air quality deterioration is the concentration of particulate matter up to 10 µm in size (PM10). So far, however, there has been little discussion on forecasting methods that have been used in PM10 research. This paper provides a comprehensive review of previous methodologies and approaches used in forecasting PM10 concentration. The reviewed literature are analysed based on the frequency of single and hybrid approaches. These approaches comprise multiple forecasting models which can be categorized as statistical and deterministic models. Ee determined the prevalent variables that are related to PM10 concentration. Hybrid approaches are exhibited as preferable methods based on the number of the reviewed publications. In addition, artificial neural network model hadthe highest number of applications both for its single and hybrid approaches. Meteorological variables are the highest variable that influenced the PM10 concentration. This study provides discussions on the future directions of the research based on the identification of amicable approaches and the influential factors contributing to PM10 concentration.

Keywords: Air quality, forecasting, particulate matter, hybrid model, meteorological variables, artificial neural network .

Introduction

It is becoming increasingly difficult to ignore the prevalence of various environmental impacts of air pollution as a result of uncontrollable urbanization and rapid pace of industrialization. Deterioration of air quality has now become a new challenge to those who are directly involved in air quality management. The air quality management includes measuring concentration of air pollutants, preparing and planning for further actions while enforcing environmental laws to mitigate the atmosphere exceedance of air pollutants. Air pollution is occurring when there exists a composition of a certain amount of compounds, which included ozone, nitrogen oxides, sulphur dioxide, carbon monoxide and, last but not least, particulate matter (Shang, 2017).

The airborne particulate matter with diameters up to 10 µm (PM10) is therefore an important compound in air quality and is believed that it could be harmful and has substantial influence on theenvironment and extreme weather and also climate change (Dumitrache *et al*., 2016; Juneng *et al.* 2011; Luo *et al*., 2018). In fact, a number of authors haveclaimed that the concentration of PM10 in the atmosphere is significantly related to the increasing numberof health consequences (Kaufman *et al*., 2002; Chaloulakou *et al*., 2003; Martuzzi *et al*., 2005; Ho, 2017). These tiny particles are able to penetrate respiratory systems of human and could disrupt the biological system (Perez and Reyes, 2002).Furthermore, World Health Organization (WHO), (2013) stated that scientific research significantly relates deterioration of human health to the effect of over exposure to these particles for both short-term and long-term(World Health Organization, 2013). These issues indicate that there is an urgent need for accurate and precise information of PM10 concentration in order to provide proper planning for air quality sustainability. Brunelli *et al*., (2007) stated that forecasting model for air pollutants at short lead-time should provide a potential tool to plan a health warning system.

Other scholars supported the indication that, it is necessary to monitor air quality in real time and to predict its trends prior to PM10 pollution (Chen *et al.,*2013; Luo *et al*., 2018; De Mattos Neto *et al.,*2014). While the behavior of PM10 concentration is always uncertain, ambiguous and volatile, many widely accepted that the suspended particles are predictable. Feng *et al*., (2015) eloquently classified the forecasting methods based on related literature into two main categories; deterministic and statistical methods. Statistical methods comprises various linear and non-linear methods to forecast PM10 concentration, e.g: artificial neural network (ANN), support vector machine (SVM), multiple linear regression (MLR) and chemistry transport model (CTM). Apparently, these two methods have different characteristics with different requirements. Chen *et al.*, (2013), Baklanov *et al.,* (2008) and Kim *et al.*, (2010) defined deterministic method as a tool that is able to simulate the mechanism of discharge, diffusion, accumulation and transmission of a pollutant with a limited number of monitoring stations in terms of animated figure. Meanwhile, the statistical models have always been favorable models due to their simplicity, eventhough they demand a large amount of historical data (Feng *et al*., 2015; Wang *et al*., 2018)

However, these two methods suffer from their own disadvantages. As a result, deterministic methods tend to yield imprecise simulation and require high computational time (Niu *et al*., 2016;Vautard *et al.,* 2007; Stern *et al*., 2008). Yet, most of the statistical approachesrequire abundant past data values to interpret inter-relationship that associates between experimental variables. Niska *et al.*, (2005) stated that, statistical approaches lack flexibilities to vary their representation in other regions with different meteorological conditions. Recently, there has been a resurgence of interest in hybrid forecasting model due to its successes in many forecasting areas including atmospheric scienceand examples include hybrid forecasting models in the area of agriculture by Jana *et al.*, (2016) and Xiong *et al*., (2017),electric power system foreceasting (Wang *et al.,* 2018), photovoltaic solar power forecasting (Eseye *et al.,* 2018), etc.The main goal of hybridzing is to curb deficiency exhibit from the selected models and exploit each benefit to produce an improvised forecasting model. The hybridization is made by coupling deterministic methods, intelligent and traditional statistical methods to enhance the precision of forecasting model (Luo *et al*.,2018) as applied by Konovalov *et al*., (2009) and Song *et al.*, (2015) in their research.

Being one of the critical compounds in air quality and environmental sustainability, particularly its threat to human health, PM10 related researches have been investigated and reviewed by many scholars. Afroz *et al*., (2003) and Costa *et al*., (2014), for example, presented their review papers with discussion on the effect of air pollution and its threatto human health. They reviewed and discussed the trans-boundary topic with the general view on the impacts of the high concentration of PM10 to human health and the environment. It is clearly indicated that these discussions focused on the issues of the PM10 airborne as the main factor of air quality deterioration. There has been a growing interest in developing robust and potential forecasting models for PM10 concentration, but the review of related forecasting models still remains limited. A difference is seen in this and the earlier reviews mentioned where the impacts on human living were the main concern. This paper comprehensively reviews forecasting methods that are normally used to investigate PM10. Besides, to the best of the authors' knowledge, there has been no attempt to review the forecasting related models and its variables in PM10 research, although a handful of researchers have published global reviews on PM10 (Vahlsing &Smith, 2012).

The scope of our survey considers research that have been conducted ~~overa period of~~from 2001 to 2018 across variability of geospatial. The total of 74 scientific articles retrieved from multiple sources including ScienceDirect, Springerlink, IEEE Xplore Digital Library, Taylor & Francis and Google Scholars were reviewed. Most of them were published in peerreviewed journals as full publications, while others were selected from book chapters and conference papers.

This paper aims to provide a clear description of forecasting tools, models, methods and information needed to forecast the airborne particulate matter. In brief, our goals are twofold: (1) to classify the application of multiple methodologies based on single approaches and hybrid approaches of forecasting models, (2) to identify the influential causal variables. It is noteworthy to mention that most of the reviewed papers that we investigated were not only limited to the investigation of PM10 concentrations but other pollutants as well. The rest of this paper is organised as follows. In the next two sections, we describe the single approaches and hybrid approaches respectively. Subsequently, some analysis on the reviewed paper is made to show the most prevalent single model, hybrid model and the most influential variables. There is also a section to provide a discussion on the results and limitation of this paper. A conclusion is made in the last section.

Single Approaches

In this section, single forecasting models applied to predict PM10 pollutants are reviewed. 36 different research articles have been reviewed.

Neural Networks

Artificial neural network (ANN) is the most typical forecasting model that has been applied worldwide in many different kinds of research fieldsand was invented by McCulloch and Pitts between 1943 to 1947 (Dedovic, 2016). This review highlights that the model has become the most prevalent applicatio in this area. Chelani *et al*., (2002) explored the predictability of ANN models to forecast PM10 concentration and toxic metals in Jaipur India. Meteorological data such as temperature, wind speed and direction, relative humidity and also temporal variable were assigned ranging from 1993 to 1998. A feed-forward ANN (FF-ANN) was characterized by three layers with back propagation learning algorithm. They justified their prediction by using root mean square of error (RMSE) and $R^2$and concluded that the predictability of ANN is superior with the capability to predict peaks and nonlinear patterns of experimental data. However, a selection of the model's architectures are very subjective particularly on the number of layers or hidden neurons.

Kukkonen *et al*., (2003) has been a good reference on the issue of forecasting with ANN. Their main study was to achieve an objective of evaluating and comparing forecasting models for hourly concentration of NO_2 and PM10 in Helsinki, Finland. The forecasting models which have been favoured by the authors were five different cha), NN-HoG (NN with homocedastic Gaussian noise), NN-Lapl (NN with noise of Laplacian distribution racteristics of ANN models; multilayer perceptron – ANN (MLP-ANN), NN-2HeG (NN with two heterogenous components), NN-3HeG (NN with three heterogenous components), statistical model, linear regression (LR) model and a deterministic model. Experimental data were considered for four years in a row from 1996 to 1999 and comprised mean per hour of traffic flow, pre-processed, meteorological parameters, concentration of pollutants (NO_x, NO_2 and O_3) and temporal variables (weekday, sine of year day, cosine of year day and hour). Meanwhile, the traffic flow information was used only by deterministic model. They presented their forecasting results by classifying the results based on different stations that the data were collected (Toolo and Vallila) and the results were analysed based on three statistical indices, the index of agreement (IA), the fractional bias (FB) and the correlation of determination (R^2). The authors concluded that the ANN with its limitations can still be considered as a reliable forecasting tool with proper characterization of the model. However, it is observed that for training phases in particular,this model is very time consuming. Another limitation is its inability to predict spatial concentration distributions in urban areas

Chaloulakou *et al*., (2003) and other researchers focused on ANN model to forecast the daily mean of airborne matter. They selected meteorological parameters from Athens, Greece such as surface temperature, relative humidity and horizontal wind speed and direction through a procedure of stepwise regression analysis. They opted not to include other air pollutants as their objective was to investigate the efficiency of forecasting in terms of meteorological and time-period parameters. Experimental parameters ranged from June 1999 to May 2001. Two ANN models were built, noted as MLP1 and the MLP2 and two LR models were used; LR1 and LR2, one with lagged PM10 information while the other was vice versa respectively. The ANNs were characterized by hyperbolic tangent activation function with levenberg-marquardt (LM) algorithm and mean squared error (MSE) as the error function during training process. They extended the knowledge that the ANN models were sufficient enough to get a sound forecast of PM10 and the lagged concentration would improve the forecast significantly. However, it depends on the characteristics that will be assigned to the ANN forecasting models.

Hooyberghs *et al*., (2004) implemented predictability of ANN to design an operational system which could alarm exceedances of daily average of ground level of PM10 concentration. The forecast was made for two-days ahead. Data of meteorological forecast were utilized, sought from related authority in ~~of~~ Belgium from 1997 to 2001. There were 20 parameters which included the classical set of meteorological parameters, boundary height layer, cloud cover and height of land transport. Thoroughly, two different parameters were mixed to perform different forecasting of PM10. As to justify the accuracy of the proposed model, they determined the RMSE valueand to provide accurate alarm of exceedances of PM10 concentration, they evaluated the value of success index (SI). To ensure validation of the experiment, they used a persistence model to perform ~~a~~ forecasting using the same set of experimental data. Their findings showed that the extra parameters on the forecasting would not have significant effects on~~to~~ the forecasting.

Hooyberghs *et al*., (2005) performed forecasting of one day ahead daily mean PM10 concentrations in Belgium by using FF-ANN. Basically, their main purpose was to investigate the relationship between number and type of input parameter towards the model's architecture. The architecture was developed with one hidden layer of four nodes characterized by the sum of squares function error by back-propagation. Two similar models but with different number of input parameters were developed. Experimental data cover theyearsfrom1997 till 2001. Parameters employed in this study were the value of PM10(day-0, 1-9 hour) and other classical meteorological parameters. Evaluations of the models' performances were made based on the value of RMSE, R^2 and SI. Finally, without rigid conclusion, they however, suggested that the latter model could be utilized as the extra input variables were available to be retrieved and would give positive significant improvement.

Corani (2005) is one of the most cited for the research paper on the issue of forecasting daily mean of PM10 concentration. The author adapted experimental data in Milan to investigate the forecasting ability of pruned-ANN and lazy learning-ANN (LL-ANN), then compared the results with FF-ANN at once. Four years of experimental data (1999-2003) consisted of four parameters, the PM10 concentration, SO_2 concentration, temperature and pressure which have undergone preprocessed procedure. Each of the proposed forecasting model has different architectures that indicated a different ability in forecasting. Apart from comparing the obtained results with the three models by indexing them with goodness indicators, the author also discussed his results in terms of threshold exceedances and compared with other relevant literature. For instance, his outcomes showed relatively insignificant ability of forecasting between the different approaches. Yet, with respect to the indicators of average goodness and threshold, the author concluded that the lazy learning and pruned neural networks were the best approaches. Nevertheless, the FF-ANN tends to be over-fitting, time-consuming and to some extent,~~of~~ difficulty in interpreting the meaning of the parameters and the relevant inputs.

Grivas & Chaloulakou (2006) developed three neural network models to predict hourly PM10concentration in Athens and Helsinki. They used lagged information of PM10 concentration, temporal variables and selected eight meteorological parameters consisted of the temperature, relative humidity, wind speed and direction, total rainfall, solar radiation and barometric pressure. The first model was built by assigning all the mentioned input variables, while the second model utilized a genetic algorithm optimization procedure (GA-MLP) to retrieve the best input variables. Then, the third was developed in the absence of meteorological parameters. They used R square, IA, RMSE and mean absolute error (MAE) to evaluate their performances and compared with multiple linear regressions (MLR) model. They discussed the results where they asserted that the optimization process was feasible with modelling and expected to produce better results of forecasting. However, in real-time forecasting conditions, some extent of compromise in performance should be expected, due to the possibility of less accurate meteorological forecasts.

Perez &Reyes (2006) have implemented ANN, persistence model and linear model to forecast the daily maximum of the 24 hour moving average of PM10 in Santiago, Chile, in which data of PM10 concentration were retrieved from five different stations. Data spanning from 2001 to 2004 during theperiod of April to August were chosen comprising one hour mean of PM10 concentrations at 6 pm and 7 pm, the observed thermal amplitude on the present day, the predicted thermal amplitude for the next day and the predicted value of "meteorological potential of atmospheric pollution" (PMCA) for the next day. To evaluate the forecasts they determined absolute percent errors (APE) and comparisons were made into the three models. In fact, they classified their forecast into three different classes which defined the number of days forecast where the authors discussed significant changes in the results of forecasting based on the classes. They summed up their study by capturing that the ANN has the edge over the linear model and persistence model for forecasting concentration of PM10 one day advancement, neither in terms of the class of the day nor the numerical value. The differences between the results were insignificant. However, there is a need to be cautious that the feed forward ANN requires good choices of experimental variables

Raimondo *et al*., (2007) implemented the non-linear statistical model ANN and support vector machine (SVM) to perform short-term forecasting of the airborne. In addition, their goal was to recognize the most relatable causal of the airborne spread by utilizing a backward selection algorithm and the Koller Sahami algorithm during the pre-processing phase of experimental data (meteorological data and pollutant concentration). The data was retrieved from a station in an urban area of Goteborg, Sweden. The pollutant measurements were sulphur dioxide (SO_2), nitrogen monoxide (NO), nitrogen dioxide (NO_2),carbon monoxide (CO), ozone (O_3) and PM10. The meteorological parameters included the sets of typical meteorological parameter with atmospheric pressure, solar radiation and precipitation of rainfall. These data were sought from January 2004 till October 2005. After selecting appropriate input parameters, two different cases with eight input parameters and 11 input parameters were implemented onto ANN and SVM with certain assigned characteristics. With the determination of correct forecasting and incorrect forecasting, the scholars concluded that the optimization procedure which sustained the selection of the best input parameters produced reliable forecasting results.

