Environ Sci Pollut Res
DOI 10.1007/s11356-016-6565-9

RESEARCH ARTICLE

Artificial neural network models for prediction of daily fine particulate matter concentrations in Algiers

M. R. Chellali[1,2] · H. Abderrahim[2,3] · A. Hamou[2] · A. Nebatti[4] · J. Janovec[1]

Received: 20 January 2016 / Accepted: 27 March 2016

Abstract Neural network (NN) models were evaluated for the prediction of suspended particulates with aerodynamic diameter less than 10-μm (PM_{10}) concentrations. The model evaluation work considered the sequential hourly concentration time series of PM_{10}, which were measured at El Hamma station in Algiers. Artificial neural network models were developed using a combination of meteorological and time-scale as input variables. The results were rather satisfactory, with values of the coefficient of correlation (R^2) for independent test sets ranging between 0.60 and 0.85 and values of the index of agreement (IA) between 0.87 and 0.96. In addition, the root mean square error (RMSE), the mean absolute error (MAE), the normalized mean squared error (NMSE), the absolute relative percentage error (ARPE), the fractional bias (FB), and the fractional variance (FS) were calculated to assess the performance of the model. It was seen that the overall performance of model 3 was better than models 1 and 2.

Responsible editor: Gerhard Lammel

✉ M. R. Chellali
redachellali@daad-alumni.de

1 Faculty of Materials Science and Technology, Slovak University of Technology in Bratislava, Bratislava, Slovakia

2 Laboratory of Environmental Science and Material Studies, University of Oran 1-Ahmed Benbella, Oran, Algeria

3 Hydrometeorological Institute for Training and Research-IHFR, Oran, Algeria

4 Institute of Science and Technology, University Center Ain Témouchent, Ain Témouchent, Algeria

Keywords Neural network · Particulate matter · Air pollution · PM_{10} emission forecasting

Introduction

Particulate matter with aerodynamic diameter less than 10 μm (PM_{10}) has negative effects on public life, vegetation, animals, and human health (Dockery et al. 1992; Molina and Molina 2004; Schwartz and Dockery 1992; Touloumi et al. 1994). Governments and worldwide organizations have tried to decrease emissions and human exposure to PM_{10} by supplying guidelines and by taking a number of jurisdictive penalties. In Algeria, daily average PM_{10} concentrations should not exceed the air quality standard regulations 50 μg/m^3 more than once a year. However, previous studies have proved that the airborne particle concentrations often surpass the air quality standards in Algeria (Khedairia and Khadir 2012; Laïd et al. 2006; Sabri and Med Tarek 2012). Therefore, it is of best interest for local authorities to forecast air quality suitably in advance so that the population can be protected during such periods. Several approaches have been used for modeling air quality measurements: deterministic analysis (numerical and analytical models) (Juda 1989; Zannetti 1989), statistical analysis (empirical and regression models) (Delfino et al. 1994; Fuller et al. 2002; Hadley and Toumi 2003; Ibarra-Berastegi et al. 2001), and physical analysis (Khare and Sharma 2002). Forecasting the daily/hourly levels of air pollutants represents a challenging task owing to the difficulty of the physical and chemical processes that control the formation, transportation, and elimination of particulate matter in the air (Ibarra-Berastegi and Madariaga 2003; Jacobson 1997; Seinfeld and Pandis 1998). Thus, on account of these difficulties, the statistical methods are expected to be more consistent predicting tools and offer better results than physical and deterministic approaches

Published online: 04 April 2016

(Benarie 1987; Juda 1989; Khare and Sharma 2002; Venkatram 1983; Zannetti 1989). Among several statistical methods, artificial neural networks (ANNs) have been shown to be quite powerful in capturing the complex relationships between the pollutant and the predictor meteorological variables, with good predictive results (Božnar et al. 1993; Chaloulakou et al. 2003; Gardner and Dorling 1999; Hadjiiski and Hopke 2000; Kolehmainen et al. 2001). One of the major advantages of neural networks is their ability to generalize faster execution speed and ease of working with high-dimensional data. The use of ANNs in the case of PM has been emphasized to the forecasting of hourly/daily average concentrations based on air pollution and weather historical information (Maier and Dandy 2000; Maier et al. 2004). Perez et al. (2000) reported predictions of hourly average concentrations of particles with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$) several hours in advance, based on data obtained at a fixed point in the downtown area of Santiago, Chile. However, results obtained with ANN showed prediction errors in the range from 30 to 60 %. In order to improve forecasts, they considered the reduction of noise in the data as necessary. McKendry (2000) has compared ANN model with traditional multiple regression (MLR) models for fine particulate matter PM_{10} and $PM_{2.5}$. He found that meteorological variables, persistence, and co-pollutant data were useful for forecasting PM concentration. Chelani et al. (2002) established an ANN model to forecast PM_{10} and noxious metals concentration measured in the city of Jaipur, India. Authors were able to predict concentrations quite reasonably. Lu et al. (2002) have developed neural networks for the prediction of hourly respirable suspended particle (RSP) concentrations collected in causeway bay area of Hong Kong. The simulation results showed the effectiveness of the approach. Tecer (2007) proposed ANNs to predict PM concentrations at two different stations in Zonguldak Province in the black sea area of Turkey. The results obtained showed that models can efficiently be used to forecast of air quality. Alternatively, Kukkonen et al. (2003) made extensive evaluation of neural network models for the prediction of nitrogen dioxide (NO_2) and PM_{10} concentrations, for two urban traffic locations in Helsinki, Finland, using selected traffic flow and preprocessed atmospheric variables as predictors. Results showed that ANN models can be useful and fairly accurate tools of valuation in forecasting air pollution concentrations in urban areas. Pires et al. (2008) have compared the performance of five linear models: MLR, principal component regression, independent component regression, quantile regression, and partial least squares regression to predict the daily mean PM_{10} concentrations. They concluded that the dataset size was an essential parameter for the assessment of models. The prediction was more efficient when using independent component regression for smaller dataset, while partial least squares regression was more efficient for larger dataset. Paschalidou et al. (2011) used MLP and radial basis function, for hourly PM_{10} concentrations in Cyprus. The model depends on heterogeneity of climatological and pollutant parameters of 2-year dataset. The estimation showed that the MLP models show the best predicting performance. In addition, Roy et al. (2011) have developed multiple regression and ANN models to predict particulate matter in different seasons at a large opencast coal mines in India. The outcomes revealed that ANN can forecast particulate matter levels better than multiple regression models.

