



Original article

Development and comparison of regression models and feedforward backpropagation neural network models to predict seasonal indoor PM_{2.5–10} and PM_{2.5} concentrations in naturally ventilated schools



Maher Elbayoumi*, Nor Azam Ramli, Noor Faizah Fitri Md Yusof

Clean Air Research Group, School of Civil Engineering, Universiti Sains Malaysia, Penang, Malaysia

ARTICLE INFO

Article history:

Received 23 March 2015

Received in revised form

24 May 2015

Accepted 25 May 2015

Available online 14 October 2015

Keywords:

Feedforward backpropagation

Indoor air quality

Multiple linear regression

Seasonal variations

ABSTRACT

A combination of multivariate statistical methods, including multiple linear regression (MLR) and feedforward backpropagation (FFBP) were used to evaluate the influence of seasons on the concentrations of indoor PM_{2.5–10} and PM_{2.5} in twelve naturally ventilated schools located in Gaza Strip, Palestine. Samples were collected by using hand held particulate matter sampler during fall, winter and spring from 2011 to 2012. Statistical results revealed that MLR models agree fairly well with the measured data with reasonable coefficients of determination (R^2) 0.58, 0.69 and 0.70 for indoor PM_{2.5} and 0.44, 0.56 and 0.57 for indoor PM_{2.5–10} during fall, winter and spring, respectively. The FFBP model results performed better than the MLR analysis in determining indoor PMs with R^2 values of 0.75, 0.78 and 0.79 for PM_{2.5} and 0.65, 0.73, and 0.78 for PM_{2.5–10} during fall, winter, and spring, respectively. The accuracy (R^2) models of the FFBP showed an improvement of 12.08%–25.56% and from 26.36% to 38.53% for prediction of indoor PM_{2.5} and PM_{2.5–10} compared to MLR models. In addition, FFBP models improved the accuracy by reducing the error (RMSE) as much as 19.35%, 7.41%, and 7.41% during fall, winter, and spring for prediction of indoor PM_{2.5}, and by 32.00%, 7.41%, and 32.00% during fall, winter, and spring, respectively for prediction of indoor PM_{2.5–10} compared with MLR. Results showed that the artificial neural network approach can be capable of accurately modeling indoor air quality in naturally ventilated buildings.

Copyright © 2015 Turkish National Committee for Air Pollution Research and Control. Production and hosting by Elsevier B.V. All rights reserved.

1. Introduction

Education is considered a crucial component of child development. Children generally spend seven hours or more every day inside school buildings. However, the regulation of indoor air quality (IAQ) in schools has been recently raised as an issue owing to various types of hazardous pollutants, such as particulate matters (PMs), volatile organic compounds, and formaldehydes. These pollutants accumulate in the school space due to overcrowding and inadequate ventilation (Janssen et al., 1999; Pegas et al., 2010; Almeida et al., 2011; Hassanvand et al., 2014).

Worldwide studies have been conducted on the effects of different indoor air pollutants on human health. Results of these

studies show that indoor air pollutants directly affect students' comfort and health (Pope et al., 2006; Stafoggia et al., 2013; Elbayoumi et al., 2014a). PMs a common indoor pollutant may cause adverse health effects that range from simple respiratory symptoms to morbidity and mortality depending on duration of exposure and concentration of pollutants (Jedrychowski et al., 2013; Pope III et al., 2013). Thus, to ensure a good health of students, accurate monitoring and prediction techniques are necessary for regulating IAQ in schools.

Several studies have investigated diurnal and seasonal PM concentration in different types of buildings. However, these studies mainly focused on monitoring but not on predicting IAQ inside buildings (Almeida et al., 2011; Braniš and Safránek, 2011; Buonanno et al., 2013). In addition, direct and long-term measurements of PM concentrations are impractical as IAQ heavily relies on local conditions such as weather changes and seasonal variations (Elbayoumi et al., 2013). In the absence of effective and efficient means to directly measure indoor PM from school buildings, mathematical prediction models might be a good alternative

* Corresponding author. Tel.: +60 4 5996227; fax: +60 4 5941009.

E-mail address: maher@usm.my (M. Elbayoumi).

Peer review under responsibility of Turkish National Committee for Air Pollution Research and Control.

to provide reasonably accurate estimates. Furthermore, modeling indoor air pollution identifies the relative contribution between outdoor source, meteorology, atmospheric concentrations, deposition, and other factors. Thus, this help in determination of the effectiveness of remediation strategies and the simulation of future scenarios (Russo et al., 2015). Particularly, time-lagged models (i.e. models that can forecast concentrations), for example, one hour ahead, or 24 h ahead, are found to give reliable predictions and allow effective alert procedures in monitoring areas (Mishra and Goyal, 2015).