Kurt *et al*., (2008) performed a forecasting of air pollution by using ANN. However, unlike other research, the study was an online forecasting. Hence, this study involved a web-based forecasting system of air pollution. The developed system comprised ~~of~~ three modules; i) data collection module, ii) forecasting module and iii) website module. The forecast required meteorological data and ten air pollution indicators. The ANN model was generated with levenberg-marquadt optimization, with seven input nodes, 10 hidden layers and three output parameters (PM10, SO_2 and CO). Experimental data were retrieved from the air pollution forecasting website from August 2005 to July 2006. The forecast was done for 10 areas located in Istanbul with three different steps ahead from one-day ahead till three-day ahead. In addition, the forecasts were performed based on three different objectives, i) to appraise the effect of independent and cumulative predictions on the model's accuracy, ii) to obtain optimum training data set size and iii) to investigate the effect of day of week as input parameter. The authors discussed their forecasting results by comparing the three output parameters and concluded that the prediction results were sufficiently gratifying.

Fang *et al*., (2011) described the feasibility of ANN to predict PM10, PM_5 and PM2.5 concentrations in thedense city of Chongqing, China. They tested two different kinds of ANNs: radial basis function and backpropagation, both with levenberg-marquadt learning algorithm. Besides the measurement of PM10 concentration, experimental data used included building height on the street, street width, wind speed and direction, traffic flow, type of vehicle proportion, position of azimuth and motor vehicles. All these parameters collected during the spring season in 2009. Outcomes of the prediction were compared between these two models in terms of *R*. Each of the pollutants gave different results accordingly to different models. The authors justified the~~ir~~ outcomes that the RBF model performed the best with PM2.5 data. Nonetheless, they did not present further discussion on the simulation results of PM10.

A study by McKendry (2002) examined the ability of the ANN model of multi-layer perceptron (MLP-ANN) to forecast the aerosol pollutant of PM10 and PM2.5. They carried out the study to verify the claim of a successful event in forecasting O_3 by using the model. Thus, the experiment applied the same model but on the particulate matter data by considering meteorological conditions and co-pollutant predictors. Experimental data were collected from an area known as Chilliwack (eastern part of Fraser Valley, Vancouver). They ran out MLR with the same data to compare the results obtained. As a consequence, they found out that the developed model, the MLP-ANN, showed only subtle performance yet unable to perform to overcome the issue from regular conventional statistical models such as predicting extreme event. However, they suggested further enhancement of the proposed model through adoption of a variety learning algorithms.

Dedovic (2016) presented a study which applied the benefits of predictability from ANN. They ran two different cases which would vary in terms of numbers of input parameter and the time span of forecasting (one-day ahead, one-week ahead and two-week ahead). These data were sought from Sarajevo, Bosnia. The first case was having input variables of humidity, wind speed, temperature, pressure and value of PM10 particles in terms of hourly from 2012. Meanwhile, the second case was extended with extra variables of PM10 (from 2010 to 2012 and its forecast values in 2013). With the architecture of ANN, the first case did three different time span of forecasting, while the second case did the forecasting of the two-ahead only. Validation of the forecasting results was made based on the R and its squared - R^2, normalized root mean square error (NRMSE) and IA whereby the authors discovered that extra addition of input variable enhanced the forecasting result even with the large time span of forecasting.

Biancofiore *et al*., (2017) ran an experiment which applied the ANN model to analyse and predict the particulate matter pollutants (both PM10 and PM2.5). Their main goal was to forecast the pollutant concentrations from one to three days ahead by utilizing experimental data of three years in a row (2011-2013) measurement which included the parameters of PM10, PM2.5 and CO concentrations and also meteorological factors from Pescara,ItalyThe meteorological factors were temperature, relative humidity, wind speed/ direction pressure. Other additional pollutants were considered such as O_3, NO, NO_2, SO_2, benzene, toluene, mxylene, 1,3-butadiene. MLR model was chosen to be compared with the recursive ANN FF-ANN. As a result, the recursive ANN was the best among the three models. They also pointed out that the usage of CO parameter in PM10 forecasting could significantly contribute to improve the final result. In addition, they revealed that the PM2.5 concentration can be predicted by using only the real-time information on PM10, excluding the parameter of PM2.5 itself.

Park *et al*., (2017) employed the data of PM10 concentration from Seoul, South Korea to validate their proposed methodologies**:** long short term forecasting (LSTM) with RMSprop and Adagrad optimization algorithm. LSTM is another class of recurrent neural network (RNN) which exploited the RNN learning to form sequence data by applying gate vector. The set of data was collected from January 2005 to March 2016 and were preprocessed through moving average method. Comparative study was held to validate the proposed model and LR model and RNN models were used to forecast the same set of data. The LTSM with RMSprop optimization had the best forecasting results in terms of the lowest value of RMSE and MSE.

Deterministic Model

Mok *et al*., (2017) defined that a deterministic model is a combination model of a meteorological model, emission model and chemistry transport model (CTM). Unlike ANN- based models where the forecasting results are obtained from a series of learnt data, the results from deterministic models are fully determined by their input variables and initial conditions. This subsection presents the research on PM10 where forecasting results ~~are~~were fully determined by the predefined variables.

Monteiro*et al*., (2007) highlighted the application of CTM known as CHIMERE that was forced by meteorological modelling system (MM5) to forecast long-term PM pollution in Portugal in 2001. Experimental data of daily PM10 were retrieved from two different regions, Lisbon and Porto with four different monitoring stations characterized by i) urban area, ii) suburban area iii) traffic and iv) industry. Furthermore, they analysed their results based on two different seasons, the winter and the summer, to discover the model behaviour. They used bias, normalized error, RMS and R^2 to present their results. The outcomes showed normalized errors and correlation coefficients (R) for stations with suburban background. Meanwhile, different seasons also gave different results whereby the summer period showed significant R value for PM10. Before they concluded that the model could be a promising tool for air pollution management, the scholars pointed out several factors that could overestimate the PM10 concentration.

Manders *et al*., (2009) had investigated two different types of models, CTM and statistical model to forecast one day ahead of PM10 concentration in the Netherlands. The CTM**,** or more well known as the LOTUS-EUROS**,** was applied to predict the PM10 concentration by enforcing meteorological data (temperature, rain, wind direction and wind speed) and anthropogenic emissions data which included the concentration of PM2.5, amount of elemental carbon and natural sea salt emissions. The datawere collected from 2004 to 2006. The researchers used another statistical model, namely ~~as~~ PROPART to simulate the same data as the LOTUS-EUROS did. Their results were evaluated based on the statistical indicators of RMSE, mean, residue, standard deviation of error (SDE), R^2, hit rate and correct alarm. The authors proved that their option to use the CTM model was right when the LOTUS-EUROS was able to embrace the ambiguity and uncertainties in most of the causal ~~of~~ PM10 emissions rather than the statistical model, PROPART.

Meanwhile, De Ruyter de Wildt *et al*., (2011) attempted to apply the usage of CTM LOTOS-EUROS to predict the release of the PM10airborne. However, their attempt was deviated from many other research where they performed their forecast up to six-day range. This model was forced by meteorological forecast data, land use data and estimates of emissions, both natural and anthropogenic. An underestimation of PM10 in this study was fixed by a simple bias correction. It was based on a LR between model results and observations. Experimental data were collected in 2009 in a few selected areas of the Netherlands (rural and non-rural). The determination of the proposed model's predictability was done by calculating its RMSE, bias and R. The authors discussedtheir results which were based on two different areas and also two different seasons: summer and winter. They discovered that the predictability of the model might decrease for more than three days ahead forecasting. Nonetheless, the stability of the prediction can still be considered as motivating. The authors have compiled a few suggestions for further improvement of the model.

Saide *et al*., (2011) performed a research to forecast the releaseof the airborne (PM10 and PM2.5) in an urban area. They applied the use of deterministic model known as Weather Research Forecasting – Chemical (WRF-Chem model) on the pollution of the PM in the area of Santiago, a city in the high central Andes with complex terrain. They varied the settings of the model technically from the source of data, area of data collection and how they collected the data. With the complexity of causal data for PM, (responsive towards heterogenous chemical, aerosoldynamics and seasonal movement), they decided to utilize the data of CO to run the PM forecasting which they claimed to be co-related tothe movement of PM concentration in the atmosphere. Several sensitivity analyses were carried out to find the optimum configuration between CO concentration and meteorological data (temperature and wind speed) and their results used to simulate 24 hour mean of the airborne and performed further experiment on two or more days of simulation. They concluded that it was vital to announce the dispersion from previous two or more days to ensure the deterministic modelis competent. Besides, they proved that it is possible to achieve two-day operational forecasts under the circumstances presented in their study.

Koo *et al*., (2012) evaluated the performance of Air Quality Forecasting System (AQFS) by using the most updated Weather Research and Forecasting (WRF) model and the United States EPA's Models-3/Community Multiscale Air Quality (CMAQ version 4.6) to forecast the concentration of PM10. The research was done in Seoul in 2010. They conducted an inter-comparison for the PM10 forecasting between the WRF and MM5 and CMAQ (older systems of forecasting models in Korea). Anthropogenic factors were considered in this paper, including burning biomass and fugitive dust. The meteorological simulations included temperature, wind speed, relative humidity and wind direction. Meanwhile, chemical simulations were also applied which included the parameters of CO, NO_x, NH_3, SO_x, volatile organic compound (VOC) and also the measurement of PM. Statistical indicators such as IA, RMSE, mean bias (MB) and normalized mean error (NME) were used. In conclusion of the study, the WRF showed agreement with observation data and it performed better than the older system with the help of supplementary data of missing biomass and fugitive dust emissions in Korea.

Kloog *et al*., (2015) presented another application of a deterministic model in predicting the concentration of the airborne particulate matters (PM2.5 and PM10). They used the satellite-derived aerosol optical depth (AOD) that was retrieved from the Multi Angle Implementation of Atmospheric Correction (MAIAC) algorithm, which was based on Moderate Resolution Imaging Spectroradiometer (MODIS). They set up their experiment in the Mediterranean region that is characterized by a high percentage of bright surfaces.Israel, which they claimed has not been tested by others, was selected. Besidesthe data of daily PM2.5 and PM10concentrations across Israel from the 2003 to 2013, another spatial causal factor that ~~were~~was used in this case which included population density, elevation, traffic density, distance from major roads, distance fromthe shoreline and the percent of open space. Meanwhile, the daily mean of meteorological factors such as air temperature, relative humidity and rainfall were collected from the authorities. The three stages of predictions were estimated by calculating thesquare root of the mean squared prediction errors (RMSPE) before the results were discussed based on different seasons (summer and winter). They also ran a sensitivity analysis to test whether their proposed model was improved spatially and temporally. As it was the first experiment that utilized MAIAC AOD for PM2.5and PM10 prediction in an area with complex geo-climate, the authors recommended further enhancement. However, apart from the limitations of the model discussed,the authors concluded that the model could be used to restructure spatially resolved long term 24 hour mean of PM2.5 and PM10 concentrations.

Hadlocon *et al*., (2015) set up a modelling of PM10 dispersion by using a computer simulation-based model~~or~~ known as algebraic models based on a Gaussian approximation. Their purpose was to validate the applicability of the model, namely as AERMOD to simulate not only the dispersion of PM10 but also that of PM2.5. The model is owned by the US Environmental Protection Agency (EPA). Meteorological data considered the wind speed and direction, ambient temperature and pressure, cloud coverage, relative humidity and heat flux. Despite having difficulties in collecting data of the daily dispersion of the pollutant from a poultry site located in West Manfield, Ohio, their findings were very encouraging, justified by good statistical agreements and composite measures (FB, normalized mean squared error (NMSE), geometric mean bias (GMB), geometric variance (GV) and fraction of 2 (FA2)) either from PM10 or PM2.5.

You *et al.,* (2016) predicted the concentration of PM10 by developing a non-linear empirical model that was essentially equipped by the MODIS-derived AOD. Basically, their prediction was based on the information of three-year data of daily PM10 concentration from 2011 to 2013. The data information was collected from 13 monitoring stations in Xi'an, a northwestern area in China. Apart from that, they utilized the data of AOD derived from the Moderate Resolution Imaging Spectroradiometer (MODIS), reanalysis data provided by the authoritiesand the surface meteorological measurements which included ground-based wind speed, relative humidity, temperature and horizontal visibility. The proposed model worked by adjusting and correcting the AOD data based on meteorological conditions. Validation was done by comparing the model with a simple linear regression on the same simulation data. They performed their analysis based on different monsoonal season and they used the indicators of R^2 to describe the relationship between the causal factors and the PM10 concentration. Further analysis on the prediction error was done by using the indicators of RMSE and APE and found out that their suggested model was performing up to three times towards accurate prediction.

Ho (2017) conducted a research of model PM10 concentration in the city of Ho Chi Minh, Vietnam. The dispersion of PM10 was modelled by using a meteorological model known as Finite Volume Model (FVM) and the Transport and Photochemistry Mesoscale Model (TAPOM). Besides the meteorological parameters, experimental data also included the source of traffic, industrial sources and area sources (household activities and small restaurants from fuel combustion). The measurements were retrievd from the 2012 data. The emissions of PM10 from these factors were calculated based on their mathematical formulae which differed according to its sources. Their analyses were then used to further analyse its impact on human health and they suggested a few steps to curb the emissions based on their experiment.

Linear Statistical Model

In this subsection, traditional statistical linear models that are normally used in forecasting studies are presented. This section is divided into several subsections where each subsection presents one specific linear method.

Multiple Linear Regressions (MLR)

Vlachogianni *et al*., (2011) applied the conventional statistical method to build a forecasting model. They stated that the statistical models would not face complexity to establish interrelationship between predictors and predictants as they did not need any information on the relationship causal. Therefore, their goal in this research was to build accurate statistical linear predictive models, utilizing stepwise MLR towards NO_x and PM10 concentration (maximum hourly and daily average concentration for the next day) based on meteorological and air quality data of Athens and Helsinki. These two cities have different characteristics of demographic aspects and climate zone. They did an evaluation based on several performance measures such as mean bias error (MBE), R, NMSE, IA, FB, FA_2, MAE, RMSE and also fractional variance. Comparative study was done by simulating the same data with ANN model (multilayer perceptron with Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS)-quasi newton algorithm) to further validate the model. The models showed encouraging results, however, the results were different depending on the characteristics of topography of case study, variables and meteorological conditions.

Stadlober *et al*., (2008) utilized MLR application of to forecast PM10. Their option was based on their claim that the MLR offered simplicity with practical feasibility and precision. The experimental data were collected from three different characteristics of the following cities: Bolzano, Kalgenfurt and Graz in Austria. Input variables included PM10 concentration, meteorological factors (temperature and wind speed) and also anthropogenic parameters between 2003 to 2006. By evaluating the performance of the proposed model in terms of a quality function as defined, the researchers claimed that their results were satisfactory and promising. However, they did mention that there was room for further improvement for the experimental model.