Generally, ANN has been proven to be successful general machine learning methods amongst other approaches. However, the neural network-based approaches have their own disadvantages. The first disadvantage is the high complexity of neural network methods which leads to large lags during the treatment. The second disadvantage is that neural network requires a period of learning before they can be employed. It has been powerless to split the cause-effect interactions of the phenomena associated with atmospheric pollution and in particular with aerosol particle by using certain types of ANN (Kolehmainen et al. 2000). The results showed that the forecasting modeling of gaseous pollutants is more consistent than that of the particles. Furthermore, Schlink et al. (2003) demonstrated that specific techniques are often good in some respects but bad in others. To overcome this problem, they recommended to use different types of neural networks together as the best concession, as these can handle nonlinear relations and can be easily improved to site particular conditions.

In this paper, ANN models with different number of neurons are investigated to identify the best model for PM_{10} forecasting at El Hamma station in Algiers, Algeria. In the evaluation section, the results of the ANN models are compared with other conventional methods.

Measurements and methods

Study area and data

Algeria is located in North Africa midway with a coastal territory of about 1600 km long Mediterranean along the Mediterranean to the north, Libya to the east, Tunisia to the northeast, Niger to the southeast, Morocco to the west, and Mauritania and Mali to the southwest. Most of its industrial regions are located in the north, which represent 20 % of its surface area and 80 % of its population. Algiers is the capital and one of the major cities of the Northern Algeria. The city has a population of about 1.5 million inhabitants more than 3 million live in its metropolitan area. Population, industrialization, and motorized rate are constantly increasing in Algiers. Road traffic is particularly intense.

Air quality and meteorological variables are monitored at an automatic post operated by local pollution network, namely SAMASAFIA (Arabic term literally means: Clear Sky). The stations provide data with a continuous temporal resolution of 24 h. In Algiers, the network is consisted of four stations: Ben Aknoun, El Hamma, Place du 1er Mai, and Bab El-Oued distributed through the city (see inset of Fig. 1). The main pollutants monitored in the Algiers's atmosphere are classified in two categories. The first pollutants are chemical, directly emitted to the air such as nitrogen oxides (NO_x), carbon monoxide (CO), carbon oxides (CO_2) and sulfur oxides (SO_x), and hydrocarbons (HC), or solid such as PM. The secondary pollutants are the elements formed by chemical reactions such as ozone (O_3). The meteorological variables obtained and examined from station include wind speed (WS), wind direction (WD), relative humidity (RH), and air temperature (temp).

In this study, the measurements are provided by SAMASAFIA network at El Hamma post. The station is near to the city center, and thus measurements are strongly influenced by car traffic and urban activities. Dataset used to predictive PM_{10} concentration were collected during the period 2002–2006 as is shown in Fig. 1. Levels have, however, substantially surpassed the recommended national ambient air quality standards of 50 μg/m^3 for residential areas. Forty-eight percent of the total data exceed the agreed limit threshold. The average annual PM concentrations have reached the highest levels ever recorded at the station. This information is of vital necessity to national decision makers to protect living system and control of air pollution.

The missing data were found to be ~25 %, mainly caused by power cuts and numerous failures in different analysis. Due to the lack in statistics, we focused our study on a period without missing data form 2002 to 2003. The dataset (2002–2003) of measured PM_{10} concentration and three main meteorological parameters (i.e., WS, RH, and Temp) used in this paper are presented in Table 1.

In Table 1, the average annual PM_{10} concentrations of 68 μg/m^3 for 2002 and 143 μg/m^3 for 2003 were estimated for the monitoring station used in the present work. These statistics illustrate the fact of quite high pollution concentrations in the city of Algiers. Sometimes, the PM_{10} levels can reach more than four to seven times the limit value (see Table 1). Furthermore, the standard deviations (δ) as compared to average values in both PM_{10} levels and climatological variables reveal high seasonal dissimilarities in meteorological conditions of Algiers. In fact, the strong variability of weather conditions has a prominent effect on PM_{10} (Azmi et al. 2010). For instance, high wind speeds and rainy weather result in lower particulate matter concentrations in the atmosphere (Akpinar et al. 2008; Bevan et al. 1991; van Wijnen et al. 1995). In addition, the increase in the temperature will stimulate the chemical reactions, resulting from the formation of finite particulate matter in the atmosphere (Wang et al. 2013).

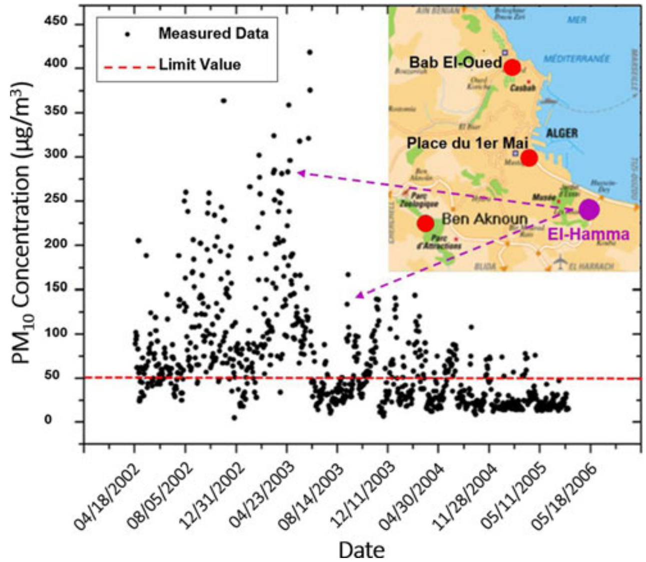

Fig. 1 PM_{10} concentration (μg/m^3) monitored at El Hamma (2002–2006). *Inset map* shows the locations of SAMASAFIA air quality monitoring networks (*purple and red circles*) in Algiers

Implementation of the model

In this study, an ANN model was developed to forecast PM_{10} concentration with the best accuracy. Prediction of future pollution levels is expected to base on the history of its individual concentration and values of some atmospheric parameters. Therefore, the timely variation of the pollutant concentration can be illustrated by the stationary time series equation:

$$\boldsymbol{Z}(\boldsymbol{t}) = \mathbf{PM}_{10}(\boldsymbol{t}) = \boldsymbol{f}\{\mathbf{PM}_{10}(\boldsymbol{t}{-}24), \mathbf{WS}(\boldsymbol{t}{-}24), \mathbf{RH}(\boldsymbol{t}{-}48), \mathbf{Temp}(\boldsymbol{t}{-}24)\} \quad (1)$$

where t is the current hour.

The ANN applied in this study is the multiple layer perception (MLP) network. This network consists of a set of interconnected layers of artificial neurons "nodes", which are arranged to form three layers: an input, hidden, and an output layers. Each layer includes one or more neurons connected with the neurons from the previous and the following layer.