Multiple linear regression (MLR) analysis is a popular technique that relates the response of a dependent variable to several independent variables. Previous studies have used MLR to establish a quantitative relationship between various predictor variables (e.g., weather, outdoor source, ventilation system, and number of students) and indoor pollutant concentrations (Chaloulakou et al., 2001; Adar et al., 2008; Braniš and Šafránek, 2011; Elbayoumi et al., 2014b). This relationship is used to understand which predictors exert the greatest effect and to forecast future values of the equation response when only the predictors and the direction of their effects are known. In spite of its success, the MLR approach can face difficulties which are known as multicollinearity problem particularly during high correlation between independent variables and the non-linear relationships problem between the variables where MLR cannot adequately estimated any non-linear relationship (Al-Alawi et al., 2008). Thus, several non-linear multivariate statistical methods (e.g., artificial neural networks, ANNs) that can approximate any non-linear relationship have been developed.

Artificial neural networks (ANN) are among the favored techniques in predicting a complex system. Several studies have demonstrated that the performance of ANN is generally superior to that of traditional statistical methods and deterministic modeling systems because of its computational efficiency, generalization ability, and limited need for prior knowledge about the modeling process structure (Nejadkoorki and Baroutian, 2011; Lal and Tripathy, 2012; Elangasinghe et al., 2014). The applications of ANN models in the field of air pollution forecasting increased in the past decade (Grivas and Chaloulakou, 2006; Sousa et al., 2007). ANN has actually been proposed to predict atmospheric concentrations of NO (Gardner and Dorling, 1999; Mishra and Goyal, 2015), ozone (Wang et al., 2003), benzene (Viotti et al., 2002), SO₂ (Sofuoglu et al., 2006), PM₁₀ (Ul-Saufie et al., 2013; Russo et al., 2015) and PM_{2.5} (Mishra et al., 2015). However, the literature shows that little attention has been paid to forecasting IAQ within buildings using either ANN or MLR.

Air quality varies from season to season because atmospheric dynamics and meteorological conditions play an important role in governing the fate of air pollutants. Several studies confirmed that the IAQ is dependent on outdoor concentrations and local conditions, such as weather changes and seasonal variations (Roberts, 2004; Kam et al., 2011). Thus, seasonal IAQ predictive model should be designed for the environment aspects. The overarching goal of this project is to present the results of the application of multivariate regression analysis and feedforward backpropagation (FFBP) in predicting seasonal indoor PM_{2.5–10} and PM_{2.5} concentration as a function of meteorological parameters and other pollutants concentration from naturally ventilated school buildings.

2. Materials and methods

2.1. Study area

Gaza Strip (365 km²) is located on the eastern coast of the Mediterranean Sea, between longitudes 34° 15' and 35° 40' east

and latitudes 29° 30' and 23° 15' north. The average density is 4400 inhabitants/km², which is estimated to quickly increase to 5835 inhabitants/km² in 2020 (UNCT, 2012). Geographically, the location among three transitional zones, namely, the sub-humid coastal zone of Israel, the semiarid plains of the northern Negev desert, and the arid Sinai desert zone of Egypt, is an important factor in shaping climate diversity, as illustrated in Fig. 1 (Shomar et al., 2008; PMD, 2012). Climatically, winter in Gaza strip is characterized by high humidity and dominated by rainfall. In contrary to winter, summer season is characterized by lack of wet precipitation and high temperature and humidity. During spring season unsettled rainfall which is associated with North African cyclones is occurred between March and April. However, the remaining months of spring season (May–June) are characterized by high humidity and the lack of wet precipitation. The last season, fall, is usually characterized by an abrupt summer-type weather in the first month (September) and unsettled winter weather during November and December months. The average daily temperature fluctuates from 24 °C in summer to 15 °C in winter. The daily relative humidity varies between 62.5% in the daytime and 83.4% at night during summer and between 51.6% in the daytime and 81.3% at night during winter. The monthly average wind speed for Gaza is 3 m/s (Koçak et al., 2010; PMD, 2012).

2.2. Description of sampling locations

The concentrations of pollutants were monitored at schools in north, middle, and south Gaza Strip from October 2011 to May 2012 to cover the three season's fall, winter, and spring. However, during July and August months schools are closed on account of summer holidays. As a result sampling was not carried out in those months. Four schools served as a sample in each location. The sampled schools were purposely selected to reflect the diverse nature of human and vehicular activities. The description of sampling locations is summarized in Fig. 1 and Table 1. In each selected school, three representative classrooms were selected for three sampling days. Sampling was conducted both inside and outside the selected classrooms during the study activities.