Generalized Linear Model

Huebnerova &Michalek (2014) performed an estimation of PM10 during the winter season by applying two generalized linear models. The first one with the log link, while the second one with gamma distribution. These two models were different in terms of type of causal parameters used where the first model was applied using covariates, ~~mean~~while the latter utilized the data from a monitoring station. Apart from that, the data of PM10 concentration and the meteorological parameters were also included comprising temperature, wind speed, wind direction and cloud cover. Experimental data were collected from two monitoring stations located in Brno Czech Republic – Zidenice and Zvonark. These two places have great exposure to the release of PM10. Two winter seasons were considered:nfrom November 2007 to March 2008 that were used to fit the forecast model and between October 2008 to March 2009 so as to assess the prediction of PM10 by the fitted model. They performed square contingency table to evaluate models and concluded that the model ~~one~~ with covariates parameters performed better.

Nonparametric Statistical Model

We also reviewed some of the statistical models that do not have any particular statistical distributions or free distributions.

Random Forest, Mixtures of Linear Regression & Nonlinear Additive Model

Jollois *et al*., (2014) identified attributes that influenced the dispersion of PM10 by using three non-parametric statistical methods: random forest, mixtures of linear regression and nonlinear additive model. Meteorological data of three consecutive years were collected from Rouen, which consisted of temperature, wind speed and direction, relative humidity and also total precipitation. Measurements of a few other pollutants were also considered such as the NO, NO_2 and SO_2, which had a substantial effect on the level of PM10concentration. They determined values of explained variance and explained deviation to describe the forecasting performance. To provide a clear picture of the resultsobtained, the authors provided the readers with reviews from a few perspectives since each model offers different benefits depending on different tasks to be accomplished. Random forest offers great ability on prediction and quantifying influential variable while the two regression models, the generalized adaptive mode (GAM) and mixtures of LR offer dependent model with flexible classification.

Multivariate Regression Splines& Discriminant Analysis

Silva *et al*., (2001) presented their research on prediction of PM10 and PM2.5in Santiago, Chile.By considering a few important meteorological factors of PM dispersion such as the measurement of ground level temperature, relative air humidity and wind speed together with data of the PM concentration, they carried out discriminant analysis (DA) onto the data that were collected from July to October 2003. Two different approaches were adapted; the parametric classification and the non-parametric classification where positive results of the predictionwereobtained. Then, they furthered their study by predicting another experimental data collected from May to September 1994 (temperature, relative humidity, wind speed, wind direction, hour of registry and concentration of PM10). However, this time a statistical model, the multivariate regression splines (MARS), was applied. To validate the approaches, they compared the ~~obtained~~ results using the mean proportion of absolute estimation error, correlation coefficient and generalized cross validation (GVC). The predictions were done based on two different step-ahead (12 hours ahead and 36 hours ahead). Finally, they concluded that the three proposed models ~~were~~ performed with remarkable accuracy of predictions.

Other Statistical Models

Smoothing Exponential Techniques

Cortina-Januchs *et al*., (2009) considered statistical smoothing exponential techniques to forecast the two most important atmospheric pollutants in Mexico, the PM10 and the SO_2. Daily concentration of the two pollutants in Salamanca, a city in northwest Mexico was predicted by applying their methodology that was segregated into two phases**:** data preprocessing phase that removed the data that exceeded a certain range and double exponential smoothing application onto the time series. Furthermore, to determine the best forecast values they applied different time window (TW) from one hour up to 24 hours for each time series. By evaluating the optimum TW, the smoothing and trend parameters, they evaluated their forecast in terms of RMSE and MAE and concluded that the proposed techniques were presentable with lessmcomputationmyet reliable to obtain a sound pollutant prediction even though with the absence of meteorological variables.

Statistical Time Varying Model

Hoi *et al*., (2009) endeavoured to predict daily averagePM10 concentrations in coastal cities located in Macau, China. They formulated a kalman filter- based model known as time-varying autoregressive model with exogenous input (TVAREX) then compared obtained results with an ANN model by levenberg-marquadtalgorithm. Meteorological data of wind speed and direction together with rainfall measurement from 2001 to 2005 were assigned within this experiment. The authorsused indicators of RMSE, MAPE, IA and $R^2$to discuss their results and summarized their findings by reviewing that the TVAREX is more favourable than the ANN for its adaptive nature.

Persistence, Generalized Additive Model & Cluster-wise Linear Model

Poggi & Portier (2011) highlighted another forecasting study using three different statistical models to forecast the daily mean of PM10 concentration. The models were persistence, generalized additive model and clusterwise linear model. This study was carried out by considering three different stations located in different areas in France. By retaining the typical set of meteorological parameter, they retrieved the daily experimental data from December through March over the period from 2004 to 2009. In brief, their methodologies started with variable selection using random forests:a non-linear and nonparametric statistical method. The forecasting procedure proceeded with a non-linear additive model and mixtures of linear regression or the clusterwise linear model. The discussion on the obtained forecasting results was done by dividing them into two categories: one with the indices of the percentage of explained variance (EV), R, MAPE, RMSE, IA and SS. The second category was evaluated with the indices of the probability of detection (POD), the false alarm rate (FAR) and the TS with a certain fixed value of range and value of threshold. Thus, among the three applied models, the clusterwise regression model performed better than others. They concluded that there is potential to provide an accurate forecast of the daily average PM10 concentration by fitting a function of meteorological variables.

Other Forecasting Models

Zeroth-level Classifier System for Data Mining

Tzima *et al*., (2009) introduced a method of forecasting namely Zeroth-level classifier system for data mining (ZCS-DM) to predict daily average PM10 concentration in Greece. Experimental data for seven years (2001 till 2006) were used, including classical meteorological attributes: temperature, wind speed and direction and humidity. Pollutant measurements of SO_2, NO_2, O_3 and CO were also considered in this experiment. Comparison study was carried by implementing other methods such as MLP, SVM and linear discriminant analysis. Determination of Kappa Index (KI) could define each performance by the proposed model. The ZCS-DM is justified as a superior model owing to its behaviour of being flexible and insensitive towards missing values.

Maximum Entropy Model

Zhang *et al*., (2017) considered a probabilistic model to assess the concentration of PM10 in Beijing, China. The model, MaxEnt - an abbreviation of maximum entropy, was adopted in this research that included various forms of parameter in the estimation of PM10. They classified their experimental data into four types; the PM10concentration, meteorological data, principal roads and land cover. These parameters were collected from various sources from the year 2008 to 2011. They revealed that the MaxEnt was able to establish a nonlinear relationship held by the experimental parameters which would help in estimating PM10 concentration. Finally, they concluded that the concentration of the airborne are varied depending on many factors, especially the traffic congestion and functionality of certain areas.

Logistic Regression, Decision Tree, Multivariate Adaptive Regression Splines and ANN

Zickus *et al*., (2002) presented multiple forecasting models to predict PM10concentration in Helsinki. Data of daily mean PM10 concentration from 1996 till 1999 were used and twelve meteorological data were sought during the years involving the measurement of temperature, wind speed and direction, friction velocity, precipitation, pressure, total cloudiness, dew point temperature and state of the ground. Before the four models performed the forecasting, the variable selections were made. Each of the forecast models (logistic regression, decision tree (CART), multivariate adaptive regression splines (MARS) and ANN) were valuated based on its index of success. The authors commented that there were indistinct differences of the performances showed by each model when none of the investigated model was able to present remarkable results of forecasting.

Chaos Theory (Local Polynomial Approximation)

Chelani & Devotta, (2006) chose a nonlinear analysis in order to predict the air pollutant. The non-linear analysis was favoured by the scholars as most of daily-life problems that need to be solved involved dynamic relationships. Hence, this paper applied the chaos theory by utilizing local polynomial approximation based on reconstructed phase space. Autocorrelation function and Fourier power spectrum were used as the first step to recognize the behaviour of daily concentration of PM10 time-series in India between 1999 until 2002. The dynamics of the time-series was additionally studied through correlation integral analysis and phase space reconstruction. As to further validate the applicability of the proposed method, a comparative studywas carried ou. They observated that the non-linear local approximation outperformed the autoregressive model. It was justified by the small relative error of prediction possessed by the local approximation model compared to the auto regressive model.

Hybrid Approaches

A hybrid forecasting model to forecast real-life problememergedin the midst of the popularity of single forecasting models. The hybridization aimed to have an ensemble of abilities from each model to overcome a single technology's weakness that would generate better results than the ones achieved with the use of each isolated technique (Zarandi *et al*. 2012). Due to this, a number of literature about hybrid forecasting models would be discussed in this section. Thirty eight published articles were reviewed in this section and twenty four different hybrid models were used with at least 18 being hybrid of ANN-based models.

Artificial Neural Network (ANN) Based Model

Artificial Neural Network-Principal Component Analysis (ANN-PCA)

Skrzypski *et al*., (2008) attempted to investigate the ANN model to forecast the next 24-hour concentration of PM10. They investigated the ability of the MLP-ANN and radial basis function (RBF-ANN) models for predicting air quality in Lodz, Poland,in relation to maximum daily concentrations of particulate matter PM10 in the air. Experimental data included the measurement of PM10 concentration and meteorological data were retrieved from November to March every year for the period between 2004 to 2007. These data were preprocessed using principal component analysis (PCA). By running the proposed model with specific optimum characterization, the authors analysed the obtained results by determining the percentage of errors between the predicted value and the real value of PM10 concentration. Despite expressing the limitations of the ANN that required a developed environmental monitoring system to do air quality prediction, the authors concluded that the model could produce fair results of forecasting air quality.

Paschalidou *et al*., (2011) extended another application of ANN models to forecast PM10 per hour in Cyprus. They employed two different types of ANN; MLP-ANN and RBF-ANN. Besides that, another model based on principal component regression was applied to simulate the same forecast. This study used pollutant and meteorological data that were retrieved from July 2006 to July 2008. Their developed ANN models were characterized by BFGS training algorithm with three layers of the ANNs comprising 15 neurons in input layer, 26 neurons in hidden layer and one neuron for output layer. Meanwhile, the statistical model was developed by considering PCA and stepwise regression analysis (SRA) for predictors' data. The main purpose of the analysis was to filter the experimental data which havea substantial effect on the PM10 dispersion. Intriguingly, the authors presented a comparative study between their results with other previous studies that used ANN forecasting models in terms of R^2 and IA. Further evaluation on the derived models were made by establishing contingency table that consisted of the bias, the false alarm rate (FAR), the probability of episode detection (POD) and the percentage correct (PC). They found out that the MLP had an edge over the other models. To provide a strong conclusion for the proposed MLP models, they performed further validation of the MLP through a series of case study (Saharan dust events) and finally were assured that the MLP was able to supply information of weak air quality one day in advance.

Voukantsis *et al*., (2011) carried out a study which applied almost similar framework of research regarding forecasting the particulate matters in the atmosphere. They applied ANN-MLP model and PCA to forecast PM10 and also PM2.5 Then, they compared their application based on two different results, which were forecasting of daily mean of PM concentration in two different cities, Thessaloniki and Helsinki. Both cities have different topographic characteristics. Meteorological data, starting from 2001 until 2003, included temperature, relative humidity, wind direction and wind speed that were selected as experimental data. To optimize the input selection, they applied a hybrid optimization procedure with incorporated grid search algorithm, result of linear-regression and ANN-MLP as the objective function. Other than the usage of the RMSE and R, they used IA and KI as statistical measures to validate the proposed forecasting models. Overall, their study found out that the proposed hybrid model showed satisfactory performances.

Neuro- Fuzzy Inference System (ANFIS)

Mihalache *et al*., (2015) drew a scheme of short-term forecasting of the PM10 concentration by applying an adaptive neuro- fuzzy inference system (ANFIS) modelling approach. They presented the next hour forecasting by characterizing architecture of ANFIS model with five layers together with rules of Takagi-sugeno. Backpropagation algorithm and hybrid learning algorithms were chosen and three different sets of experimental data were applied (retrieved from three urban areas in Romania). There was also an adjustment towards the type of membership function used, in which the triangle and Gaussian type were chosen in this study. The data consisted of the value of PM10 concentration in four different ranges; current concentration until three previous hours. Experimental results were discussed based on the origins of the data, meanwhile the evaluation of the models was based on the value of testing error (mean and standard deviation). Finally, they concluded that different sets of data favoured different adjustment of the model's architecture to obtain an optimum forecasting result.

Meanwhile, Mihalache & Popescu (2016) proceeded with their investigation by conducting comprehensive experiment of forecasting the next hour of PM10 concentration with three different architectures of ANFIS. They developed their experimental scheme exactly as the previous research with only slight addition on parameter of variables. Experimental data were gathered from an urban area in Romania, Ploiesti and comprised ~~of~~ hourly values for PM10 concentration, temperature and relative humidity. In addition, the three models were different in terms of their input parameters: the first model with only PM10 concentration and the other two models with an addition of temperature and relative humidity. Moreover, this study justified their conclusion based on comparison held on the three forecasting models by using statistical indices such as RMSE, IA, R and R^2. They concluded that the second model was the best prior to its performance with the smallest RMSE and the biggest values for IA, R and R^2.

ANN- ARIMA

Wongsatham & Changnam (2016) expanded the application of ANN and ARIMA by establishing two hybrid models based on the two methods. Firstly, they extended the development of ARIMA to ARIMAX (an ARIMA model with exogenous variables) and the extension is furthered with the development of hybrid model based on ARIMAX – namely ARIMAX-ANN and another one based on ANN called ANN-ARIMAX. Experimental dataincluded PM10 concentration, four pollutants – CO, O3, NO, SO2 and four meteorological factors – gas wind, temperature, relative humidity and pressure. This data were sought from Chiang Mai, Thailand from 2011 till 2013. By evaluating mean of error for each forecasting, the authors discussed that the hybrid model clearly showed improvement in forecasting results comparable to the results of isolated models (ANN, ARIMA and ARIMAX). However, they opined that it was hard to decide which hybrid model was better because the performances of hybrid models were dependent on the nature of the problem.

Díaz-Robles *et al*., (2008) provided ~~us~~ another option of integrating ANN model with a superior statistical model, the ARIMA. With the experimental data from Temuco, Chile, which consisted of the measurement of PM10 concentration and meteorological data (temperature, rainfall, pressure, solar radiation and wind speed), they forecasted the next 24-hour of PM10 concentration. Comparison was madeon the proposed model by using a deterministic MLR model as a benchmark. Various indicators were used to make the justification which finally showed that the hybrid model had the edge over another model. It offered an ability to capture non-linear pattern of data, however, it was irrefutable that the type of experimental data had a great influence on the results of forecasting.

Artificial Neural Network-Partial Least Squares(ANN- PLS)

Kim *et al*., (2009) are among the scholars who paid their attention on applying the ANN model as a forecasting model. However, unlike the others, their paper was focused on indoor concentration of PM10 whereby experimental data were collected from a subway station located in Seoul, South Korea. They performed the forecasting by using recurrent neural networkand compared with another two forecasting by using FF-ANN and regression. In addition, their study performs a preprocessing step on experimental data by applying partial least squares (PLS) and determined the most important variables in the forecasting. Initially, the authors chose nine variables, which were daily mean values of NO, NO_2, NO_x, CO, CO_2, temperature, humidity, PM10 and PM2.5. After the preprocessing step, the variables were then reduced to three called variable importance in projection (VIP)**:** the PM10, PM2.5 and NO_x. Finally, they compared the forecasting results in terms of RMSE when they found out that the recurrent-ANN gave more accurate forecasting result rather than the other model. Indeed, they concluded the preprocessing method contributed to significant changes in the accuracy of the results.