Table 1 Data summary for key monitored parameters, El Hamma station (2002/2003)

Parameters	Average 2002/2003	Minimum	Maximum	δ
PM_{10} (μg/m^3)	68/143	24/67	205/359	33.8/67.2
WS (m/s)	4.4/4.2	1.7/1.7	9.7/8.1	1.8/1.7
RH (%)	68/72.6	31/39	84/83	11.6/10.1
Temp (°C)	21.3/19.1	15.4/16.5	26.5/24.8	2.3/1.7

PM_{10} articulate matter with aerodynamic diameter less than 10 μm, *WS* wind speed, *RH* relative humidity

The input layer has as many nodes as the number of input variables (e.g., meteorological parameters). The main work of the hidden nodes is to process the data and encrypt the information in the system (Fontes et al. 2014). Once the net sum at a hidden neuron is determined, a signal in every artificial neurons of the following layer is then delivered, using a function of linear integration of the incoming inputs (Kim and Gilley 2008). The function is known as activation function and is crucial for the determination of the output. For instance, if transfer function is linear, an input of meteorological and pollution parameters are transformed to an output with both negative and positive values, but if the function is exponential, the output can have only positive values. This study will use a sigmoid activation function in hidden neurons and an exponential function in output neuron, respectively. Mathematically, a sigmoid activation function has the following form:

$$f(t) = \frac{1}{1+e^{-t}} \tag{2}$$

Afterward, the shaped signal is conducted to every node in the following layer to produce new targets known as outputs. In view of these disparities between target and output, new weights are then directed to transfer functions to calculate new outputs, until the new output programs are close enough with the successful target performance.

An archetypal MLP network is shown in Fig. 2. The dataset used covers the period 2002–2003. The hourly raw data were based on pollutants and meteorological parameters of the day before at each monitoring site. Therefore, a preprocessing phase was mandatory to manage the primary data and generate a reliable database having all preferred parameters for each days. The validity of this implementation was evaluated in MATLAB 7 neural network toolbox, by indexing date and consolidating their corresponding values. This is because of its capacity to work with matrix/arrays and vector variables (Hagan et al. 1996). The program offers several facilities such as simulation, algorithm developments, graphical presentation, normalization, and demoralization methods (Roy 2012). There are no universal rules to choose on the network architecture. The number of hidden nodes is a vital factor in the model: If the numbers of hidden neurons are too small, the training data may fail to converge to a minimum in the process data space, but if a neuron number are too large, the network may over fit our training model. To develop ANN model and to avoid over-fitting, a cross validation approach is commonly recommended (ASCE 2000). The method implicates subdividing data into supplementary subgroups, executing the analysis on one subgroup (training data), and validating the analysis on the other subgroup (validation data and testing test) (Bowden et al. 2002; Comrie 1997; Gardner and Dorling 1998). Five percent of the total data were not used, because they were missing mainly due to the power cuts and various failures in the measure instruments. Missing data is a problem which poses a serious problem for the quality of the network. Due to that, it was difficult to assess correctly the daily PM_{10} concentrations. Therefore, 95 % of the measured data were used and randomly divided into three sets, training, validation, and testing. Epochs are usually needed before the error becomes adequate to PM_{10} concentrations. An entire pass through all of the input training vectors is named an epoch. Once such an epoch has ensued without errors, training data is then completed.

Several experiments were performed to determine the best combination of the optimal number of hidden layers, neurons per layer, learning rate, learning algorithm, the activation function, and errors. In this work, the feed-forward neural network with five neurons, ten neurons, and 15 neurons in hidden layer shows preeminent prediction on the approach. Networks with more than two hidden layers did not match well with the measured pollutant levels. The ANN architecture and input variables for the three models are shown in Table 2.

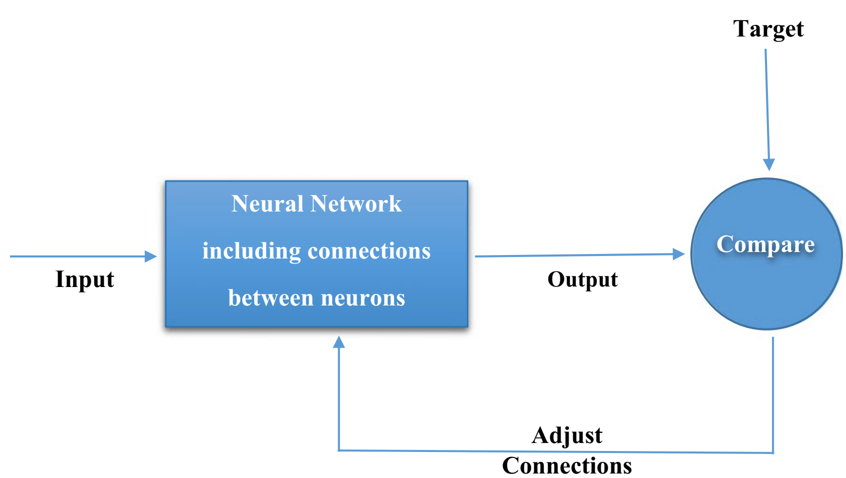

Fig. 2 A typical multiple layer perception (MLP) structure

In order to evaluate the performance of constructed ANN, several statistical parameters were computed during the forecasting process. The recommended parameters (Karppinen et al. 2000; Petersen 1997; Willmott 1981) are mean absolute error (MAE), root mean square error (RMSE), index of agreement (IA), the squared correlation coefficient (R^2), and the fractional bias (FB). The mathematical formulas for each of these statistical parameters are calculated as follows:

The root mean squared error (RMSE)

$$\mathbf{RMSE} = \left[\overline{(\boldsymbol{P_i}-\boldsymbol{O_i})^2}\right]^{1/2} \tag{3}$$

The normalized mean squared error (NMSE)

$$\mathbf{NMSE} = \frac{\overline{(\boldsymbol{O_i}-\boldsymbol{P_i})^2}}{\overline{\boldsymbol{O}}*\overline{\boldsymbol{P}}} \tag{4}$$

The mean absolute error (MAE)

$$\mathbf{MAE} = \frac{1}{\boldsymbol{n}}\sum_i|\boldsymbol{P_i}-\boldsymbol{O_i}| \tag{5}$$

The squared correlation coefficient (R^2)

$$\mathbf{R}^2 = \frac{\left[\sum_i\left(\boldsymbol{P_i}-\overline{\boldsymbol{P}}\right)\left(\boldsymbol{O_i}-\overline{\boldsymbol{O}}\right)\right]^2}{\sum_i\left(\boldsymbol{O_i}-\overline{\boldsymbol{O}}\right)^2\sum_i\left(\boldsymbol{P_i}-\overline{\boldsymbol{P}}\right)^2} \tag{6}$$

The index of agreement (IA)

$$\mathbf{IA} = \mathbf{1}-\frac{\overline{(\boldsymbol{P_i}-\boldsymbol{O_i})^2}}{\left[\left|\boldsymbol{P_i}-\overline{\boldsymbol{O}}\right|+\left|\boldsymbol{O_i}-\overline{\boldsymbol{O}}\right|\right]^2} \tag{7}$$

The fractional bias (FB)

$$\mathbf{FB} = \mathbf{2}*\left(\frac{\overline{\boldsymbol{O}}-\overline{\boldsymbol{P}}}{\overline{\boldsymbol{O}}+\overline{\boldsymbol{P}}}\right) \tag{8}$$

where P_i and O_i are the predicted and observed concentrations, respectively. The overbar refers to the average over all hourly values. The statistical parameters IA and R^2 indicate the correlation of two time sequences of values, whereas FB indicates the agreement of the mean values. The lower IA and R^2 are, the weaker is the degree of the agreement between two time sequences of values.