2.3. Pollutants measurements and instrumentation

The measurements were taken in each site for three consecutive days during school hours, specifically, from 7:00 A.M. to 12:00 P.M. during winter and spring and from 12:00 P.M. to 5:00 P.M. during fall. The samplers were placed inside the classroom opposite the blackboard at least 1 m from the wall and at least 1.5 m from the floor (Blondeau et al., 2005; WHO, 2011). For outdoor sampling, the samplers were placed at the front side of the building, usually near the playground area. The mass concentration of particles (PM_{2.5} and PM_{2.5–10}) was monitored by using a handheld optical particle counter (HAL-HPC300). The monitor measures particulate size by laser light scattering. Air with multiple particle sizes passes through a flat laser beam produced by an ultra-low maintenance laser diode. A three-channel pulse height analyzer for size classification detects the scattering signals. Quality control checks, including inspection of the instruments and zero concentration check, were performed every three days using Zero-Count Filter (HAL, 2012). A Kanomax IAQ Monitor was used to measure CO and CO₂ concentrations. The IAQ monitor was calibrated in the Islamic University–Gaza laboratory for CO₂ and CO in accordance with manufacturer instructions. Ventilation rate was calculated using the indoor concentration of CO₂ as a surrogate of ventilation levels per occupant (Kulshreshtha and Khare, 2011; WHO, 2011).