ANN – SVM

Siwek *et al*., (2009) explored the benefits of integrating forecasting models to obtain high precision prediction the next day of daily PM10 concentration. The main frameworks of their research were based on homogenized predictors, hybridizing using the method based on neural model, blind source separation (BSS) technique and wavelet decomposition to finalize the results of forecasting using single model and perform a better forecasting. In brief, they used MLP, RBF and SVM integrated either directly or using the wavelet decomposition approaches. Experimental data retrieval was made for three years in a row, 2005 to 2007 collected form the south region of Warsaw, Poland. Besides the classical set of meteorological parameters (the temperature, direction and strength of the wind, humidity), the air pressure measurement was also considered in this experiment. Empirical data obtained were evaluated based on RMSE, MAE and MAPE and compared the eight different forecasting models, each was differentiated by its development (direct-MLP, direct-RBF, direct-SVM; wavelet-MLP, wavelet–RBF, wavelet-MAPE, BSS integrator and SVM integrator). The best forecasting result corresponded to the

SVM integrators, while the best single predictors were shown by SVM-wavelet. Their main goal was achieved when significant improvement of the forecasting precision was obtained from the proposed alteration to the single predictors.

ANN-LR

Sfetsos & Vlachogianni (2010) used a new model by presenting a combination of the LR and ANN. This combination was described by a generalized equation that can be performed in linear or non-linear behaviour. It helps to figure out the most optimum variable within the forecasting. The experimental data were taken from Athens (five monitoring stations) and included seven meteorological variables over the period of 2000 until 2004. They made a forecast on hourly concentration data for each hour on daily basis. The applicability of this proposed model was then measured by statistical indicators. The forecast of each stationfrom where data were taken was analysed accordingly based on three different groups of model; (i) univariate statistical model, ii) multivariate model based on conventional variables and iii) multivariate model based on subsampling forecast data. From their study, they captured that the enhancement within the proposed model produced a better forecasting result with the help of optimized variables rather than the forecast of PM10with the sole model.

Artificial Neural Network-Fuzzy C-means (ANN – FCM)

Vega-corona *et al*., (2011) sketched their research to forecast the next hour of PM10 concentration by using four main steps. Initially, they started their methodology by providing data of forecasting which included PM10 pollutant concentrations, wind speed and its direction, temperature and relative humidity in Salamanca, the most polluted city in Mexico. They performed the fuzzy C-means (FCM) clustering algorithm to attain the relationships between variables. After that, they decided on the best structure of the MLP-ANN and designed the architecture with assigned criteria for the input layer, hidden layer and output layer, learning rate, the choice of activation function, number of training set, number of epoch, performance function and number of test set. At the final stage, evaluations were made by comparing their forecast with different time window while some of the models have the same number of cluster. The best model with the smallest value of RMSE and MAE were revealed with time-windows size of 2 hours and 3 clusters and the benefits of clustering the time series was clearly significant to the forecast.

Artificial Neural Network-Wavelet Transformation

InSiwek & Osowski, (2012) a forecasting framework was developed by involving various kinds of neural predictors, algorithms and linear model to perform 24 hours mean concentrations of PM10. Different neural networks included the MLP, RBF, Elman networks (EN) and support vector machine for regression (SVR) together with linear models. Then, they generated predicted time series, which were subjected to errors. The main idea of this study was to integrate each of the time series by decomposing them using wavelet transformations. Experimental data of meteorological variables from Warsaw, Poland ranging from 2006 to 2010 were discussed as they claimed that the variables had greatly influenced ~~on~~ the airborne concentration. They determined the cross correlation coefficient, statistical t-test and also variance inflation factor (VIF) test before they proceeded to the final forecast. At the end of the forecasting stage, the authors judged the obtained results based on five statistical measures; the MAE, MAPE, RMSE, R and IA. Consequently, they concluded that the proposal of utilizing the wavelet preprocessing gave significant improvement towards the forecast that had been made. In fact, the approach allowed the forecast model to be less reliant on the air pollution data.

***Artificial Neural Network*(*ANN-NNM*)**

Perez (2012) extended his investigation over the performance of ANN in forecasting PM10 by hybridizing with the nearest method (NNM). This paper had a similar framework from the previous research, like its experimental input variables, unless for its methods (the hybridization) and data spanning within a period from 2006 to 2011. The author found that with the help of NNM, the forecast of maximum concentration of one-day ahead for PM10 by using ANN showed improvement and suggested it as an important tool for air quality control.

Artificial Neural Network(ANN – GM)

Antanasijević *et al*., (2013) described the application of an ANN to predict annual PM10emissions and compared the results obtained with the same simulation made by conventional MLR and principal component regression (PCR). To enhance the development of the proposed model, an optimization model, genetic algorithm were put together with the model to choose the best variables before they proceeded to the forecast the data of PM10 emissions in 26 European Union countries from 1996 to 2006. Uniquely, this paper attempted to use input variables which comprised ~~of~~ sustainability and economic (industrial) parameters. These include the measurement of GDP, energy consumption, incineration of wood, rate of motorization, coal and lignite production of coal, lignite, paper, paperboard, roundwood, sawnwood, refined copper, aluminium, pig iron, crude steeland NPK fertilizers. The obtained results were compared to other conventional models which were running the same simulation in this research. As a result, the optimized ANN is much better than the non-optimized ANN. Nevertheless, both ANNs outperformed the other comparison models.

Artificial Neural Network-Time-delay Added Evolutionary Forecasting (ANN-TAEF)

De Mattos Neto *et al*., (2014) developed an intelligent systems model that was hybridized to perform a robust forecasting towards particulate matter concentration. It was called as TAEF (Time-delay Added Evolutionary Forecasting) methodwhich was particularly developed to consider the existence of random walk phenomenon from certain time series. They considered MLP-ANN to be used as the predictor meanwhile parameters of number of time lags representing the series, number of hidden units and the algorithm to perform the training of the predictor,were adjusted accordingly to time distortion. It consisted of two different phases: parameter optimization and phase adjustment. The first phase was where they applied the genetic method to find the most optimum parameters while the phase adjustment played its role to determine the final forecasting output with minimum error by adjusting the effect of time-delay within the time series. They provided the error analysis by using typical statistical indicators of MSE, MAPE and IA. Other indicators were included such as the U of Theil Statistics (Theil), the ARV measure and the POCID measure. They did numerical analyses based on two different stations that the experimental data (daily mean concentrations of PM2.5 and PM10) were collected from the year 2001 to 2003; the Kallio station and the Vallila station were located in Helsinki. Their outcomes were very impressive and they suggested that their proposed method would be a good alternative to enhance air pollutant forecast.

Artificial Neural Network- Clustering Algorithms

Meanwhile, Cortina–Januchs *et al*., (2015) explored the predictability of MLP-ANN that was an ensemble with two clustering algorithms: the K-means and FCM. The implementation of the two algorithms was done to alleviate preparation of experimental data which consisted of the PM10 pollutant concentrations, wind direction, wind speed, temperature and relative humidity. These data were retrieved from three different stations located in Salamanca, the most polluted city in Mexico. By developing the proposed model with certain criteria, it predicted the average concentration of PM10 for the next 24 hours and the obtained results were compared with a simple MLP and a LR in terms of four statistical indicators: the MSE, MAE, R and IA. The primary aim of this paper was achieved when they obtained significant decrement of the forecast error by the utilization of clustering algorithm onto the ANN model.

Artificial Neural Network-SVM and Taylor Expansion

Wang *et al*., (2015) presented a research regarding modification of forecasting models. Whilst the usage of ANN is quite typical to forecast PM10, they made some alterations to the model and the other two AI models, the support vector machine and Taylor expansion model. They integrated the models into another hybrid model that was expected to work efficiently, better than the traditional one. They validated their novel models through the data of the daily average of PM10 and SO_2 (from June 2008 to May 2010) together with their meteorological factors, including temperature, the dew point, atmospheric humidity, sea level pressure and the wind speed. The data set was collected from four stations located in Taiyuan, China. In conclusion, their experimental results were very encouraging with small values of statistical indicators (MAE, RMSE, DA, IA, FA_2 and fractional variance) from the forecasting.

Artificial Neural Network- Synoptic Patterns

Hur *et al*., (2016) conducted a forecasting of PM10 by using the ANN integrated with a method known as synoptic patterns. The method could demonstrate statistical experimental parameters. The case study held in Seoul, South Korea graded the concentration of PM10 into four different classes (low, moderate, high and very high).The parameters chosen in this study included the six meteorological variables such as pressure vertical velocity, temperature, humidity, wind speed and geopotential height. As they claimed that there was a high correlation between the concentration of PM10 in China and the PM10concentration in South Korea, they included the PM10 concentration in China as the experimental data. The research was done during the cold season in the area of study over a period of 2001 to 2014, from October to March. They discovered that that the synoptic patterns analysis showed that the chosen meteorological variables were able to classify the PM10 grades by utilizing cosine similarities for each PM10grade. Thus, it ~~is~~means that the synoptic patterns were adequate ~~as~~ to forecast PM10grade in Seoul. As they compared their results between the approaches of using synoptic patterns only as predictors with the approach that used thirteen predictors, the suggested patterns expressed an improved predictability for a certain PM10 grade. Meanwhile, the other grades were good without synoptic patterns. Therefore, they concluded that their suggested method would be very helpful and it was necessary to consider the method and also other predictors for another grade of PM10 (high and very high).

Artificial Neural Network-GA-RF

Dotse *et al*., (2017) performed forecasting of daily PM10 exceedances in Brunei by exploiting the predictability of backpropagation-ANN. The predictability of the model was enhanced by integrating the forecasting processes with GA-RF optimization to set optimum attributes. Experimental data were retrieved from four monitoring stations across Brunei during the period between 2009 to 2013. Initially, they considered rainfall, temperature, relative humidity, wind speed and wind direction as experimental data. They used statistical indices of R, IA, MBE, RMSE and MAE to evaluate the forecast performances. Additionally, they applied true predicted rate (TPR), the positive rate (FPR), the false alarm rate (FAR) and the success index (SI) in order to determine threshold exceedances. The evaluations were made based on different sources of experimental data, where different sources showed different results. Obviously, the assessment of the forecasting results proved that the application of genetically optimized random forests–back propagation-ANN prediction model was worthwhile when tolerant results were measured by the statistical indices.

Artificial Neural Network-Variational Mode Decomposition

The expansion of ANN model was furthered by Luo *et al*., (2017)who introduced a decomposition-ensemble based error correction model combining a signal processing technique variational mode decomposition (VMD) known as fast ensemble empirical mode decomposition (FEEMD) with error correction and quick search with extreme learning machine (CS-ELM). The ELM is another kind of FF-ANN and the proposed integrated model worked by having the FEEMD to decompose the forecast error sequence into a number of modes in order to separate and extract the multiple frequency components that existed in the time series of forecast errors. Then, they utilized CS algorithm to optimize the weights and thresholds between input layer and hidden layer of ELM model. Experimental data retrieved were two time series of daily PM10 concentration from Beijing and Harbin, China from January 2015 till August 2016. The data were analysed into linear transformation before being further processed. For the purpose of evaluation, they used MAPE, RMSE, MAE, Theil's inequality coefficient (TIC) and directional change statistic (DStat). For a comprehensive comparative study, two experiments were conducted with the first experiment consisting of five forecasting models listed as: ELM, CS-ELM, FEEMD-CS-ELM, FEEMD-CS-ELM-OEC, FEEMD-CS-ELM-DEC. The OEC and DEC denoted two different kinds of error correction techniques based on CS-ELM model and decomposition-ensemble model respectively. Meanwhile**,** the second experiment consisted of three state-of-the-artmodels including ARIMA, EEMD-GRNN and CEEMD-VMD-DE-ELM models, selected as the comparison models. From their results, different origins of experimental data gave different forecasting results and good forecasting results were obtained as a consequence of the hybridizing forecasting model with signal processing technique together with error correction.

Deterministic-based Models

Chemistry Transport Model- Regression

Konovalov *et al*., (2009) attempted to combine the usage of a deterministic model (a chemistry transport model known as CHIMERE) and statistical model (regression model) to perform robust predictions of the 24 hours average concentration of PM10. They exploited the ability of statistical data to correct the outcomes from the CHIMERE model. Their experimental data were seven meteorological parameters: two components of horizontal wind speed, humidity, boundary layer height, precipitation, optical attenuation, near surface temperature. Apart from that, experimental data included anthropogenic emissions of aerosols and gaseous species. Other parameters were taken into account such as dry deposition sea salt aerosol formation, re-suspensions of small particles and wet scavenging of gaseous species and aerosols. The experiment was done by considering the data from Central and Western Europe. Specifically, they did the forecast from November-March and May-September during the years between 2003 to 2006. They measured their error by using RMSE and the coefficient of determination, R^2 after performing the research by simulating the data with original CHIMERE, the sole statistical model and the combination of both of the models. The analysis was based on two different seasons, the hot and cold seasons and are also classified into three different groups**:** the urban area, suburban and rural area which significantly differed in terms of PM10 dispersion. They aimed to improve the error of their forecast and it was achieved by merging the two different types of forecasting models.

Regression –AOD-MODIS & CALIPSO

You *et al*., (2015) developed an establishment of a relationship between PM10 concentration in Xi'an, China and AOD. They exploited the usage of PM10 concentration data collected from January 1st, 2011 to December 31st, 2013. These data were sought from various sources, including from 13 monitoring stations in Xian, set of satellite data (MODIS fire count and

CALIPSO) and meteorological parameters such as wind speed, temperature, horizontal visibility and relative humidity from two different meteorological stations. Additionally, planetary boundary layer height (HPBL) was also used in the experiment. Geographically weighted regression (GWR) was applied to determine a local R^2 for each PM10 monitoring site on a daily basis. There were two versions of the GWR: the GWR with AOD as the only independent variable and the GWR with meteorological parameters as the addition variables. The accuracies obtained from the proposed models were measured by comparing its errors of estimations in terms of RMSE and APE. Besides, the cross validation method was also used in this study. The authors summarized their findings by highlighting the benefits of MODIS fire count and CALIPSO to provide informative data for high precision of estimating PM10 concentration and exploring the intricate mesh relationship between AOD and the pollutant.

MLR-MODIS (AOD)

Aneja *et al*., (2016) integrated a few methods before they developing the MLR model for predicting surface air quality in a mountaintop region of Roda in Virginia, which ~~is~~ depends mainly on coal-mining. The model comprised satellite-based method (MODIS-AOD and WRF model) and LR. One of their goals was to find a significant relationship of meteorological factors on the spread of particulate matter in the atmosphere. Experimental results were achieved by deriving the relationship equation between AOD and the PM2.5which is expected to improve the precision of forecasting the surface particulate matter. The data of PM2.5 were used before PM10 to validate the regression relation. The authors concluded that their proposed techniques might have augmented characteristics of air quality in the area of interest.