Table 2 ANN topology and input variables for various models

Model	Network topology (nodes per layer)	Input variables
Model 1	4-5-1	Wind speed
Model 2	4-10-1	Temperature
Model 3	4-15-1	Relative humidity Emission level

In addition, the performance of each model was evaluated by calculating the absolute relative percentage error (ARPE) and the fractional variance (FS) between predicted and measured values as follows:

The absolute relative percentage error (ARPE)

$$\mathbf{ARPE} = \left|\frac{\boldsymbol{P_i}-\boldsymbol{O_i}}{\boldsymbol{O_i}}\right| \tag{9}$$

The fractional variance (FS)

$$\mathbf{FS} = 2*\left(\frac{\boldsymbol{\delta_{O_i}}-\boldsymbol{\delta_{P_i}}}{\boldsymbol{\delta_{O_i}}+\boldsymbol{\delta_{P_i}}}\right) \tag{10}$$

The δ_O and δ_P are the standard deviation of observations and predictions, respectively.

Results and discussion

In this research work, the model was trained with small data sets due to the incomplete data base available without lack. The performances of the constructed models were assessed by calculating R^2, RMSE, δ, slope, and IA. Table 3 summarizes the model performance in the training and testing phase with different numbers of neurons after obtaining optimal architecture for ANN.

R^2 values between the measured and casted concentrations of different pollutants varied from 0.61 to 0.86 for the training data set and from 0.43 to 0.80 for the prediction data set. The corresponding IA was from 0.66 to 0.93 for the training data set and from 0.46 to 0.85 for the prediction data set. As a result of the power term, RMSE is more appropriate to illustrate the presence of significant underpredictions or overpredictions. The RMSE varied from 6.16 to 12.33 for both training and prediction data sets. The difference between the standard deviation of the predicted concentration (δ_P) and the observed concentration (δ_O) is minimum. In addition, the slope of trend line was calculated to assess the model performance. The slope values fluctuated from 0.62 to 0.95 for the training data set and from 0.7 to 1.3 for the prediction data set. These results indicate rather good fits between measured and modeled pollutant levels in most of the cases. The produced scatter plots and metrics reflect the real forecasting ability of the models. In general, the scatter plots are fairly symmetrical for the numerous types of ANN produced. This elucidates that the models are reproducing the variation in the test data set with a reasonable correctness.

Table 3 Values of calculated statistical parameters for the constructed ANN models

	Statistical parameter	Model 1	Model 2	Model 3
Training	RMSE	9.31	8.74	12.33
	R^2	0.65	0.61	0.86
	Slope	0.62	0.63	0.95
	δ	8.26	8.26	8.26
	IA	0.69	0.66	0.93
Prediction	RMSE	6.16	11.47	10.75
	R^2	0.43	0.80	0.75
	Slope	0.7	1.1	1.3
	δ	8.19	8.28	8.42
	IA	0.46	0.85	0.81

RMSE root mean square error, *IA* index of agreement

After recurrent trials, the best prediction on validation data set was achieved at 23 epochs with the learning rate of 0.05 for model 3. Likewise, for models 1 and 2, the best prediction on validation data was attained at 11 epochs with the learning rate of 0.05 and 10 epochs with the learning rate of 0.05, respectively. Table 4 summarizes quantitatively the performance of the model with different number of neurons. On the whole, it is obvious that the model shows good agreement between the two data sets (i.e., the observed and predicted values). Furthermore, the relatively low values of the MAE—13.59, 17.13, and 10.80 μg/m^3—as well as the low RMSE values—17.72, 21.13, and 13.78 μg/m^3—for models 1, 2, and 3, respectively, reflect the small deviation among the predicted and the observed concentrations. Moreover, the low values of FS reveal the fact that the variance of the MLP predictions is equal to the variance of the observed value. The low values of NMSE (0.0037, 0.0042, and 0.0051) are another indicator of the good performance of the MLP approach, as well as the FB values, which are close to zero in the three cases. The performance of these models viewed by the mean absolute percentage error (MAPE) over the total 948 prediction hours

Table 4 Performance indicators of conventional models

Statistical parameter	Model 1	Model 2	Model 3
Learning rate	0.05	0.05	0.05
NMSE	0.0042283	0.005156	0.003756
MAE	13.58829	17.1334	10.79852
δ	3.68623	4.139251	3.28611
FS	0.008808	−0.00208	−0.01881
FB	0.0069322	−0.01424	−0.04604
IA	0.91	0.87	0.96
RMSE	17.72	21.13	13.78
ARPE	0.198	0.2507	0.158043
R^2	0.72174	0.6029	0.85358
Slope	1.05994	0.95237	0.8868

NMSE normalized mean squared error, *MAE* mean absolute error, *FS* fractional variance, *FB* fractional bias, *IA* index of agreement, *RMSE* root mean square error, *ARPE* absolute relative percentage error