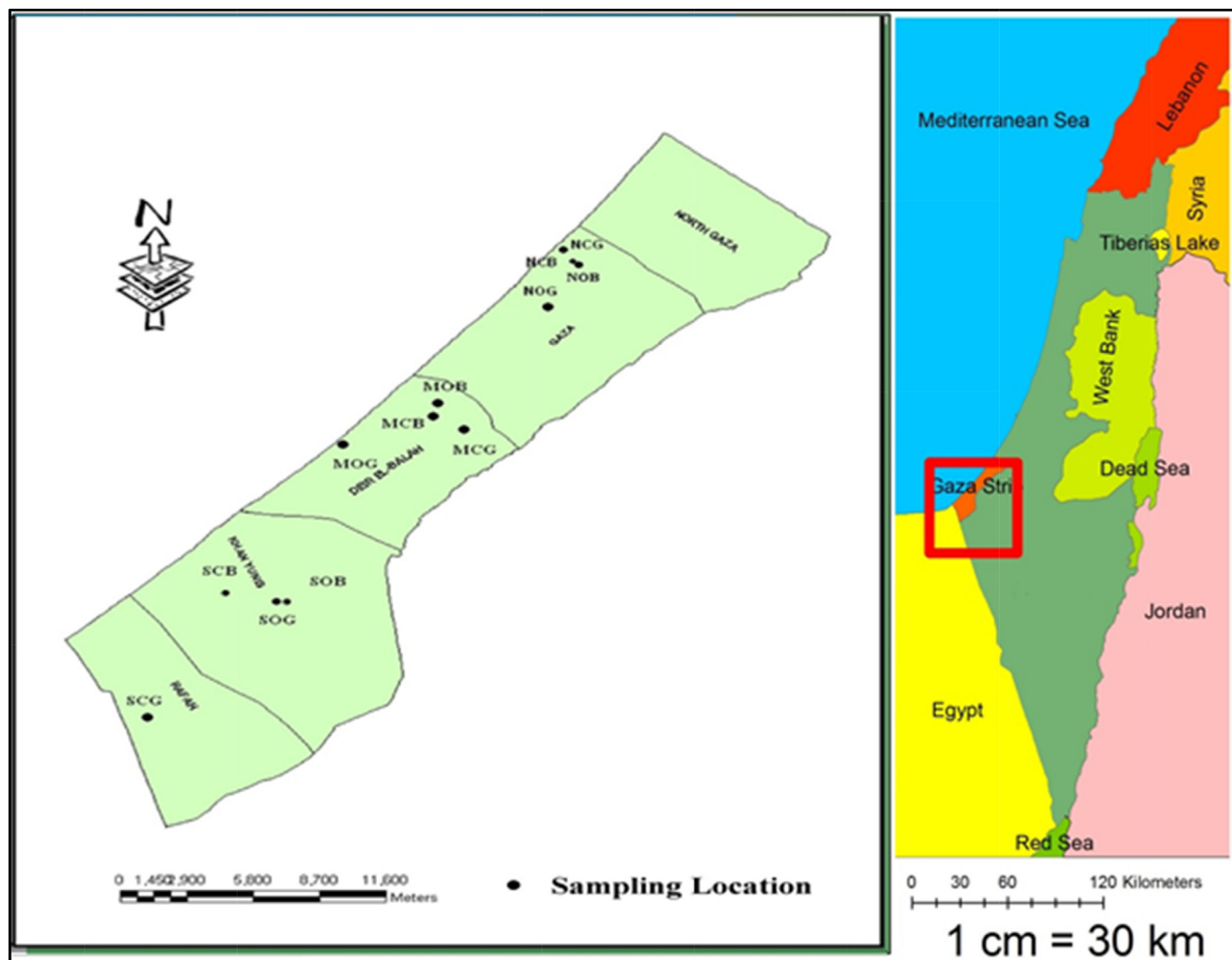


Fig. 1. Map of Gaza strip and the monitoring schools.

2.4. Meteorological data

Meteorological parameters (temperature, humidity and wind speed) were simultaneously recorded along with pollutants measurements. A Kanomax IAQ Monitor was used to measure temperature and relative humidity, while Smart Sensor Electronic Anemometer was used for wind speed.

2.5. Data interpretation

A vital step in the development of a forecast indoor air model is the choice of input parameters (Jef et al., 2005). A number of statistical methods can be applied to select the most appropriate set of input parameters for FFBP; among such methods are stepwise regression (FSR), principal component analysis, and cluster analysis

Table 1
Characteristic of monitoring schools.

Location	Code on the map	Number of students	Distance from main road (m)	Width of main road (m)	Description
North	NOB	623	43	10	Medium population density/medium traffic volume
	NCG	1183	50	10	High population density/medium traffic volume
	NCB	1066	30	20	Medium population density/high traffic volume
	NOG	883	58	20	Medium population density/medium traffic volume
Middle	MCB	733	43	10	High population density/medium traffic volume
	MCG	903	50	20	High population density/medium traffic volume
	MOG	1024	50	16	Low population density/very low traffic volume
	MOB	712	65	10	High population density/medium traffic volume
South	SCG	578	55	12	High population density/medium traffic volume
	SCB	729	50	10	High population density/medium traffic volume
	SOG	1132	40	20	Medium population density/high traffic volume
	SOB	1448	55	30	Medium population density/high traffic volume

(Wilks, 2011). These methods are important in reducing the number of input variables into the models, thus considerably diminishing redundant information, instabilities, and over-fitting. In this study, the criteria that used to select the input parameters was performed in two stages. In the first stage, univariate analysis was used to determine factors affecting indoor PM levels taking into consideration the meteorological conditions and the pollutants measured in the monitoring schools. In the second stage, the environmental factors that were found to be associated with indoor PMs were incorporated into FSR technique by using the spatial (school) and seasonal data set. During FSR, which starts with the variable that is most correlated with the target, additional variables are added which most accurately predict the target with the previously selected variables (Wilks, 2011; Russo et al., 2015). The procedure stops when any new variable does not significantly reduce the prediction error. Significance is measured by a partial F-test applied at 5% (Wilks, 2011). All 12 potential predictors for indoor PM_{2.5–10} and PM_{2.5} were first considered. The use of the FSR for each monitored school in every season reduced complexity by retaining substantially less variables. Then, the most common variables among the schools and seasons were used by MLR and ANN approaches.

For calibration and validation purposes cross-validation is used for each season data for both MLR and ANN techniques. In other words, each season data set is divided to 3 parts 70%, 15% and 15% for construction of the neural network, testing and model validation, respectively. To reduce variability, 10 rounds of cross validation are performed and the validation results are averaged over the rounds. With such procedures, one can ascertain the stability and forecasting capability of the model as well as limiting problems such as overfitting (Russo et al., 2015). After the calibration and testing procedure, the models are used to produce forecasts for the daily average of PMs concentration. For this purpose 15% of data set is left out in order to be used for evaluation of models performance. Finally, the performance indicators were used to compare the observed values with the predicted values. Data analysis was carried out using the statistical software Statistical Package for Social Science (SPSS) version 22 and MATLAB version 10.