AOD-SVM

Stafoggia *et al*., (2017) discussed ~~on~~ the utilization of an experimental framework which consisted of four-staged assembled-model to predict daily concentration of PM10 for the years 2006 until 2012 at one kilometer per squaredresolution over Italy. The stages included the calibration of AOD with the PM10concentration and before the outcomes of the model were used to forecast daily PM10 over cell-days with or without the available AOD. The last stage was the expression of the impact of local sources of air pollution which might be able to inscribe specific estimations where the data were available. This stage applied a machine learning algorithm, SVM, to make further predictions of what had been done in the first stage. This study included many parameters which were the cause of the increment of PM10 concentration. Daily mean of PM10 concentration data were collected from 686 monitoring stations, while the others, such as meteorology data, AOD, HPBL, monthly data on Normalized Difference Vegetation Index (NDVI) and also spatial predictors came from many different sources. The spatial predictors included the emissions from industries, population density, road traffic, land cover terms, elevation and impervious surfaces. The outcomes were discussed by comparing the R^2 and RMSPE year by year. The authors concluded that their goal for the research was achieved with a promising improvement made after completing all the four stages.

AURORA-Optimal Interpolation

Mauricio Agudelo *et al*., (2015) proposed a new improvement on a deterministic quality model known as AURORA (Air Quality Modelling in Urban Region using an Optimal Resolution Approach) by assimilating a technique known as optimal interpolation (OI). The improvement was made to enhance its skill in estimating PM10 concentration and was validated by PM10 concentration data sought from Belgium, Luxemborg and a few parts of the Netherlands, France and Germany. They evaluated their experiments by using statistical indices of RMSE, R and mean bias. They presented their forecasting results by three models: the AURORA, AURORA with OI and AURORA with ensemble kalman filter (EKF). From analyses of forecasting error, there was absolute improvement shown by the AURORA-OI and AURORA-EKF. However, to compare between the use of OI and EKF, the OI attributes to the largest error reduction in this case study.

WRF-AERMOD

Kesarkar *et al*., (2007) overcame the problem of scarcity of data to forecast PM10 dispersion by proposing coupling method between WRF and AERMOD. Since the AERMOD functionality requires the presence of meteorological observations, the utilization of WRF was expected to provide a support system that could attribute necessary data. By adopting experimental data from Pune, India ranges from 13th to 17th April of 2005, they compared obtained simulated and observed temperature and wind fields that showed the WRF managed to support generation of meteorological input attained for AERMOD.

LR-based Model

MLR-LWC

Demuzere & van Lipzig, (2010), presented their study that discovered the applicability of a forecast model based on statistical method - the MLR to hindcast levels of $O_3$and PM10. The first regression-based model is the MLR that was developed by the authors in their past research, then the MLR that integrated with one type of classification known as Lamb weather type. The third model was developed by using the Lamb weather which hadundergone through a data stratification process before being integrated with MLR. To evaluate the methods, they performed several statistical tests such as the models skewness and kurtosis. Experimental data from Cabauw, the Netherlands were used for the years 2001 until 2006, where the last two years were applied as training data. As another PM10research, the meteorological data that were important to the forecast study being used within this research are temperature, relative humidity, cloud cover, wind speed and direction and shortwave downward radiation. The meteorological data were mainly focused on low resolution gridded-data. Consequently, the authors discussed their results separately based on the ozone concentration and also the PM10 concentration. The results were supporting for the O_3 concentration, but somehow the authors addressed further improvement of the proposed model to be applied to the PM10 concentration. In fact, they urged that the poor performance of the PM10 forecasting might be affected by the scarcity of data of causal variable as the predictors.

MLR-PCA

Ul-Saufie *et al*., (2013) held a research combining different forecasting models and methods before validating the model by predicting experimental data of PM10concentration. They did the forecast of the next day, next-two days and next-three days by adapting meteorological parameters and other pollution concentration. They examined the predictability of hybrid models that were built by merging the combination of MLR with PCA and the combination backpropagation FF-ANN with PCA. Experimental data were collected from Negeri Sembilan, a state in Malaysia,. A comparison analyses were done in order to see how the suggested modifications were able to enhance air quality prediction. The authors claimed that they were able to reach their goal by utilizing the PCA techniques into the two different models and achieved significant decrement of the forecast error especially for the next three-day prediction.

Mixture of LR–GAM

Misiti *et al*., (2013) estimated the daily average of PM10 concentration from three different French cities. They mixed several models of LR to form a clusterwise model and predicted the pollutants by using the classical meteorological attributes such as temperature, wind speed and pressure. In addition, generalized additive model (GAM) was used as an estimator in this experiment. Cross validation was made onto the forecasts which varied in terms of time horizons. The authors claimed that their proposed techniques are applicable and have been used by the local authorities for monitoring air quality.

Stepwise Regression –Wavelet analysis

Chen *et al*., (2013) attempted to develop a hybrid model (ensemble and enhanced) to forecast PM10 concentration in 23 cities across the eastern part of China. The hybrid model was established by exploiting the superior methods of stepwise regression (SR) model and wavelet analysis. Experimental data from 2005 to 2010 retrieved included the AOD and four-times daily, meteorological factors, including surface temperature, potential temperature, precipitable water, pressure, relative humidity, sea level pressure, wind direction, humidity and total cloud cover. Correlation analyses were made between PM10 and each meteorological factor ~~at~~ before the fitted temporal meteorological factors were chosen. In a nutshell, the framework of this study analysed input parameters (PM10, meteorological parameters and satellite-derived AOD) based on the wavelet principle to decompose them into detail and approximation components in six scales. Then, PM10 forecasting model of detail and approximation component at each scale was established by using stepwise regression. Again, for the final forecast, the predicted results of the aforementioned were reconstructed based on the theory of wavelet decomposition. Next, the single stepwise regression (SSR) and combined SR-wavelet were used as an enhanced regional model for PM10 forecasting. Its performance was measured and analysed spatially and temporally. The third process was to use the enhanced model to investigate the PM10 concentration in each city, yet the best-fitted model for the cities was selected and hybridized as an ensemble for forecasting the PM10 concentration in eastern China. Evaluations were made by measuring the value of RMSE, MAE, MRE, IA, missing rate, vacancy rate and accuracy rate. The scholars concluded their research by elucidating their findings of high precision accuracy was a reflection that the ensemble and enhanced model was a reliable forecasting tool.

Land Use* Regression – *Meteorological Factors Regression

Liu *et al*., (2015) depicted an application of combining land use regression (LUR), meteorological factor regression (MFR) and backpropagation ANN to estimate level of PM10 concentration in Changsa, an urban area located in China. Since the LUR did not include meteorological factors, the MFR and BPNN were used to include the factors in the forecast. Meteorological factors such as temperature, wind speed, percentage of haze, mist, precipitation, cloud coverage and sun together with daily concentration of PM10 in the period from April 2013 to April 2014 and annual average concentrations in 2010 were used to develop MFR. During this step, the backpropagation ANN showed better performance than the MFR. Then outcomes of the proposed combination techniques showed remarkable improvement temporally and spatially to estimate PM10 concentration that the authors recommended its applications in epidemiological study.

Autoregressive-based model

SARFIMA – GARCH

Reisen *et al*., (2014) demonstrated their study that applied a statistical model to predict daily average of PM10 concentration in a metropolitan city of Ciaracica, Brazil The model known as seasonal autoregressive fractionally integrated moving average (SARFIMA) that used more than one fractional parameter and heteroscedastic errors to ensure the model was able to fit the seasonality, volatility and long-range dependency. They did further inventions by integrating the use of generalized autoregressive conditional heteroscedastic (GARCH) model to be blended well with the errors and multiple behaviours of the PM10data especially the long-memory of data. Raw experimental data were measured from 1st January 2005 to 31st December 2009. The authors used the semi-parametric procedure in order to estimate fractional parameters. In the final stage, they did a comparison study between the SARMA and SARFIMA-GARCH models to validate the model that was more accurate. By using the statistical indicators of RMSE and symmetrical MAPE (sMAPE), the results showed positive outcomes in which the adjusted model outperformed those non-fractional SARMA model in terms of forecasting errors.

Seasonal ARIMA

Hamid *et al*., (2016) demonstrated another alternative in predicting PM10 concentration mainly for short term time span of forecasting. Their main goal was to use seasonal autoregressive integrated moving average models (Seasonal ARIMA) and Akaike Information Criterion (AIC) to distinguish the best time series model to predict the release of PM10 in Nilai, Negeri Sembilan, Malaysia. Experimental data of hourly PM10 concentration from April 2008 to March 2009 were considered by the scholars. Then, three error indicators – RMSE, MAPE and normalized average error (NAE) were applied ~~as~~ to determine accuracy of the forecast and discovered that the seasonal ARIMA with the presence of three parameters (autoregressive part of order, moving average part of order and seasonal moving average part of order) was the most appropriate model to forecast PM10 concentration in the region of this case study.

Adaptive Nonlinear State-Space & Kalman Filter

Zolghadri & Cazaurang, (2006) conducted a study to develop a modelling system based on an adaptive non-linear state-space based filter. An extended Kalman filter was adapted to solve the filtering equations. Then, the model was tested over a set of empirical data of daily mean of PM10 concentrations and its causal parameters which were collected from Bordeaux, a city ~~located~~ in France. Meteorological parameters were used in this study such as wind speed, wind direction and temperature. Meanwhile, other pollutants released were also included such as the CO, NO and NO_2. The authors provided numerical analysis by using statistical indicators of the absolute mean error (AME), R^2 and the IA. They found that the application of their proposed method was able to capture the nonlinearity and dynamic behaviour of the PM10 concentration and its causal. Nonetheless, they mentioned that the rate of having good forecast relied on the choice of input variables and initial state.

Fuzzy-based Model

Domanska & Wojtylak, (2012) depicted a study on forecasting pollutants by integrating the method of distance function and the theory of fuzzy numbers. The pollutants included the PM10, PM2.5, SO_2, NO, CO and O_3 retrieved from Poland. Their methodologies included weather forecast (January 2002) and meteorological variables (January 2004 for cloud cover, wind speed pressure, temperature, water vapor pressure, humidity and wind vector) and in fact, they also considered the experts' opinions. The forecasting was made for next 12-hour, 24-hour and 36-hour before being evaluated with average of the error, maximum of error, alpha-standardization of error and alpha-beta-standardization error. Finally**,** the results were obtainedanalysed based on the fixed band representing the level of PM10 concentration in which each of the forecasts had different outcomes.

Meanwhile,Domańska & Wojtylak, (2014) extended their research to forecast hourly-values of the concentrations, of PM10, SO_2, O_3 with a more comprehensive method. In fact, they renamed their previous air pollution forecasting model (APF) with extended-APFM. Basically, the extensions were made on an additional parameter of wind direction and its synoptic stations. On the basis of the previous model, there are eight main steps to accomplish the forecasting which is called explorative forecasting. Two years of experimental data collected from January 2011 to December 2012 went through three classes of forecast**:** one-day ahead, two-day ahead and three-day ahead. Typical error indices were then used to evaluate the forecasting such as the RMSE and MAE. In spite of the intricate procedures of the proposed model, the authors highlighted that the e-APFM would be useful for stations in the absence of data to obtain outstanding forecasting outcomes.

Grey-based Model

Qingxin *et al*., (2009) proposed on a new innovation on grey forecasting model (GM(1,1)) to simulate the data of PM10 concentration that had been released into the air in Harbin City, China. The new invention model was metabolic GM(1,1) which inherits good criteria from the conventional one but enhanced with the ability to adapt the latest improvement of simulation data. Data of PM10released were used, from 2001 until 2007 and projection data were from 2008 until 2012. As a result, the authors who applied testing criteria of relative error, correlation degree testing, infinitesimal probability error testing and ratio of mean square deviation concluded that the enhancement was satisfactory with higher accuracyachieved in the forecasting.

Sequential Aggregation

Auder *et al*., (2016) provided another option in this study area of PM10 forecasting. They opted to identify reliable methods of sequential aggregation to build decision making tool to forecast PM10 concentration. Sequential forecasting lets the result of forecasting from one model be integrated with another ~~one~~ to build a heterogeneous one where in this paper ten forecasting models became one forecasting framework ensemble. The models included six statistical models (four statistical models from Air Normand operational forecasting model and two MLR models), three numerical models (PREVAIR at 10 km scale, PREVAIR at 5 km scale and ESMERALDA at 3 km scale) and one persistence model. These models were assigned as M1 to M9 and P for the persistence model. Two methods of aggregation, EWA (Exponential Weighted Average) and RR (Ridge Regression) were applied. Experimental data adapted from 15 stations in Normandy (a city in northwest France) starting from 2013 until 2015 for daily mean of PM10concentration. Furthermore, the daily mean temperature, the daily mean atmospheric pressure, the daily mean wind speed and the daily maximum gradient of temperature were assigned to expand the family of experts by further combining more experts. They evaluated the proposed performance by using RMSE and threat score (TS) as the statistical indicators. In brief, they concluded that such a method significantly improved the experts in many angles.

Analysis On the Reviewed Papers

This section is devoted to the discussionof detail analyses on the reviewed papers with respect to some important features within the study area of forecasting PM10concentration. Many methods were suggested. The findings show that forecasting PM10 concentration is reliable and can help to realize sustainability of air quality. To aid readers' understanding of the forecasting model that had been applied, we sorted the approaches into two main categories; the single forecasting model and the hybrid forecasting model. However, the task was cumbersome and intricate since there is an abundance of research that applied multiple different models without integrating them. . Hence, each category is discriminated by the method of single and hybridapproaches which comprise ~~of~~ multiple numbers of statistical and deterministic methods within both categories. For a better understanding, the analysis is carried out by recognizing the most prevalent forecasting models, both for the single and hybrid approaches that have been reviewed in this study. Then the most important causal variables were identified, which are believed to have great impact on the results of forecasting.

Most Prevalent Methods

Thirty six of the reviewed papers have applied single model. Figure 1 depicts the percentage of the frequency of the forecasting models according to ~~its~~ categories. It is clear that the black box model, ANN is a preferable model to forecast the PM10 concentration. As can be seen from Figure 1, 42% of the published articles used ANN as a single approach. The ANN as an AI model accounted for the highest percentage as it offers flexibility and superior ability in forecasting and modelling. Meanwhile, the deterministic model is ranked second when it has 25% of the overall reviewed publications within a single category. Linear statistical models represent 8% and 5% are for non-parametric statistical model. In addition, there are 8% of the model categorized under other statistical model which represented ~~for~~ time varying and cluster-wise model. Other sub-category that accounted for 11% represented the reviewed publications that investigated other types of machine learning apart than ANN and a few of them reviewed more than one forecasting model in their research. It comprises seven reviewed papers and each applied single methods/ modelsof different kinds. In the meantime, hybrid approaches have equally attracted researchers. Total number of the reviewed papers that applied hybrid models is thirty eight. Figure 2 represents the percentage of hybrid models.