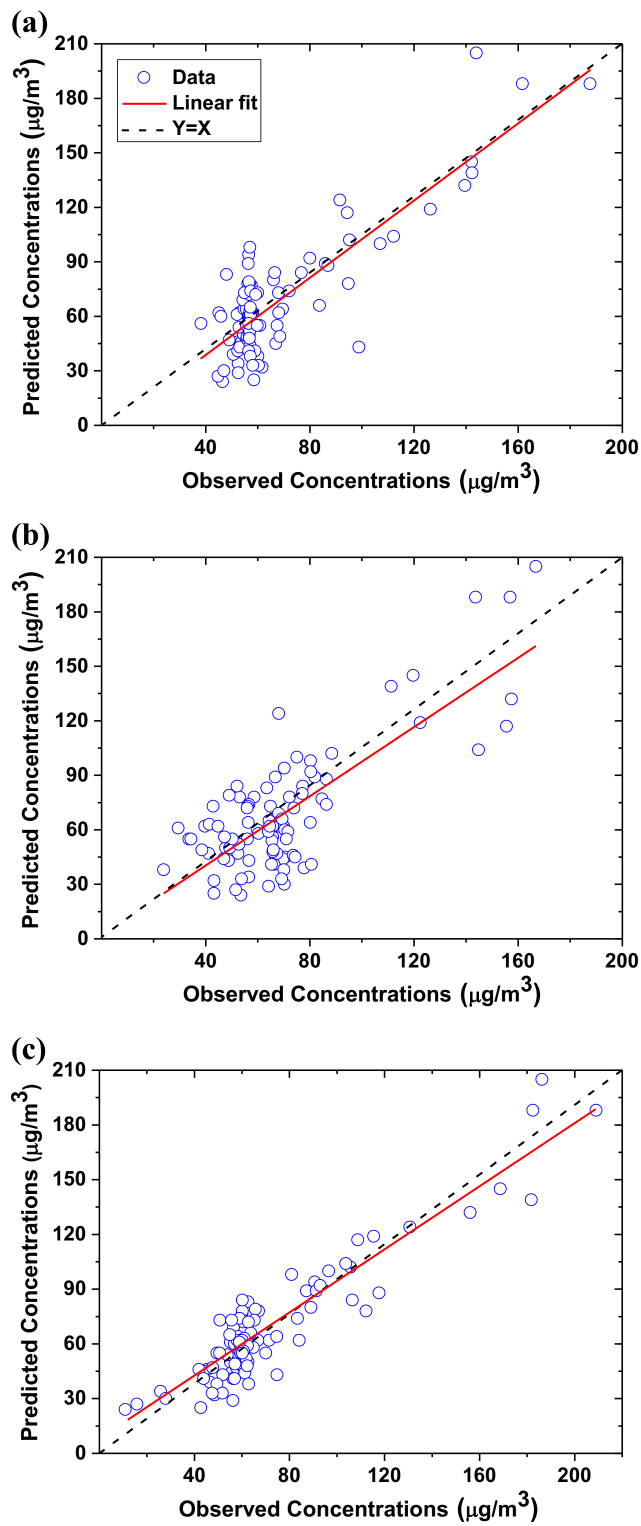

Fig. 3 Scatter plots comparison between the predicted and the observed PM_{10} concentrations in models 1 (**a**), 2 (**b**), and 3 (**c**)

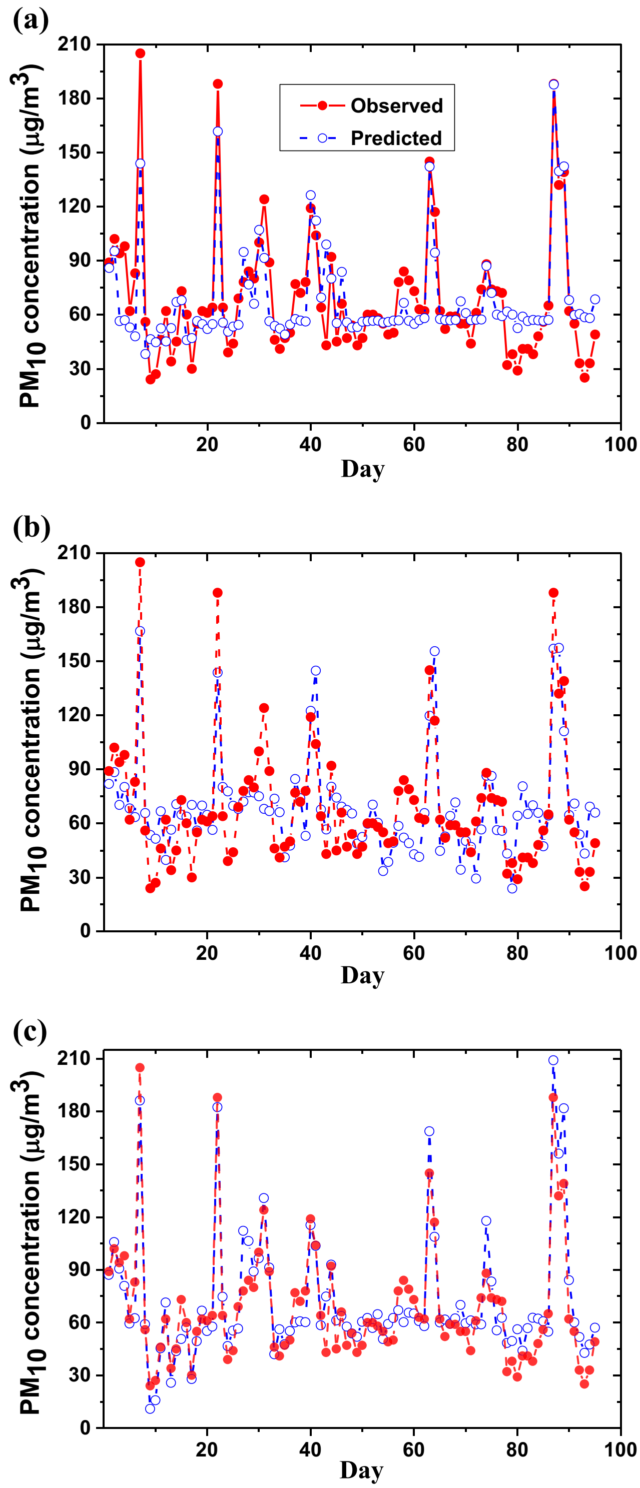

Fig. 4 Daily time series comparison of the measured and predicted concentrations of PM_{10} for a 95-day period in the model 1 (**a**), 2 (**b**), and 3 (**c**)

is 20, 25, and 16 %. These results could be considered satisfactory given that the criteria used for the input pattern selection were quite restrictive. Performance of the model may be improved if more input parameter combinations or inclusion