2.6. Modeling details

2.6.1. MLR

Stepwise multiple regressions were carried out for PM_{2.5–10} and PM_{2.5}. The results were checked for multicollinearity by examining the variance inflation factors (VIFs) of the predictor variables. To check for first-order autocorrelation problem, Durbin–Watson statistic was used. MLR can be expressed according to Equation (1).

$$y = b_0 + b_1x_{1i} + b_2x_{2i} + \dots + b_kx_{ki} + \varepsilon \quad (1)$$

where b_k is the regression coefficients, x_k is the explanatory variables, $i = 1, 2, \dots, k$, and ε is the stochastic error associated with the regression (Agirre-Basurko et al., 2006). To verify the adequacy of the statistical model, residuals (or errors) were checked to ensure the normal distribution of data with zero mean and constant variance (Al-Alawi et al., 2008).

2.6.2. FFBP

Multilayer perceptron (MLP) is one of most common neural network architecture with feedforward backpropagation network topologies. It offers an efficient learning environment, minimizes error between the target and the obtained values (Ul-Saufie et al., 2013). As shown in Fig. 2, the network of FFBP usually consists of an input layer, several hidden layers, and an output layer. Each

layer consists of several operating neurons, each of which is connected to every neuron in nearby layers through adaptable synaptic weights that determine the strength of the relationship between two connected neurons. In each layer, every neuron sums all the inputs that have been received from previous layers and forms the neuron output through predefined activation or transfer function. Learning is defined as the ability of a network to change weights by backpropagation algorithm through two phases. In the forward phase, the training data set is propagated through the hidden layer and comes out of the neural network through the output layer. The output values are then compared with actual target output values. The error between the output layer and the actual values are calculated and propagated back toward the hidden layer (Ul-Saufie et al., 2013). In the backward phase, derivatives of network error (with respect to the networks) are fed back to the network and used to adjust the weights to reduce errors with each iteration, thus improving the FFBP models and prompting the neural model to produce the desired outputs.

The net architecture in this study consists of three-layer perceptron model. The first input layer contains the input variables that were selected by FSR method to reduce complexity by retaining substantially fewer variables. The second architecture layer is the hidden layer. The common problems in hidden layer architecture are to identify the number of hidden layers, to identify neurons values and to choose the suitable activation function. The optimum number of neurons is important because too few neurons will contribute to under-fitting, whereas too many neurons lead to overfitting. A fixed scientific solution for the design of an optimal ANN model does not exist. Using Equation (2), the initial number of neurons was 9 neurons, and the number was increased until a relatively stable and optimal value was achieved (Yang et al., 2005).

$$n_h = 2 \times n_i + 1 \quad (2)$$

where n_i is the number of input neurons, and n_h is the number of hidden neurons. The second problem in building the architecture of the hidden layer is deciding suitable transfer function. According to Kriesel (2007) the transfer function in the hidden layers, must use a nonlinear transfer function (otherwise the result end up with only linear separable solutions). Therefore, the optimal activation functions are obtained using sigmoid transfer function (Ul-Saufie et al., 2013). The most common sigmoid functions are linear (purelin), log-sigmoid (logsig) and hyperbolic tangent sigmoid (tansig). The logistic function will generate values close to 0.0. If the argument of the function is substantially negative, the output of hidden neuron will be close to zero, and as a result lowering the learning rate for all subsequent weights which will almost stop learning. Tansig can produce both positive and negative values which will generate a value close to -1.0 , and thus will maintain learning (Kriesel, 2007). In this study, Tansig and purelin functions were used to build the architecture of the hidden layer. The third architecture layer is the output layer which consists of the target of the forecasting model. In this research, both indoor PM_{2.5–10} and PM_{2.5} were used as output variable.

2.6.3. Performance indicators

The analysis of prediction performance typically involves calculation of errors between observed and predicted values. In this study five performance indicators were used which are normalized absolute error (NAE), root mean square error (RMSE), prediction accuracy (PA), coefficient of determination (R^2) and index of agreement (IA). Normalized absolute error (NAE) and root mean square error (RMSE) were used to find the error of the model where value closer to 0 indicated a better model. Meanwhile, the other