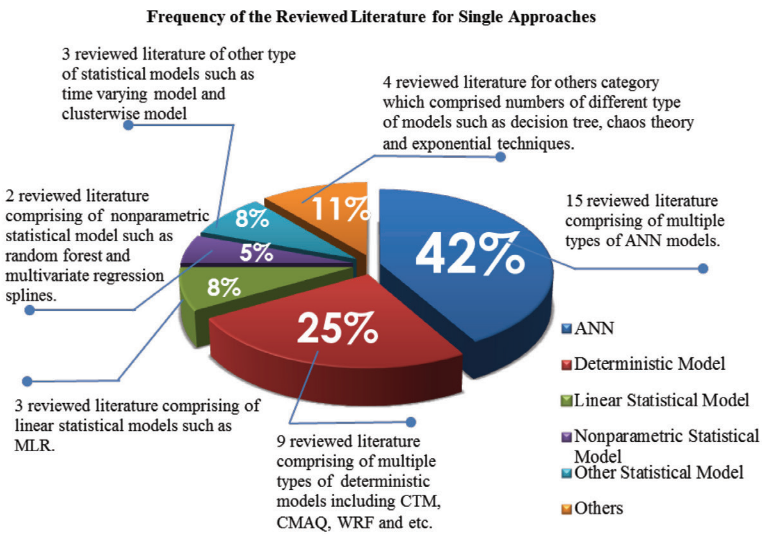

Figure1: Frequency of single approaches

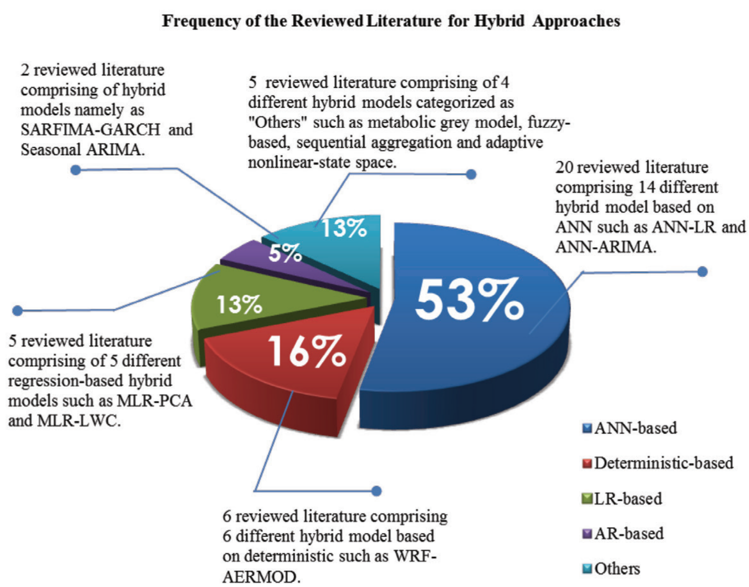

Figure 2: Frequency of hybrid approaches

The ANN-based hybrid model has the highest percentage with 53 % of publications in the hybrid category. As expected, the deterministic-based hybrid model is the second choice with 16% of its utilization to forecast PM10 concentration. Regression-based models represent hybrid models that were synchronized with MLR and land use regression which accounted up to 13%. Fuzzy-based hybrid models and autoregressive-based models and other hybrid models accounted for 5% of the reviewed publications in this category. Other~~s~~ sub-category represents for fuzzy-based, grey-based, sequential aggregation and adaptive non-linear state-space based model and accounted for 13%, with one paper for each hybrid model. Table 1 shows the frequency of single approaches and hybrid approaches according to the type of articles (journal and conference/proceeding articles). It can be seen that ANN-based models are the popular models in the forecasting studies of PM10.

Most Influential Causal Variables

As discussed earlier, the selection of causal variables of the dispersion of PM10concentration is a crucial step. Many of the reviewed papers were having two milestones in which to investigate the forecasting models and to explore the effect of certain variables as experimental data onthe forecasting result. For instance, Hooyberghs *et al*., (2004) carried out an investigation to determine how different variables as experimental input could have an effect on forecasting result. Figure 3 illustrates the percentage of the causal variables which have been placed into four different main categories**:** the meteorological variables (denoted by MET), the pollutant variables, the anthropogenic variables, the AOD variables and other variables which include chemical reaction and temporal.

Meteorological variables are defined as the causal of PM10 dispersion by the factors of natural meteorology. Most of the researchers have related this factor as a major factor that drives the concentration of PM10 in the atmosphere. The most common variables are temperature, wind speed, wind direction and relative humidity. However, there are also other meteorological factors which have been investigated by many scholars to forecast PM10 concentration such as solar radiation, total rainfall, cloud coverage, percentage of haze and

Table 1: Frequency of single and hybrid approaches in the reviewed papers.

	Approach	**Journal paper**	**Conference/ Proceeding paper**	**Total**	**% of single/hybrid model**	**% of total**
Single approach	ANN	11	4	15	42	20
	Deterministic	9	0	9	25	12
	Linear statistical	3	0	3	8	4
	Non-parametric	2	0	2	5	3
	Other statistical	2	1	3	8	4
	Others	3	1	4	11	5
Hybrid approach	ANN-based	14	6	20	53	27
	Deterministic-based	5	1	6	16	8
	LR-based	5	0	5	13	7
	AR-based	1	1	2	5	3
	Others	1	4	5	13	7

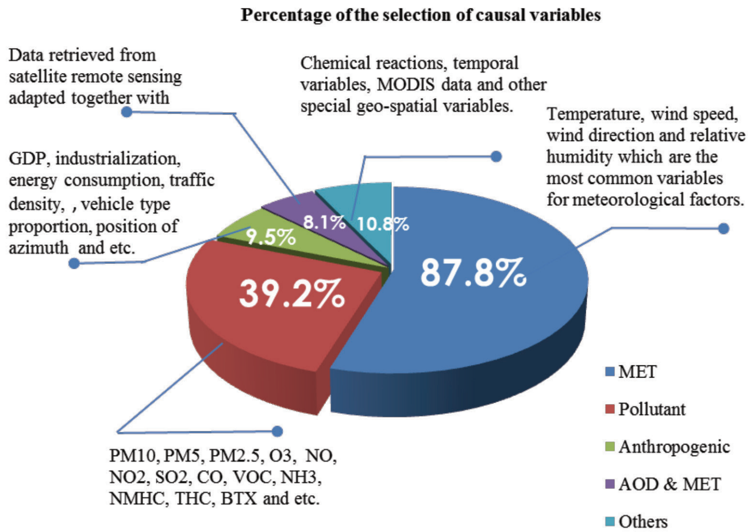

Figure 3: Percentage of causal variables

others~~.~~ This category of experimental data has the highest percentage of ~~its~~ utilization (87.8%).

The pollutant variables are defined as ~~the~~ released ~~of~~ hazardous gaseous other than PM10 such as PM2.5, O_3, NO_2 and others (see Fig 3). There are 39.2% of the reviewed publications discussed on this category. Other than these two major attributes to the dispersion of PM10, anthropogenic factors have also co-related to the level of PM10 concentration. The factors such as GDP, industrialization, energy consumption and traffic density claimed to have substantial influence on~~to~~ the level of PM10 concentration. Thus, they include the anthropogenic causal as additional variables with 9.5% of the reviewed publications utilized the factor. Apart from that, AOD variable is the retrieved data of aerosol optical depth by enforcing the technology of satellite remote sensing. There are 8.1% of the reviewed publications that applied this type of variable as additional attributes instead of meteorological causal. Meanwhile, a class of others would consist of the variables from chemical reactions, temporal variables, Moderate Resolution Imaging Spectroradiometer (MODIS) data and other special geo-spatial variables. The utilization of these variables would overlap with each other particularly, in the case where researchers aim to investigate multiple factors contributing to PM10 dispersion.

Analysis of Area of Case Study

Since factor of demographic could influence the dispersion of PM10 concentration in atmosphere, we also analysed the distribution of the reviewed papers according to the area of case study. Table 2presents the evaluation of the number of publications based on area of case study.

The period of publications and total number of citations of the articles are also included in the table. It is noteworthy that there are research that explored the forecasting of PM10 in more than one area of case study. As can be seen, China has been the leading country in terms of number of publications followed by Helsinki (Finland), France and Greece with 6 publications for each of the area. Rapid industrial development and other economic activities could be the reason for high dispersion of PM10 in China. Thus, ithas become a concern/issue to be explored. The greatest number of citations are from Helsinki, Finland (807 citations) followed by Greece (615 citations), Santiago, Chile (614 citations), Belgium (253 citations) and China (213 citations).

Table 2: Distribution by area of case study by years

Area of case study	Number of publications	Citations	Year
China	10	213	2009-2018
Helsinki	6	807	2002-2014
France	6	82	2006-2016
Greece	6	615	2006-2011
Poland	5	102	2008-2014
Santiago	5	614	2001-2012
South Korea	4	29	2009-2017
The Netherlands	4	99	2009-2015
Mexico	3	10	2009-2015
Belgium	3	253	2004-2015
India	3	180	2002-2007
Romania	2	9	2015,2016
Malaysia	2	59	2013-2016
Italy	2	177	2005-2017
Brunei	1	1	2017
Virginia	1	3	2016
Sweden	1	9	2007
Istanbul	1	101	2008
Vancouver	1	67	2002
Bosnia	1	2	2016
Portugal	1	52	2007
Israel	1	28	2015
Ohio	1	5	2015
Vietnam	1	1	2017
Austria	1	76	2008
Czech Republic	1	9	2014
Cyprus	1	71	2011
Thailand	1	1	2016
Brazil	1	23	2014
Luxemborg & Germany	1	1	2015

Year of Publications and Publishing Journals

Table 3 depicts the number of journal papers and conference/ proceeding papers over the period 2001 to 2018.

It is can be seen, that the trend is the increasing number of papers in total except for the last year range on the list. This would be because ~~of~~ the range represents only three years compared to others which represent a range of five years. The trend can be interpreted as that the awareness regarding air pollution issue especially the dispersion of PM10 in atmosphere has risen over time. The analysis of the reviewed papers is extended by analysing the sources of the journal articles (database) that have been retrieved, excluding the conference/proceeding

Table 3: Distribution of publications.

Year	Journal papers	Conference/ Proceeding papers	Total
2001-2005	8	1	9
2006-2010	14	7	21
2011-2015	25	4	29
2016-2018	12	3	15

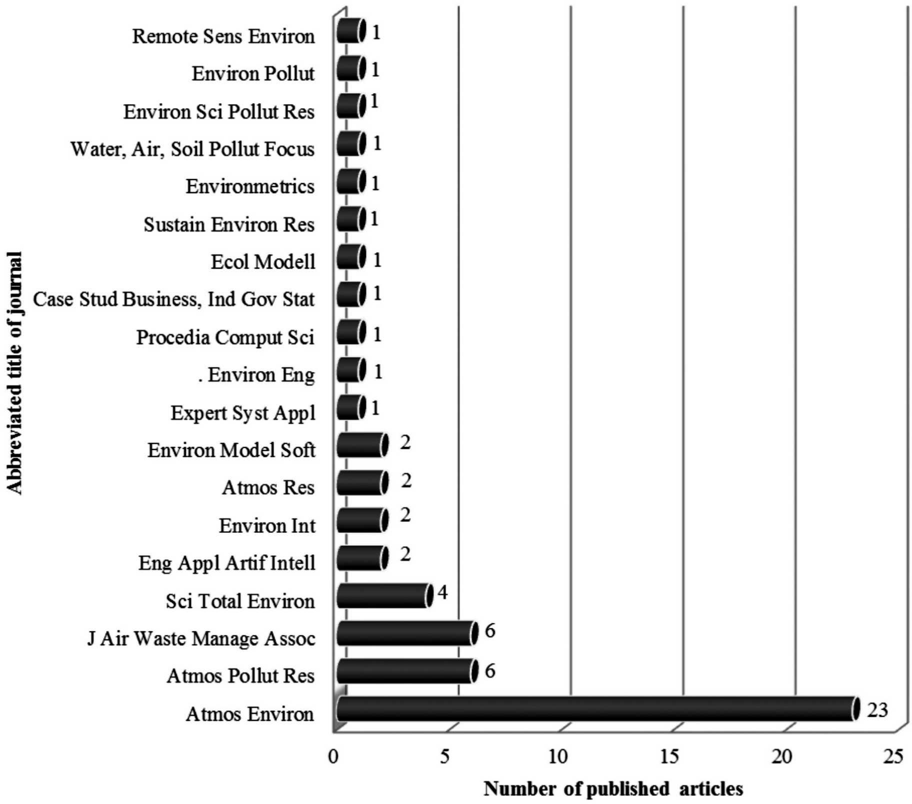

Figure 4: Graphical distribution of publications by publishing journals

papers. Figure 4 shows the distribution of published articles based on journaltitles.

With respect to this analysis, *Atmospheric Environment* is the leading journal in this field with 23 published papers. The two journals which accounted for six published articles that have been reviewed are *Atmospheric Pollution Research* and *Journal of the Air and Waste Management*. As can be seen in Figure 5, the rest of the journals published less than five articles which accounted for at least 40% from the total number of journal articles.

Discussion and Limitation

Owing to the environmental and human health concerns, PM10 pollution has gained much attention in recent decades as a threat to the universe. When faced with adverse threat from the dispersion of PM10in the atmosphere, having a well-structured plan for management of the air pollutant is very crucial. The process includes forecasting the pollutants that require promising potent tools to support the planning and management, hence the evolutionary of forecasting methods and techniques over the decades. This study is aimed at providing deep and comprehensive insights into developing PM10 forecasting models, which may be influenced by many types of causal variables. Accordingly, the forecasting models can be segregated by single and hybrid approaches. The single approach profiles of ten different single models with thirty six reviewed literature**,** regardless of whether they are statistical or deterministic methods. Meanwhile, the hybrid approach profiles 7 different hybrid models with thirty eight reviewed literature. The main focus was to highlight the most preferable method for single approach and also for hybrid approach. On the other hand, the most influential variables in forecasting were also determined. Therefore, in the fourth part of this work, a detailed analysis of the forecasting models that were available in the literature for forecasting the PM10 concentration was conducted. Summarizing the discussion on the field of forecasting PM10 concentration may help researchers get the general idea about models for further enhancement in this area. The discussions were arranged in the following points as to highlight the results and also to provide some insights on the advantages of the methods used.

i. It is found that the hybrid approaches exhibit good PM10 forecasting ability based on the number of reviewed literature. The crucial point to successful forecasting with hybrid model depends on the suitability of any two or more of the models to be integrated.

ii. To be precise, ANN forecasting model has become the most prevalent model both in single and hybrid approaches. This intelligent model represents one type of statistical method. The widespread use of ANN is irrefutable since decades ago, from sociology to engineering. The emergence of different types of ANN model offer different abilities which could influence the ability of forecasting. Its flexibility is always adaptive and its ability to handle complex conditions attract researchers to explore the model thoroughly(Jafar *et al.* 2010).

iii. Deterministic methods have accounted as the second preferable method ~~also~~ both for single and hybrid approaches. They have a major advantage in terms of data coverage which could be performed as uniform spatial (Konovalav *et al.,* 2009).

iv. However, there are more single applications compared to hybrid applications. The reason behind this could be because of its intricacy which is dependant on the knowledge of pollutant sources and imprecise description of physico-chemical processes that can lead to intolerance forecasting outputs (Vautard *et al.* 2007; Stern *et al.* 2008).

v. An attempt also has been made by Konovalov *et al.*(2009) to hybridize the deterministic forecast with statistical method. They achieved significant enhancement of the PM10 forecasting with maximum reduction of forecasting error. This outcome demonstrates that the proposed hybrid model is feasible to attain PM10 forecast with high level of accuracy.

vi. Nonetheless, Kukkonen *et al.*(2003) performed their study by comparing ANN models and deterministic models and generalized their conclusions that both models have certain inherent limitations but with appropriate spatial location and time-specific data. The ANN employment requires less computational effort rather than utilizing deterministic model computations.

vii. As a conventional method, the classical statistical model such as MLR and clusterwise LR recorded fewer numbers of single approach to forecast PM10 concentration. This is probably due to its forsaken and our limitation of data retrieval for the past two decades, starting from 2001 till recent times.

viii. In spite of being conventional, its applications are deemed to be reasonable and beneficial for hybrid approaches with respect to more reviewed literature in this category compared to the single category.

viv. It is worthwhile to note that very few attempts have been made on other AI models The AI models keep evolving and are varied like fuzzy logic and grey model.

vix. There are a numberof driver factors that could influence the accuracy of forecasting models. One of them is the type of experimental data. It is found that meteorological variables have high correlation relationship with the precision of the forecasting. The majority of reviewed literature have applied the forecasting by considering the meteorological factor.

x. The most common meteorological variables are temperature, wind direction, wind speed and total rainfall. Of the forecasting which are based on these factors, the one which considered any of theseas the main factor(s) provides reliable results. However, these findings are not exhaustive and very subjective.

xi. Some of the reviewed literature describe their variables in a very general manner, such as wind direction and temperature while there are cases where meteorological variables were meticulously used such as cosine and sine transformations vector of the wind and evaluate temperature statistically (maximum, minimum and average). Hence, the general review of the causal variables.

xii. This extensive review are based on on the retrieval of published articles and conference/ proceeding papers from certain databases for almost two decades.

xiii. Selection strategy was used based on a few keywords such as "forecasting/prediction/ modelling PM10" and "forecasting/ prediction/ modelling air quality". Thus, this selection strategy could lead toa different number of publications. Moreover, one of our limitations in this study is related to very limited access to certain database. We believe that there must be other literature that could contribute good points this review paper but wistfully unable to be comprised.

xiv. The review is analyzed by only two major issues within the area of PM10 forecasting; i) Methods/Models development and ii) type of causal variables of PM10 dispersion. Yet, we believe there are others factors which should be reviewed an in-depth, manner for example, how certain forecasting error is measured.

xv. In terms of measuring error forecasting, different use of statistical error indicators may attain slight deviation in data accuracy which could effect researchers' judgment.