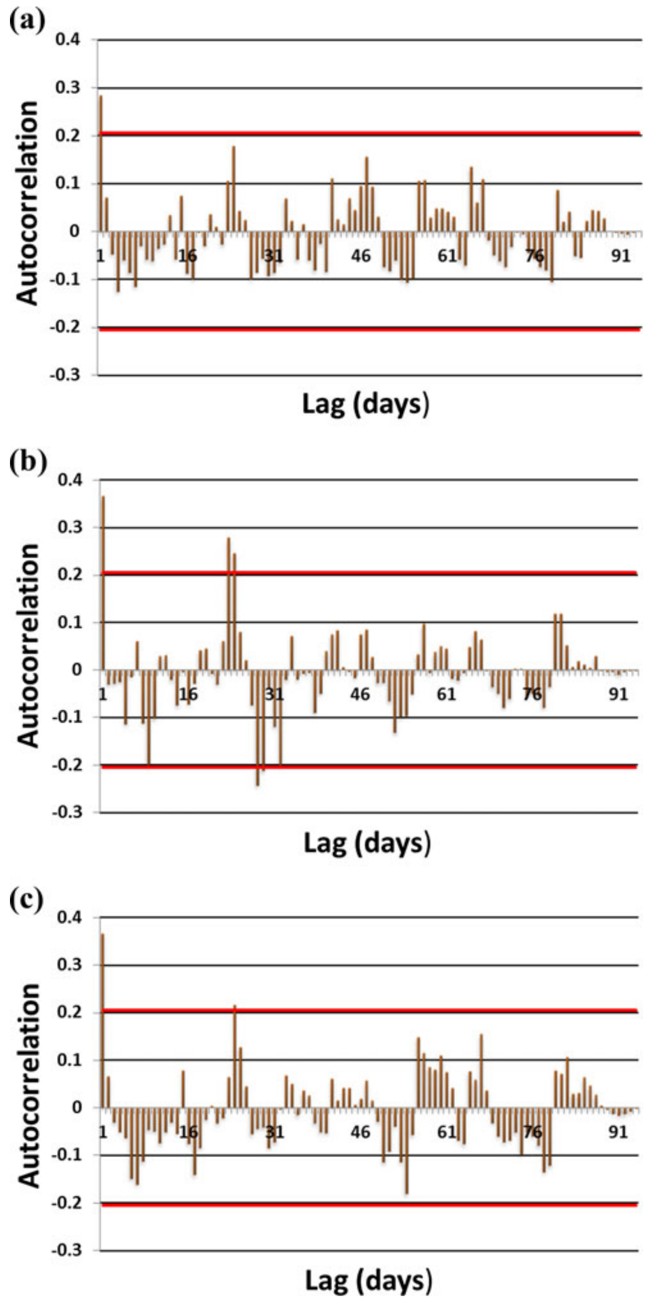

Fig. 5 Autocorrelation plot of PM_{10} for models 1, 2, and 3. The *red dashed lines* are called the critical lines

of more related parameters could be tested. The IA values are close to 1, which explains that more than 99 % of the model predictions are error free. The best predictions are achieved in model 3, which has an IA of 0.96 for 2-year (2002–2003) forecasts. Furthermore, the standard deviation is minimum in model 3. This means that the approach is mimicking the variation in the test data set with an adequate accuracy.

Nevertheless, comparison of the model performance in Table 4 clearly shows that model 3 has much better forecast capabilities. Figure 3 illustrates the scatter plots of the predicted versus observed PM_{10} emission with different number of neurons. The PM_{10} concentration predicted by optimized ANN for model 3 was highly correlated with the measured levels, with R^2 of 0.85 and slope of 0.89 (see Fig. 3c). Even though the predicted levels for models 1 and 2 shown in Fig. 3a, b were less accurate compared to model 3 optimization, coefficients have shown generally a rather good correlation to the measured levels, with R^2 of 0.60–0.72 and slope of 0.95–1.05 for models 2 and 1, respectively. Thus, it can be concluded that the overall performance of model 3 is better than the other models.

Figure 4 shows the time history graph of the predicted and observed data for a 95-day period. In appearance, there is slight scatter between the predictions and measurements in the three cases. This may be considered as "noise" and could be caused by some instantaneous, individual incidents. Here also, the model 3 (see Fig. 4c) seems to perform better than the other ones (see Fig. 4a, b). Generally, these time series graphs confirm the superior performance of the approaches in predicting the studied pollutants' concentrations. The developed neural network forecasts satisfactory the PM_{10} peaks for all cases, which are supported by the stated statistical predictors, both for the training and the validation data set.

To check whether the models can represent the data, the autocorrelation function (ACF) was used. ACF is an essential tool to evaluate the degree of the data dependence and to select a suitable model reflecting this characteristic. ACF is resulted by shifting the original PM10 against lags (number of days) and then valuing the correlation of the original time series with each one of the moved forms. The autocorrelation function will vary between −1 and +1. If ACF is close to near ±1, the data will indicate a stronger correlation. Figure 5 illustrates the autocorrelation functions for models 1, 2, and 3. The red dashed lines in the plots (Fig. 5) are called the critical lines computed as $\pm 2/\sqrt{N}$, where N represents the number of the lags or the number measured data. The autocorrelations were calculated up to Lag 94. As shown in Fig. 5, approximately 95 % of the data plots fit between the error lines. The autocorrelations were meaningfully different from zero at approximately 5 % significance level. These results are consistent, supporting the robustness of the obtained models (see Fig. 4). This finding was also consistent with results obtained by other researchers (see Table 5). Kolehmainen et al. (2000) used hybrid neural network (NN) modeling to obtain the desirable hourly forecast of PM_{10} concentrations in Kuopio, Finland, and reported an IA equivalent to 0.47. Kukkonen et al. (2003) settled five NN models for the prediction of PM_{10} and NO_2 concentrations in Helsinki, Finland, and reported R^2 values of 0.28 to 0.42 and IA values of 0.67 to 0.77, relying on the site, the input configuration, and the training algorithm. Alternatively, Chaloulakou et al. (2003) have developed an MLR and premeditated numerous statistical metrics using a self-governing testing set of data. The reported values of MAE, RMSE, and R^2 were 12.62–16, 16.94–21.9, and 0.47–0.65, respectively. Aldrin and Haff (2005) estimated a model on log scale on PM10 concentrations for four different stations. The reported R^2 coefficient varied between 0.48 and 0.70. Corani (2005) used feed-forward NNs, pruned NNs, and lazy learning to forecast daily fine particulate matter in Milan, Italy, and an IA value of 0.94 and MAE value of 8.55 were reported. Grivas and Chaloulakou (2006) developed NNs for the predictions of hourly levels of PM_{10} in Athens, Greece, and the R^2 values were evaluated between 0.50 and 0.67, RMSE values between 12.16 and 17.06, and the IA values between 0.80 and 0.89, based on the location. Papanastasiou et al. (2007) developed a NN model to air quality in the medium-sized city of Volos, Greece, and reached an R^2 value equal to 0.61, RMSE value 11.37, and IA value equal to 0.78. In addition, Kurt et al. (2008) reported an error percentage of 43 % for prediction of sulfur dioxide (SO_2) in Istanbul, whereas errors between 9 % in our approach. Hrust et al. (2009) established an empirical NN model to forecast both gaseous and particulate pollutants in Zagreb, Croatia, and

Table 5 Previous studies on the forecasting of PM_{10} levels through NNs

Area	R^2	RMSE	IA	References
Algiers, Algeria	0.60–0.85	13.78–21.13	0.87–0.96	This work
Kuopio, Finland	n/a	n/a	0.47	Kolehmainen et al. (2000)
Helsinki, Finland	0.28–0.42	n/a	0.67–0.77	Kukkonen et al. (2003)
Athens, Greece	0.47–0.65	16.94–21.9	n/a	Chaloulakou et al. (2003)
Oslo, Norway	0.48–0.70	n/a	n/a	Aldrin and Haff (2005)
Milan, Italy	0.94	n/a	0.94	Corani (2005)
Athens, Greece	0.50–0.67	12.16–17.06	0.80–0.89	Grivas and Chaloulakou (2006)
Volos, Greece	0.61	11.37	0.78	Papanastasiou et al. (2007)
Zagreb, Croatia	0.72	19.3	0.91–0.97	Hrust et al. (2009)
Switzerland	0.72	n/a	n/a	Barmpadimos et al. (2011)
Makkah, Saudi Arabia	0.52	84	n/a	Munir et al. (2013)