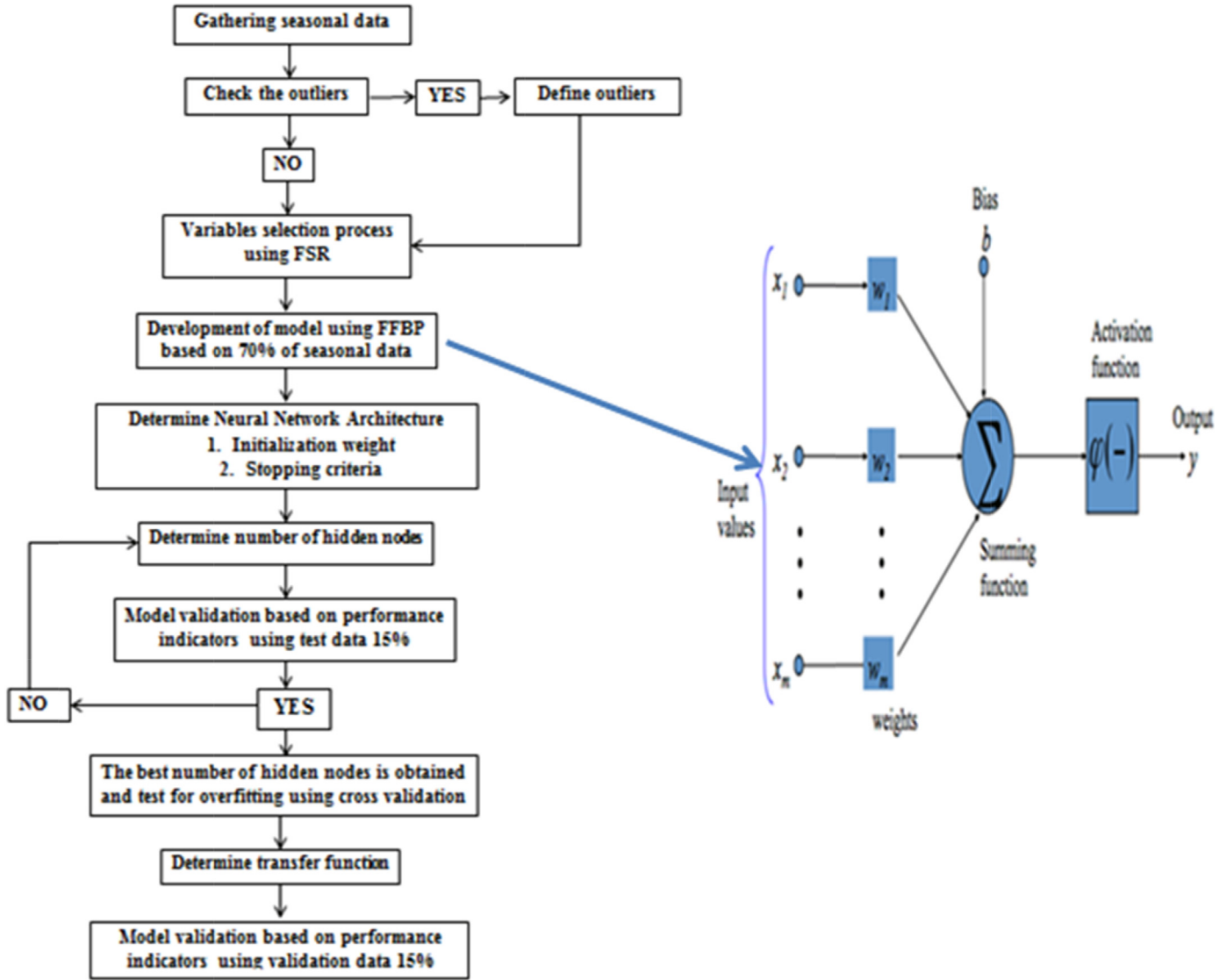


Fig. 2. The scheme and architecture of FFBP models of the predictive monitoring and the diagnosis of indoor air indoor PM_{2.5–10} and PM_{2.5} concentrations.

three performance indicators, i.e. index of agreement (IA), prediction accuracy (PA) and coefficient of determination (R^2) were used to check the accuracy of the model result, where a higher accuracy is given by value closer to 1. The equations of performance indicators are given in Equations (3)–(7) (Karppinen et al., 2000; Gervasi, 2008):

The Normalized Absolute Error (NAE).

$$NAE = \frac{\sum_{i=1}^N |P_i - O_i|}{\sum_{i=1}^N O_i} \quad (3)$$

The Root Mean Square Error (RMSE).

$$RMSE = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (P_i - O_i)^2} \quad (4)$$

The Coefficient of Determination (R^2)

$$R^2 = \left(\frac{\sum_{i=1}^N (P_i - \bar{P})(O_i - \bar{O})}{N \cdot S_{pred} \cdot S_{obs}} \right)^2 \quad (5)$$

The Prediction Accuracy (PA)

$$PA = \frac{\sum_{i=1}^N (P_i - \bar{P})}{\sum_{i=1}^N (O_i - \bar{O})} \quad (6)$$

The Index of Agreement (IA)

$$IA = 1 - \left[\frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \right] \quad (7)$$

Where N is the number of sample, O_i is the indoor PM_{2.5–10} or PM_{2.5} concentrations values and P_i is the predicted PM_{2.5–10} or PM_{2.5}, \bar{P} is the average of predicted indoor PM_{2.5–10} or PM_{2.5}, \bar{O} is the average

of measured indoor $PM_{2.5-10}$ or $PM_{2.5}$. S_{pred} is a standard deviation of the predicted $PM_{2.5-10}$ or $PM_{2.5}$. S_{obs} is a standard deviation of the observed $PM_{2.5-10}$ or $PM_{2.5}$.