However, in many cases decision on the most appropriate forecasting model is always indecisive because of the volatility behaviour of the contaminant. It is hoped that these reviews and opinions would provide a more subtle and in-depth analysis to fill the gaps between the present study and future research. Somehow, we are also anticipating to extend the review by concentrating on other~~certain~~ issues relatable to forecasting PM10, for example, the impact of pre-processing techniques onthe forecast and how different topography of case study could affect the forecast.

Conclusions

In general, the aim of this review paper is to provide information on the evolution of forecasting PM10 concentration. Intriguingly, we attempted to implement the whole forecasting processes starting from the initial step until the last step. A total of 74 publications in the field of PM10 forecasting which have emerged over the past two decades have been comprehensively reviewed and analysed accordingly to several aspects that become apparent within this research stream. Every forecasting model was sorted out based on two categories; single approaches and hybrid approaches. These two categories were then divided into several other sub-categories mainly based on statistical linear and non-linear models (intelligence machine learning) and deterministic models. Based on our analysis, AI model of ANN seems to be the most favorable model be it for its single application or for hybrid application. Furthermore, the hybrid models have been given more attention compared to the single models. In fact, the hybrid models based on ANN have the highest numbers of the overall reviewed publications. An inference drawn from this finding is that the hybrid model is more persuasive and its emergence is a concern in order to develop better forecasting tools for better air quality management. We believe that these reviews, analyses and discussions would be useful information for the readers, particularly those who are interested in searching for a good forecasting model that can be used in practical situations. For future research direction, we could get a clearer picture of some other aspects of forecasting PM10 concentration. Perhaps, this review could provide insights for future research in exploringmore promising methods or models for forecasting air quality. Most of the existing computational intelligent forecasting models are sometimes neglecting the importance of pre-processing stage. Prior to further analysis, data need to be cleaned up and outliers freed so that pre-processing technique could be included in future research. Furthermore, most of the reviewed articles did not thoroughly discuss the methods used in the process of variables selection. Selection of influential variables with a newest selection technique such as PROMETHEE could be one of the possible techniques. Apart from that, the latest performance indicators such as sensitivity analysis and ROC curve area are among the parameters that could be ventured in future research. These forecasting parameters are vital in forecasting studies as these parameters become a catalyst that empower forecasting models to work in an optimum manner and produce good forecasting tools.

Acknowledgements

The authors are grateful to the Ministry of Higher Education, Malaysia for their financial support via MyPhD sponsorship.

References

Afroz, R., Hassan, M. N., & Ibrahim, N. A. (2003). Review of air pollution and health impacts in Malaysia. *Environmental Research*, *92*, 71-77. *doi: 10.1016/S0013-9351(02)00059-2.*

Aneja, V. P., Pillai, P. R., & Isherwood, A. (2016). Particulate matter pollution in the coal-producing regions of the Appalachian Mountains: Integrated ground-based measurements and satellite analysis. *Journal of Air Waste Management Association, 67*, 1-10 . *doi: 10.1080/10962247.2016.1245686.*

Antanasijević, D. Z, Pocajt, V. V., & Povrenović, D. S. (2013). PM10 emission forecasting using artificial neural networks and genetic algorithm input variable optimization. *Science Total Environment*, *443*, 511-519 . *doi: 10.1016/j.scitotenv.2012.10.110.*

Auder, B., Bobbia, M., Poggi, J.M., & Portier, B. (2016). Sequential aggregation of heterogeneous experts for PM10 forecasting. *Atmosphere Pollution Research*, *7*, 1101-1109. *doi: 10.1016/j.apr.2016.06.013.*

Baklanov, A., Mestayer, P. G., & Clappier, A. (2008). Towards improving the simulation of meteorological fields in urban areas through updated/advanced surface fluxes description. *Atmosphere Chemistry Physics*, *8*, 523-543. *doi: 10.5194/acp-8-523-2008.*

Biancofiore, F., Busilacchio, M., & Verdecchia, M. (2017). Recursive neural network model for analysis and forecast of PM10 and PM2.5. *Atmosphere Pollution Research*, 8:652-659. *doi: 10.1016/j.apr.2016.12.014.*

Borrego, C., Monteiro, A., & Ferreira, J. (2007). Chapter 5.6 Long-term aerosol simulation for Portugal using the CHIMERE model. *Development Environtal Science*, 6:534-547. *doi: 10.1016/S1474-8177(07)06056-1.*

Brunelli, U., Piazza, V., &Pignato, L.(2007) Two-days ahead prediction of daily maximum concentrations of SO2, O3, PM10, NO2, CO in the urban area of

Palermo, Italy. *Atmosphere Environment*, 41, 2967-2995. *doi: 10.1016/j. atmosenv.2006.12.013*.

Chaloulakou, A., Grivas, G., & Spyrellis, N. (2003) Neural network and multiple regression models for PM10 prediction in Athens: a comparative assessment. *Journal of Air Waste Management Association*, 53, 1183-1190. *doi: 10.1080/10473289.2003.104-66276.*

Chelani, A. B., & Devotta, S. (2006). Nonlinear Analysis and Prediction of Coarse Particulate Matter Concentration in ambient Air. *Journal of Air Waste Manage Assoc*, *56*, 78-84. *doi: 10.1080/10473289.2006.10464432.*

Chelani, A. B., Gajghate, D. G., & Hasan, M. Z. (2002). Prediction of ambient PM10 and toxic metals using artificial neural networks. *Journal of Air Waste Manage Assoc*, *52*, 805-810. doi: 10.1080/104-73289.2002.10470827.

Chen, Y., Shi, R., Shu, S., &Gao, W. (2013) Ensemble and enhanced PM 10 concentration forecast model based on stepwise regression and wavelet analysis. *Atmosphere Environment*, *74*, 346-359. *doi: 10.1016/j.atmosenv.2013.04.002.*

Corani, G. (2005) Air quality prediction in Milan: Feed-forward neural networks, pruned neural networks and lazy learning. *Ecological Modelling*, *185*, 513-529. *doi: 10.1016/j.ecolmodel.2005.01.008.*

Cortina-Januchs, M. G., Barrón-Adame, J. M., Vega-Corona, A., & Andina, D. (2009). Smoothing exponential techniques applied for pollutant concentration prediction. *IECON Proceedings of Industrial Electronics Conference,* 3405-3409. *doi: 10.1109/IECON.2009.5415394.*

Cortina-Januchs, M.G., Quintanilla-Dominguez, J., Vega-Corona, A., &Andina, D. (2015) Development of a model for forecasting of PM10 concentrations in Salamanca, Mexico. *Atmosphere Pollution Research*, 6, 626-634. *doi: 10.5094/APR.2015.071.*

Costa , S., Ferreira, J., Silveira, C., Costa, C., Lopes, D., &Relvas H. (2014) Integrating health on air quality assessment - review report on health risks of two major European outdoor air pollutants: PM and NO2. *Journal of Toxicology and Environmental Health Part B Critical Review*, 17, 307-340.

De Mattos Neto, P.S.G., Madeiro, F., Ferreira, T.A.E., & Cavalcanti, G.D.C. (2014). Hybrid intelligent system for air quality forecasting using phase adjustment. *Engineering Applications in Artificial Intelligence, 32*, 185-191. *doi: 10.1016/j. engappai.2014.03.010.*

De Ruyter de Wildt, M., Eskes, H., & Manders, A. (2011). Six-day PM10 air quality forecasts for the Netherlands with the chemistry transport model Lotos-Euros. *Atmospheric Environment*, *45*, 5586-5594 . *doi: 10.1016/j.atmosenv.2011.04.049.*

Dedovic, M. M. (2016). Forecasting PM10 concentrations using neural networks and system for improving air quality. 2016 XI *International Symposium on Telecommunications* (BIHTEL), Sarajevo, 1-6. *doi: 10.1109/BIHTEL.2016.7775721.*

Demuzere, M. & van Lipzig, N. P. M. (2010). A new method to estimate air-quality levels using a synoptic-regression approach. Part I: Present-day O3 and PM10 analysis. *Atmospheric Environment*, *44*, 1341-1355 . *doi: 10.1016/j.atmosenv.2009.06.029.*

Díaz-Robles, L. A., Ortega, J. C., & Fu, J. S. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment*, *42*, 8331-8340. *doi: 10.1016-/j.atmosenv.2008.07.020*.

Domańska, D.&Wojtylak, M. (2014) Explorative forecasting of air pollution. *Atmospheric Environment, 92*. *doi:10.1016/j.atmosenv. 2014.03.04.*

Domanska, D.&Wojtylak, M. (2012) Application of fuzzy time series models for forecasting pollution concentrations. *Expert Systems with Applications*, *39*, 7673-7679. *doi: 10.1016/j.eswa.2012.01.023*

Dotse, S. Q., Petra, M. I., Dagar, L., & De Silva, L.C. (2017). Application of computational intelligence techniques to forecast daily PM 10 exceedances in Brunei Darussalam. *Atmospheric Pollution Research*, 1-11. *doi: 10.1016/j.apr.2017.11.004.*

Dumitrache, R.C., Iriza, A., & Maco, B. A. (2016). Study on the influence of ground and satellite observations on the numerical air-quality for PM10 over Romanian territory. *Atmospheric Pollution Research, 143*, 278-289. *doi: 10.1016/j.atmosenv.2016.08.063.*

Eseye, A.T., Zhang, J., & Zheng, D. (2018). Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information. *Renewal Energy* 118:357-367. *doi: 10.1016/j.renene.2017.11.011.*

Fang, M., Zhu, G., Zheng, X., & Yin, Z. (2011). Study on air fine particles pollution prediction of main traffic route using artificial neural network. *Proceedings of International Conference on Computing Distributive Control Intelligent Environment Monitoring* , CDCIEM 2011, 1346-1349. *doi: 10.11-09/CDCIEM.2011.431.*

Feng, X., Li, Q.,& Zhu, Y. (2015). Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment*, *107*, 118-128.

Grivas, G.&Chaloulakou, A. (2006). Artificial neural network models for prediction of PM10 hourly concentrations, in the Greater Area of Athens, Greece. *Atmospheric Pollution Research* , *40*, 1216-1229. *doi: 10.1016/j.atmosenv.2005.10.036*.

Hadlocon, L.S., Zhao, L.Y., & Bohrer G. (2015) Modeling of particulate matter dispersion from a poultry facility using AERMOD. *Journal of Air Waste Management Association*, 65, 206-217.doi: 10.1080/10962247.2014.986306.

Hamid, H. A., Yahaya, A. S., & Ramli, N. A. (2016). Short Term Prediction of PM $_{10}$ Concentrations Using Seasonal Time Series Analysis. *MATEC Web Conf* 47:5001. *doi: 10.1051/matecconf/20164705001.*

Ho, B. Q. (2017). Modeling PM 10 in Ho Chi Minh City, Vietnam and evaluation of its impact on human health. *Sustainable Environment Research*, 27, 95-102. *doi: 10.1016/j.serj.2017.01.001.*

Hoi, K. I., Yuen, K. V., & Mok, K. M. (2009). Prediction of daily averaged PM10 concentrations by statistical time-varying model. *Atmospheric Pollution Research*, *43*, 2579-2581. *doi: 10.1016/j.atmosenv.2009.-02.020.*

Hooyberghs, J., Mensink, C., Dumont, G., & Fierens, F. (2004). Short term PM 10 forecasting : a survey of possible input variables. In: Brebbia, C.A. (Ed.), *Air Pollution XII, Series: Advances in Air Pollution*, vol. 14. WIT Press, p. 171-178.

Hooyberghs, J., Mensink, C., & Dumont, G. (2005). A neural network forecast for daily average PM10 concentrations in Belgium. *Atmospheric Pollution Research*, *39*, 3279-3289. *doi: 10.1016/j.atmosenv.2005.01.050*

Huebnerova, Z. & Michalek, J. (2014) Analysis of daily average PM10 predictions by generalized linear models in Brno, Czech Republic. *Atmospheric Pollution Research*, *5*, 471-476. *doi: 10.5094/APR.20-14.055.*

Hur, S. K., Oh, H. R., & Ho, C. H. (2016) Evaluating the predictability of PM10 grades in Seoul, Korea using a neural network model based on synoptic patterns. *Environmental Pollution*, *218*, 1324-1333. *doi: 10.1016/j.envpol.2016.08.090*

Jafar, R., & Shahrour, I., Juran, I. (2010). Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. Math Comput Model 51, 1170-1180. *doi: 10.1016/j.mcm.2009.12.033*

Jana, R. K., Sharma, D. K., & Chakraborty, B. (2016). A hybrid probabilistic fuzzy goal programming approach for agricultural decision-making. *International Journal of Production Economics*, 173, 134-141. *doi: 10.1016/j.ijpe.2015.12.010.*

Jollois, F. X., Poggi, J. M., & Portier, B. (2014). Three Non-Linear Statistical Methods for Analyzing PM10 Pollution in Rouen Area. Case Stud Business, *CSBIGS*, *3*, 1-17

Juneng, L., Latif, M.T., & Tangang, F. (2011). Factors influencing the variations of PM10 aerosol dust in Klang Valley, Malaysia during the summer. *Atmosperic Environment*, 45:4370-4378 . *doi: 10.1016/j.atmosenv.2011.05.045*

Kaufman, Y.J., Tanré, D., & Boucher, O. (2002). A satellite view of aerosols in the climate system. *Nature*, 419, 215-223.