RMSE root mean square error, *IA* index of agreement realized a R^2 value of 0.72, a RMSE value of 19.3, and an IA equal to 0.91. Barmpadimos et al. (2011) developed a model to examine the influence of meteorology on PM_{10} for each season of the year. An average value of 0.72 and 0.52 was measured for the R^2 coefficient and MSE, respectively. Besides, Munir et al. (2013) have settled a model and compared the predicted and the observed PM_{10} concentrations on an independent testing dataset. They have reported 0.52 and 84, as the values of R^2 coefficient and RMSE, respectively.

On the whole, it is apparent that the ANN with different neurons, using the simple sigmoid as activation function, resulted as a very effective approach to forecast PM_{10} absorptions in the urban area of Algiers and can be valuable to city planners. One limitation of the current model is that even the NNs have surpassed the classical approaches in terms of performance, the selection of network architecture (neurons, number of hidden layers, and their interconnection) still a problem during ANN training stage. The quality of PM_{10} predictions by the neural network method can be upgraded by increasing the regularity of extreme values in the training process, either by including wide observations in the training data set or by including each periodic case more than a few times.

Conclusion

In the present study, a modeling effort was conducted in order to examine the potential of artificial neural network to predict the daily average PM_{10} concentrations. The quality and reliability of the developed models were evaluated via several statistical indexes (RMSE, NMSE, MAE, ARPE, FS, FB, R^2, IA). Comparing the three approaches, the model 3 showed slightly better skills in forecasting PM_{10} concentrations than models 1 and 2.

The overall models' results, as well as their ability to predict particulate matter, show an effective use on the operational level for obtaining 24-h forecasts. The presence of quality meteorological predictions is viewed as a vital premise for the models' successful application for real-time forecasts. Series of sensitivity tests revealed that the models' stability does not vary much from the experimental values. The developed ANN model can be used not only for predicting PM_{10} concentration but also for simulating different situations of PM_{10} emission by altering the values of the input variables. The results obtained can be used by the local community and governments to support policy, and for the development of plans for improving air quality at regional and national levels. Supplementary researches are planned to include other environmental quality pointers such as emission of ozone and acid oxides, via the economical and industrial indicators, and in applying new methods for input optimization, such as principal constituent and correlation analysis.

Acknowledgments The authors would like to thank the Algerian National Observatory of Environment and Sustainable Development (ONEDD) for providing access to the meteorological data used in this study.

References

Akpinar S, Oztop HF, Akpinar EK (2008) Evaluation of relationship between meteorological parameters and air pollutant concentrations during winter season in Elazığ, Turkey. Environ Monit Assess 146:211–224

Aldrin M, Haff I (2005) Generalised additive modelling of air pollution, traffic volume and meteorology. Atmos Environ 39:2145–2155

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000) Artificial neural network in hydrology, part I and II. J Hydraul Eng 5(2):115–137

Azmi SZ, Latif MT, Ismail AS, Juneng L, Jemain AZ (2010) Trend and status of air quality at three different monitoring stations in the Klang Valley, Malaysia. Air Qual Atmos Health 3:53–64

Barmpadimos I, Hueglin C, Keller J, Henne S, Prevot A (2011) Influence of meteorology on PM10 trends and variability in Switzerland from 1991 to 2008. Atmos Chem Phys 11:1813–1835

Benarie MM (1987) The limits of air pollution modeling. Atmos Environ 21:1–5

Bevan MAJ, Proctor CJ, Baker-Rogers J, Warren ND (1991) Exposure to carbon monoxide, respirable suspended particulates, and volatile organic compounds while commuting by bicycle. Environ Sci Technol 25:788–791

Bowden GJ, Maier HR, Dandy GC (2002) Optimal division of data for neural network models in water resources applications. Water Resour Res 38:2–1–2–11. doi:10.1029/2001WR000266

Božnar M, Lesjak M, Mlakar P (1993) A neural network-based method for the short time predictions of ambient SO_2 concentrations in highly polluted industrial areas of complex terrain. Atmos Environ 27B:221–230

Chaloulakou A, Grivas G, Spyrellis N (2003) Neural network and multiple regression models for PM10 prediction in Athens. A comparative assessment. J Air Waste Manage Assoc 53:1183–1190

Chelani AB, Gajghate DG, Hasan MZ (2002) Prediction of ambient PM_{10} and toxic metals using artificial neural networks. J Air Waste Manage Assoc 52:7, 805–810. doi:10.1080/10473289.2002.10470827

Comrie A (1997) Comparing neural networks and regression models for ozone forecasting. J Air Waste Manag Assoc 47:653–663

Corani G (2005) Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning. Ecol Model 185:513–529

Delfino RJ, Becklake MR, Hanley JA, Singh B (1994) Estimation of unmeasured particulate air pollution data for an epidemiological study of daily respiratory morbidity. Environ Res 67:20–38

Dockery DW, Schwartz J, Spengler JD (1992) Air pollution and daily mortality: association with particulates and acid aerosols. Environ Res 59:362–373

Fontes T, Silva LM, Silva MP, Barros N, Carvalho AC (2014) Can artificial neural networks be used to predict the origin of ozone episodes? Sci Total Environ 488–489:197–207

Fuller GW, Carslaw DC, Lodge HW (2002) An empirical approach for the prediction of daily mean PM10 concentrations. Atmos Environ 36:1431–1441

Gardner MW, Dorling SR (1998) Artificial neural networks (the multilayer perceptron)-a review of applications in the atmospheric sciences. Atmos Environ 32:2627–2636

Gardner MW, Dorling SR (1999) Neural network modeling and prediction of hourly NOx an NO_2 concentrations in urban air in London. Atmos Environ 31:709–719

Grivas G, Chaloulakou A (2006) Artificial neural network models for prediction of PM_{10} hourly concentrations in the Greater Area of Athens. Greece Atmos Environ 40:1216–1229

Hadjiiski L, Hopke PK (2000) Application of artificial neural networks to modeling and prediction of ambient ozone concentrations. J Air Waste Manage Assoc 50:849–901

Hadley A, Toumi R (2003) Assessing changes to the probability distribution of sulphur dioxide in the UK using a lognormal model. Atmos Environ 37:1461–1474