3. Result and discussion

3.1. Seasonal characteristic of particulate matter

The data that reflect the seasonal data at the three locations are arranged in a matrix format and summarized in Table 2. The average six-hour of indoor and outdoor $PM_{2.5-10}$ concentrations for all the schools during the study period were $350 \pm 197 \mu\text{g}/\text{m}^3$ and $150 \pm 98 \mu\text{g}/\text{m}^3$, respectively. Meanwhile, the average indoor and outdoor concentrations for $PM_{2.5}$ were $104 \pm 85 \mu\text{g}/\text{m}^3$ and $61 \pm 51 \mu\text{g}/\text{m}^3$, respectively as shown in Table 2. A comparison between the averages of $PM_{2.5-10}$ and $PM_{2.5}$ concentrations to WHO guidelines (50 and $25 \mu\text{g}/\text{m}^3$ for PM_{10} and $PM_{2.5}$, respectively) indicates that the indoor and outdoor concentrations of $PM_{2.5-10}$ of all the schools (100%) exceeded the standards. Meanwhile, the indoor and outdoor concentration of $PM_{2.5}$ showed some variation. The higher outdoor concentration may be attributed to the presence of unpaved playgrounds in the monitored schools and to the attribution of long range aerosol transportation. The presence of large spaces of unpaved playgrounds around the school buildings may increase the outdoor and indoor $PM_{2.5-10}$ and $PM_{2.5}$ concentrations by resuspension of deposited coarse particle or delay in the deposition process due to student's activity (Ozkaynak et al., 1996; Blondeau et al., 2005; Diapouli et al., 2007; Stranger et al., 2008; Heudorf et al., 2009; Goyal and Khare, 2011; Chithra and Nagendra, 2012). In addition, the location of Gaza Strip between the Sinai and the Negev deserts and in front of the Mediterranean Sea affects the aerosol concentrations governed by several important phenomena, including long-range aerosol transportation (Matvev et al., 2002), seasonal dust storms (Dayan et al., 1991; Koçak et al., 2010; Zereini and Wiseman, 2010), sea salt aerosol formation (Krom et al., 2004). Furthermore, the lower rainfall rates in the area especially during fall and spring seasons promote the elevation of road dust in unpaved playgrounds which become as a permanent reservoir for fugitive dust (Zereini and Wiseman, 2010).

During the study period, the average indoor PM concentrations were greater than the average outdoor PM concentration. This was in good agreement with the previous studies which reported that the indoor PM levels were higher than outdoor ones which due to student's activities including writing on the blackboard using chalks, erasing chalks using duster and cleaning habits (Ramírez et al., 2012; Habil et al., 2013; Zwoździak et al., 2013; Razali et al., 2015). Several studies revealed that the penetration factors are higher for NVBs than for mechanically-ventilated buildings because NVBs buildings have windows, doors, ventilators, cracks and leaks in the building envelope (Branis et al., 2005; Chen and Zhao, 2011). Furthermore, different parameters may directly influence and increase indoor pollutants level such as differences in building envelope tightness and seasonal effects (Poupard et al., 2005; Goyal and Khare, 2009), and building design (Ashmore and Dimitroulopoulou, 2009).

The $PM_{2.5-10}$ and $PM_{2.5}$ levels exhibit seasonal variation during the monitoring period as illustrated in Fig. 3. The mean indoor and outdoor $PM_{2.5}$ concentrations during winter (192 and $122 \mu\text{g}/\text{m}^3$, respectively) were much higher than $PM_{2.5}$ levels during fall (55 and $26 \mu\text{g}/\text{m}^3$, respectively) and spring (59 and $20 \mu\text{g}/\text{m}^3$, respectively). The three seasons displayed statistical differences ($p < 0.001$) in mean $PM_{2.5}$ by using ANOVA and Duncan multiple range test. Further, the mean outdoor $PM_{2.5-10}$ concentration during winter ($125 \mu\text{g}/\text{m}^3$) was much higher than the $PM_{2.5-10}$ levels during spring ($83 \mu\text{g}/\text{m}^3$) and fall ($71 \mu\text{g}/\text{m}^3$). Higher outdoor concentrations in

Table 2
Average values of environmental data during the three seasons in the three locations in Gaza strip.