Kesarkar, A.P., Dalvi, M., Kaginalkar, A., & Ojha A. (2007). Coupling of the Weather Research and Forecasting Model with AERMOD for pollutant dispersion modeling. A case study for PM10 dispersion over Pune, India. *Atmosperic Environment*, *41*, 1976-1988 . *doi: 10.1016/j.atmosenv.2006.10.042.*

Kim M., Kim, Y., Sung, S., & Yoo, C. (2009). Data-driven prediction model of indoor air quality by the preprocessed recurrent neural networks. *ICCAS-SICE*, 1688-1692.

Kim, Y., Fu, J. S., & Miller, T. L. (2010). Improving ozone modeling in complex terrain at a fine grid resolution - Part II: Influence of schemes in MM5 on daily maximum 8-h ozone concentrations and RRFs (Relative Reduction Factors) for SIPs in the non-attainment areas. *Atmosperic Environment*, *44*, 2116-2124. *doi: 10.1016/j.atmosenv.2010.02.038.*

Kloog, I., Sorek-Hamer, M., & Lyapustin, A. (2015). Estimating daily PM2.5 and PM10 across the complex geo-climate region of Israel using MAIAC satellite-based AOD data. *Atmosperic Environment*, 122:409-416. *doi: 10.1016/j.atmosenv.2015.10.004*

Konovalov, I.B., Beekmann, M., Meleux, F.(2009) Combining deterministic and statistical approaches for PM10 forecasting in Europe. *Atmosperic Environment*, 43:6425-6434. *doi: 10.1016/j.atmosenv.2009.06.039.*

Koo, Y.S., Kim, S.T., Cho, J.S., & Jang, Y.K. (2012). Performance evaluation of the updated air quality forecasting system for Seoul predicting PM10. *Atmosperic Environment*, *58*, 56-69. *doi: 10.1016/j.atmosenv.2012.02.004.*

Kukkonen, J., Partanen, L., & Karppinen, A. (2003). Extensive evaluation of neural network models for the prediction of NO 2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmosperic Environment*, *37*(32), 4539-4550 . *doi: 10.1016/S1352-2310(03)00583-1.*

Kurt, A., Gulbagci, B., Karaca, F., &Alagha, O. (2008) An online air pollution forecasting system using neural networks. *Environment International*, *34*(5), 592-598. *doi: 10.1016/j.envint.2007.12.020*

Liu, W., Li X., & Chen, Z. (2015). Land use regression models coupled with meteorology to model spatial and temporal variability of NO2 and PM10 in Changsha, China. *AtmospericEnvironment* ,*116*, 272-280. *doi: 10.1016/j.atmosenv.2015.06.056*

Luo, H., Wang, D., & Yue, C. (2018). Research and application of a novel hybrid decomposition-ensemble learning paradigm with error correction for daily PM10 forecasting. *Atmospheric Research*, *201*, 34-45. *doi: 10.1016/j.atmosres.2017.10.009*

Manders, A. M. M., Schaap, M., & Hoogerbrugge, R. (2009). Testing the capability of the chemistry transport model LOTOS-EUROS to forecast PM10 levels in the Netherlands. *Atmosperic Environment,* 43:4050-4059. *doi: 10.1016/j.atmosenv.2009.05.006*

Martuzzi, M., Mitis, F., Iavarone, I., & Serinelli, M. (2005). Health impact of PM10 and ozone in 13 Italian cities. *World Health Organization Europe Report. Retrieved from* http://www.euro.who.int/__data/assets/pdf_file/0012/91110/E88700.pdf

Mauricio Agudelo, O., Viaene, P., & De Moor, B. (2015). Improving the PM10 estimates of the air quality model AURORA by using Optimal Interpolation, Research Council KUL: CoE PFV/10/002 (OPTEC), PhD/Postdoc grants • Flemish Government: iMinds Medical Information Technologies SBO 2. IFAC-Papers OnLine, *48*, 1154-1159. *doi: 10.1016/j.ifacol.2015.12.287.*

McKendry, I. G. (2002). Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM10 and PM2.5). Forecasting. *Journal of Air Waste Management Association*, 52, 1096-1101. *doi: 10.10-80/10473289.2002.10470836.*

Mihalache, S. F. & Popescu, M. (2016). Development of ANFIS Models for PM Short- term Prediction. Case Study. 2016 8th *International Conference on Electronics, Computers and Artificial Intelligence*, 1-6. *doi: 10.1109/ECAI.2016.7861073.*

Mihalache, S.. F., Popescu, M., & Oprea, M. (2015). Particulate Matter Prediction using ANFIS Modelling Techniques. 2015 19th *International Conference on System Theory, Control and Computing* (ICSTCC), Cheile Gradistei, 2015, 895-900. *doi: 10.1109/ICSTCC.2015.7321408.*

Misiti, M., Misiti, Y., Poggi, J., & Portier, B. (2013). Mixture of linear regression models for short term PM10 forecasting in Haute Normandie (France). Case Study Business, Ind. Gov. Stat., *6*, 47-60.

Mok, K. M., Miranda, A. I., & Yuen, K.V. (2017). Selection of bias correction models for improving the daily PM10 forecasts of WRF-EURAD in Porto, Portugal. *Atmosphric Pollution Research*, *8*, 628-639. *doi: 10.1016/j.apr.2016.12.010.*

Monteiro, A., Miranda, A. I., Borrego, C., Vautard, R., Ferreira, J., & Perez, A. T., 2007. Long-term assessment of particulate matter using CHIMERE model. *Atmosperic Environment, 41*, 7726-7738. *https://doi.org/10.1016/j.atmosenv.2007.06.008.*

Niska, H, Rantamaki, M, & Hiltunen, T. (2005). Evaluation of an integrated modelling system containing a multi-layer perceptron model and the numerical weather prediction model HIRLAM for the forecasting of urban airborne pollutant concentrations. *Atmosperic Environment*, *39,* 6524-6536 . *doi: 10.1016/j.atmosenv.2005.07.035*.

Niu, M., Wang, Y., Sun, S., & Li, Y. (2016) A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM2.5concentration forecasting. *Atmosperic Environment,* 134:168-180. *doi: 10.1016/j.atmosenv.2016.03.056.*

Park, J., Lee, K., Kim, K., & Son, U. (2017). Density Forecast Model Using Long Short Term Memory. 9th*International Conference on Ubiquitous and Future Networks* (ICUFN), Milan,2017,p.576-581. *doi: 10.1109/ICUFN.2017.*

Paschalidou,A.K., Karakitsios, S., Kleanthous, S., &Kassomenos, P.A. (2011) Forecasting hourly PM10 concentration in Cyprus through artificial neural networks and multiple regression models: Implications to local environmental management. *Environmental Science & Pollution Research*, 18:316-327. *doi: 10.1007/s11356-010-0375-2*

Perez P., &Reyes J. (2002) Prediction of maximum of 24-h average of PM10 concentrations 30 h in advance in Santiago, Chile. *Atmospheric Environment*, 36:4555-4561. *doi: 10.1016/S1352-2310(02)00419-3.*

Perez, P., &Reyes J. (2006) An integrated neural network model for PM10 forecasting. *Atmosperic Environment,* 40:2845-2851. *doi: 10.10-16/j.atmosenv.2006.01.010*

Perez, P. (2012) Combined model for PM10 forecasting in a large city. *Atmos Environ* 60:271-276. *doi: 10.1016/j.atmosenv.2012.06.024.*

Poggi, J.M,& Portier, B, (2011) PM10 forecasting using clusterwise regression. *Atmosperic Environment*, 45:7005-7014. *doi: 10.10-16/j.atmosenv.2011.09.016*

Qingxin F., Ying L., & Nanqi R. (2009). Main air contaminant PM10 in Harbin City. *Environmental Engineering,* 10-13.

Raimondo, G., Montuori, A., & Moniaci, W., (2007). A Machine Learning Tool to Forecast PM10 Level. In *Proceedings of the AMS 87th Annual Meeting*, San Antonio, TX, USA, 13-18 January 2007.

Reisen, V.A., Sarnaglia, A. J. Q., & Reis, N.C. (2014). Modeling and forecasting daily average PM10 concentrations by a seasonal long-memory model with volatility. *Environment Modelling & Software*, 51, 286-295. *doi: 10.1016/j.envsoft.2013.09.027.*

Saide, P.E., Carmichael, G.R., & Spak, S.N. (2011). Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRF-Chem CO tracer model. *Atmosperic Environment*, 45, 2769-2780. *doi: 10.10-16/j.atmosenv.2011.02.001*

Sfetsos, A., & Vlachogiannis, D. (2010). A new methodology development for the regulatory forecasting of PM10. Application in the Greater Athens Area, Greece. *Atmosperic Environment* , 44, 3159-3172 . *doi: 10.1016/j.atmosenv.2010.05.028.*

Shang, H. L. (2017). Functional time series forecasting with dynamic updating: An application to intraday particulate matter concentration. *Economic Statistics* 1, 184-200. *doi: 10.1016/j.ecosta.-2016.08.004.*

Silva, C., Perez, P., Trier, A. (2001). Statistical modelling and prediction of atmospheric pollution by particulate material: Two nonparametric approaches. *Environmetrics,* 12, 147-159. *doi: 10.1002/1099-095X(200103)12.*

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 Application in Artificial Intelligence*, 25, 1246-1258. *doi: 10.1016/j.engappai.2011.10.013.*

Siwek, K., Osowski, S., Garanty, K., & Sowinski, M. (2009). Ensemble of Predictors for Forecasting the PM10 Pollution. In. Proceedings of VXV *International Symposium on Theoretical Electrical Engineering* (ISTET), p. 318-322.

Skrzypski, J., Jach-Szakiel, E., & Kaminski, W. (2008). Neural models for prediction of maximum daily particulate matter PM10 concentration in the air in big cities as ecological safety management tools. Proceedings of 19th *International Conference on System Enginering,* 141-146. *doi: 10.1109/ICSEng.2008.13.*

Song, Y., Wang, X., Lei, L., & Xue A. (2015). A novel similarity measure on intuitionistic fuzzy sets with its applications. *Applied Intelligence*, 42:252-261. *doi: 10.1007/s10489-014-0596-z*

Stadlober, E., Hörmann, S., & Pfeiler, B. (2008). Quality and performance of a PM10 daily forecasting model. *Atmospheric Environment*, 42, 1098-1109. *doi: 10.1016/j.atmosenv.2007.10.073.*

Stafoggia, M., Schwartz, J., & Badaloni, C. (2017). Estimation of daily PM10 concentrations in Italy (2006-2012) using finely resolved satellite data, land use variables and meteorology. *Environmental International* 99, 234-244. *doi: 10.1016/j.envint.2016.11.024.*

Stern, R., Builtjes, P., & Schaap, M. (2008). A model inter-comparison study focussing on episodes with elevated PM10 concentrations. *Atmosperic Environment*, 42, 4567-4588. *doi: 10.1016/j.atmosenv.2008.01.068*

Tzima, F. A., Niska, H., Kolehmainen, M., & Karatzas, K. D. (2009). Development and evaluation of data mining models for air quality prediction in Athens, Greece, *Information Technologies in Environmental Engineering.* 331-344. *doi: 10.1007/978-3-540-88351-7.*

Ul-Saufie, A. Z., Yahaya, A. S., Ramli, N. A.(2013). Future daily PM10 concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). *Atmosperic Environment*, *77*, 621-630. *doi: 10.1016/j.atmosenv.2013.05.017.*

Vahlsing, C&Smith, K.R. (2012). Global review of national ambient air quality standards for PM10 and SO2 (24 h). *Air Quality & Atmospheric Healing*, 5, 393-399. *doi: 10.1007/s11869-010-0131-2.*

Vautard, R., Builtjes, P. H. J., Thunis, P. (2007). Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry transport models over four European cities within the City Delta project. *Atmosperic Environment*, 41, 173-188. *doi: 10.1016/j.atmosenv.2006.07.039.*

Vega-corona, A. (2011). Prediction of PM10 concentrations using Fuzzy c-Means and ANN. *37th Annual Conference of the IEEE Industrial Electronics Society,* 2607-2611. *doi: 10.1109/IECON.2011.6119735*

Vlachogianni, A., Kassomenos, P., & Karppinen, A. (2011). Evaluation of a multiple regression model for the forecasting of the concentrations of NOx and PM10 in Athens and Helsinki. *Science Total Environment*, *409*, 1559-1571. *doi: 10.1016/j.scitotenv.20-10.12.040*

Voukantsis, D., Karatzas, K., & Kukkonen, J. (2011). Intercomparison of air quality data using principal component analysis and forecasting of PM10 and PM2.5 concentrations using artificial neural networks, in Thessaloniki and Helsinki. *Science Total Environment*, *409*, 1266-1276. *doi: 10.1016/j.scitotenv.2010.12.039*

Wang, J., Yang, W., Du, P., & Li, Y. (2018). Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system. *Energy, 148. doi: 10.1016/j.enconman.2017.07.065*

Wang, P., Liu, Y., Qin, Z., & Zhang, G. (2015). A novel hybrid forecasting model for PM10 and SO2 daily concentrations. *Science Total Environment*, *505*, 1202-1212. *doi: 10.1016/j.scitotenv.20-14.10.078.*

World Health Organization (2013). Review of evidence on health aspects of air pollution - REVIHAAP Project. World Heal Organ 309.

Wongsathan, R., & Chankham, S. (2016). Improvement on PM-10 Forecast by Using Hybrid ARIMAX and Neural Networks Model for the Summer Season in Chiang Mai. *Procediaof Computer Science*, 86, 277-280. *doi: 10.1016/j.procs.2016.05.062.*

Xiong, T., Li, C., & Bao, Y. (2017). Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China. *Neurocomputing*, *275*, 2831-2844. *doi: 10.1016/j.neucom.2017.11.053.*

You, W., Zang, Z., & Zhang, L. (2016). A nonlinear model for estimating ground-level PM10 concentration in Xi'an using MODIS aerosol optical depth retrieval. *Atmospheric Research*, *168*, 169-179. *doi: 10.1016/j.atmosres.2015.09.008.*

You, W., Zang, Z., Zhang, L.. (2015). Estimating ground-level PM10 concentration in northwestern China using geographically weighted regression based on satellite AOD combined with CALIPSO and MODIS fire count. *Remote Sensing of Environment*, 168:276-285. *doi: 10.1016/j.rse.2015.07.020*

Zarandi, M.F., Hadavandi, E., & Turksen, I. (2012). A hybrid fuzzy intelligent agent-based system for stock price prediction. *Intternational Journal of Intelligent System*, 27: 947-969.

Zhang, Z., Hu, M., & Ren, J.. (2017). Probabilistic assessment of high concentrations of particulate matter (PM 10) in Beijing , China. *Atmospheric Pollution Research*, 4-11. *doi: 10.1016/j.apr.2017.04.006*

Zickus, M., Greig, A. J., & Niranjan, M. (2002). Comparison of four machine learning methods for predicting PM10 concentrations in Helsinki, Finland. *Water, Air, Soil Pollution Focus* 2, 717-729.

Zolghadri, A., & Cazaurang, F. (2006). Adaptive nonlinear state-space modelling for the prediction of daily mean PM10 concentrations. *Environmental Modelling & Software 21*, 885-894. *doi: 10.1016/j. envsoft.20-05.04.008*