Hagan MT, Demuth HB, Beale M (1996) Neural network design. PWS Publishing Company, Boston

Hrust L, Klaic ZB, Krizan J, Antonic O, Hercog P (2009) Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations. Atmos Environ 43:5588–5596

Ibarra-Berastegi G, Madariaga I (2003) Traffic congestion and ozone precursor emissions in Bilbao (Spain). Environ Sci Pollut Res 10: 361–367

Ibarra-Berastegi G, Elias A, Madariaga I, Agirre E, Uria J (2001) Short term time forecasting of hourly ozone, NO_2 and NO levels by means of multiple linear regression modelling. Environ Sci Pollut Res 4:250

Jacobson MZ (1997) Development and application of a new air pollution modeling system II-aerosol module structure and design. Atmos Environ 31:131–144

Juda K (1989) Air pollution modelling. In: Cheremisinoff PN (ed) Encyclopedia of environmental control technology, vol 2, Air pollution control. Gulf Publishing Company, Houston, pp 83–134

Karppinen A, Kukkonen J, Elolahde T, Konttinen M, Koskentalo T (2000) A modelling system for predicting urban air pollution, comparison of model predictions with the data of an urban measurement network. Atmos Environ 34:3735–3743

Khare M, Sharma P (2002) Modelling urban vehicle emissions. WIT Press, Southampton, p 212

Khedairia S, Khadir MT (2012) Impact of clustered meteorological parameters on air pollutants concentrations in the region of Annaba. Alger Atmos Res 113:89–104

Kim M, Gilley JE (2008) Artificial neural network estimation of soil erosion and nutrient concentrations in runoff from land application areas. Comput Electron Agric 64:268–275

Kolehmainen M, Martikainen H, Hiltunen T, Ruuskanen J (2000) Forecasting air quality parameters using hybrid neural network modelling. Environ Monit Assess 65(1–2):277–286

Kolehmainen M, Martikainen H, Ruuskanen J (2001) Neural networks and periodic components used in air quality forecasting. Atmos Res 35:815–825

Kukkonen J, Partanen L, Karppinen A, Ruuskanen J, Junninen H, Kolehmainen M, Niska H, Dorling S, Chatterton T, Foxall R, Cawley G (2003) Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with a deterministic modelling system and measurements in central Helsinki. Atmos Environ 37:4539–4550

Kurt A, Gulbagci B, Karaca F, Alagha O (2008) An online air pollution forecasting system using neural networks. Environ Int 34:592–598

Laïd Y, Atek M, Oudjehane R, Filleul L, Baough L, Zidouni N, Boughedaoui M, Tessier J-F (2006) Health effects of PM_{10} air pollution in a low-income country: the case of Algiers. Int Tuberc Lung Dis 10(12):1406–1411

Lu WZ, Wang WJ, Fan HY, Leung AYT, Xu ZB, Lo SM, Wong JCK (2002) Prediction of pollutant levels in Causeway bay area of Hong Kong using an improved neural network model. J Environ Eng 128:1146–1157

Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ Model Softw 15:101–124

Maier HR, Morgan N, Chow CWK (2004) Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. Environ Model Softw 19:485–494

McKendry I (2000) Evaluation of artificial neural networks for fine particulate pollution (PM_{10} and $PM_{2.5}$) forecasting. J Air Waste Manage Assoc 52:1096–1101

Molina MJ, Molina LT (2004) Megacities and atmospheric pollution. J Air Waste Manage Assoc 54:644–680

Munir S, Habeebullah T, Seroji A, Morsy E, Mohammed A, Saud W, Abdou A, Awad A (2013) Modeling particulate matter concentrations in Makkah, applying a statistical modeling approach. Aerosol Air Qual Res 13:901–910

Papanastasiou DK, Melas D, Kioutsoukis I (2007) Development and assessment of neural network and multiple regression models in order to predict PM_{10} levels in a medium-sized Mediterranean city. Water Air Soil Pollut 182:325–334

Paschalidou AK, Karakitsios S, Kleanthous S, Kassomenos PA (2011) Forecasting hourly PM10 concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management. Environ Sci Pollut Res 18:316–327

Perez P, Trier A, Reyes J (2000) Prediction of $PM_{2.5}$ concentrations several hours in advance using neural networks in Santiago, Chile. Atmos Environ 34:1189–1196

Petersen RL (1997) A wind tunnel evaluation of methods for estimating surface rougness length at industrial facilities. Atmos Environ 31:45–57

Pires J, Martins F, Sousa S, Alvim-Ferraz M, Pereira M (2008) Prediction of the daily mean PM_{10} concentrations using linear models. Am J Environ Sci 4:445–453

Roy S (2012) Prediction of particulate matter concentrations using artificial neural network. Resour Environ 2:30–36

Roy S, Adhikari GR, Renaldy TA, Jha AK (2011) Development of multiple regression and neural network models for assessment of blasting dust at a large surface coal mine. Environ Sci Technol 4:284–301

Sabri G, Med Tarek K (2012) Combination of artificial neural network models for air quality predictions for the region of Annaba, Algeria. Int J Environ Stud 69:79–89

Schlink U, Dorling S, Pelikan E, Nunnari G, Cawley G, Junninen H, Greig M, Foxall R, Eben K, Chatterton T, Vondracek J, Richter M, Dostal M, Bertucco L, Kolehmainen M, Doyle M (2003) A rigorous inter-comparison of ground-level ozone predictions. Atmos Environ 37:3237–3253

Schwartz J, Dockery DW (1992) Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145:600–604

Seinfeld JH, Pandis SN (1998) Thermodynamics of aerosols. In: Atmospheric chemistry and physics. Wiley, New York

Tecer LH (2007) Prediction of SO_2 and PM concentrations in a coastal mining area (Zonguldak, Turkey) using an artificial neural network. Pol J Environ Stud 16:633–638

Touloumi G, Pocock SJ, Katsouyianni K, Trichopoulos D (1994) Short term effects of air pollution on daily mortality in Athens. A time series analysis. Int J Epidemiol 23:957–967

van Wijnen JH, Verhoeff AP, Jans HWA et al (1995) The exposure of cyclists, car drivers and pedestrians to traffic-related air pollutants. Int Arch Occup Environ Health 67:187–193

Venkatram A (1983) Uncertainty in predictions from air quality models. Bound-Layer Meteorol 27:185–196

Wang F, Fang CS, Zhang Y, Ma Z, Wang J (2013) Study on relationship between PM_{10} concentration and temperature condition in Longyan city. Appl Mech Mater 295–298:1556–1559

Willmott CJ (1981) On the validation of models. Phys Geogr 2:184–194

Zannetti P (1989) Simulating short–term, short-range air quality dispersion phenomena. In: Cheremisinoff PN (ed) Encyclopedia of environmental control technology, (Air Pollution Control), vol 2, 159–195