Location	$PM_{2.5}$ (in) ($\mu\text{g}/\text{m}^3$)	$PM_{2.5}$ (out) ($\mu\text{g}/\text{m}^3$)	$PM_{2.5-10}$ (in) ($\mu\text{g}/\text{m}^3$)	$PM_{2.5-10}$ (out) ($\mu\text{g}/\text{m}^3$)	CO (in) (ppm)	CO (out) (ppm)	CO ₂ (in) (ppm)	CO ₂ (out) (ppm)	RH (in) (%)	RH (out) (%)	Temp (in) (°C)	Temp (out) (°C)	VR (l/s, person)	WS (m/s)
Fall														
North	47.85 ± 15.52	22.46 ± 6.53	255.01 ± 92.81	52.95 ± 24.42	0.34 ± 0.34	0.52 ± 0.55	719.14 ± 75.00	380.49 ± 31.09	58.83 ± 8.01	57.86 ± 7.37	27.42 ± 1.55	27.65 ± 1.48	9.54 ± 4.48	3.56 ± 1.39
Middle	67.44 ± 32.46	33.61 ± 14.50	350.13 ± 139.95	91.34 ± 38.65	0.40 ± 0.38	0.49 ± 0.42	686.17 ± 142.70	348.23 ± 19.62	62.58 ± 4.28	58.99 ± 3.95	27.43 ± 1.06	28.33 ± 1.32	10.98 ± 7.04	3.40 ± 1.58
South	50.30 ± 15.67	23.57 ± 7.82	309.53 ± 88.94	70.50 ± 18.85	0.37 ± 0.37	0.47 ± 0.42	728.17 ± 106.19	387.5 ± 41.70	59.69 ± 2.85	59.41 ± 2.545	26.34 ± 1.48	26.47 ± 1.92	9.11 ± 2.87	3.69 ± 1.53
Winter														
North	188.44 ± 90.53	133.30 ± 63.68	278.73 ± 141.81	106.49 ± 78.96	1.18 ± 0.84	1.42 ± 0.98	1248.37 ± 284.76	607.30 ± 58.37	61.17 ± 9.69	57.27 ± 16.04	14.58 ± 1.54	14.11 ± 2.59	6.40 ± 3.48	2.79 ± 2.30
Middle	215.19 ± 93.73	142.67 ± 68.51	319.08 ± 191.07	118.79 ± 75.89	1.30 ± 1.15	1.45 ± 1.17	1175.65 ± 327.09	584.91 ± 79.35	66.35 ± 8.22	67.02 ± 11.32	14.48 ± 1.61	14.00 ± 2.40	7.20 ± 3.30	3.27 ± 2.18
South	190.17 ± 66.06	128.39 ± 33.56	286.05 ± 217.94	115.09 ± 108.14	1.14 ± 0.99	1.28 ± 1.04	1259.92 ± 367.97	614.74 ± 50.28	64.87 ± 13.54	62.61 ± 13.46	15.06 ± 2.08	14.29 ± 2.17	6.37 ± 2.77	3.41 ± 2.97
Spring														
North	56.50 ± 27.37	19.68 ± 9.62	106.70 ± 38.18	71.45 ± 21.84	0.85 ± 0.43	1.03 ± 0.51	971.22 ± 140.43	491.36 ± 35.17	74.86 ± 21.12	68.92 ± 20.86	19.83 ± 4.09	18.96 ± 4.42	7.72 ± 2.86	2.70 ± 1.81
Middle	63.65 ± 28.97	21.43 ± 10.30	155.10 ± 111.68	79.81 ± 48.48	0.74 ± 0.54	0.87 ± 0.63	922.17 ± 175.39	467.74 ± 41.65	82.86 ± 14.63	76.67 ± 17.05	18.8±2.44	17.87 ± 3.06	8.64 ± 3.26	2.77 ± 2.28
South	56.09 ± 16.48	19.26 ± 5.39	149.34 ± 80.58	95.13 ± 43.35	0.79 ± 0.48	1.09 ± 1.28	977.12 ± 179.68	502.90 ± 34.26	80.21 ± 12.90	73.84 ± 15.41	18.70 ± 2.75	17.73 ± 3.44	7.90 ± 2.61	2.48 ± 2.03

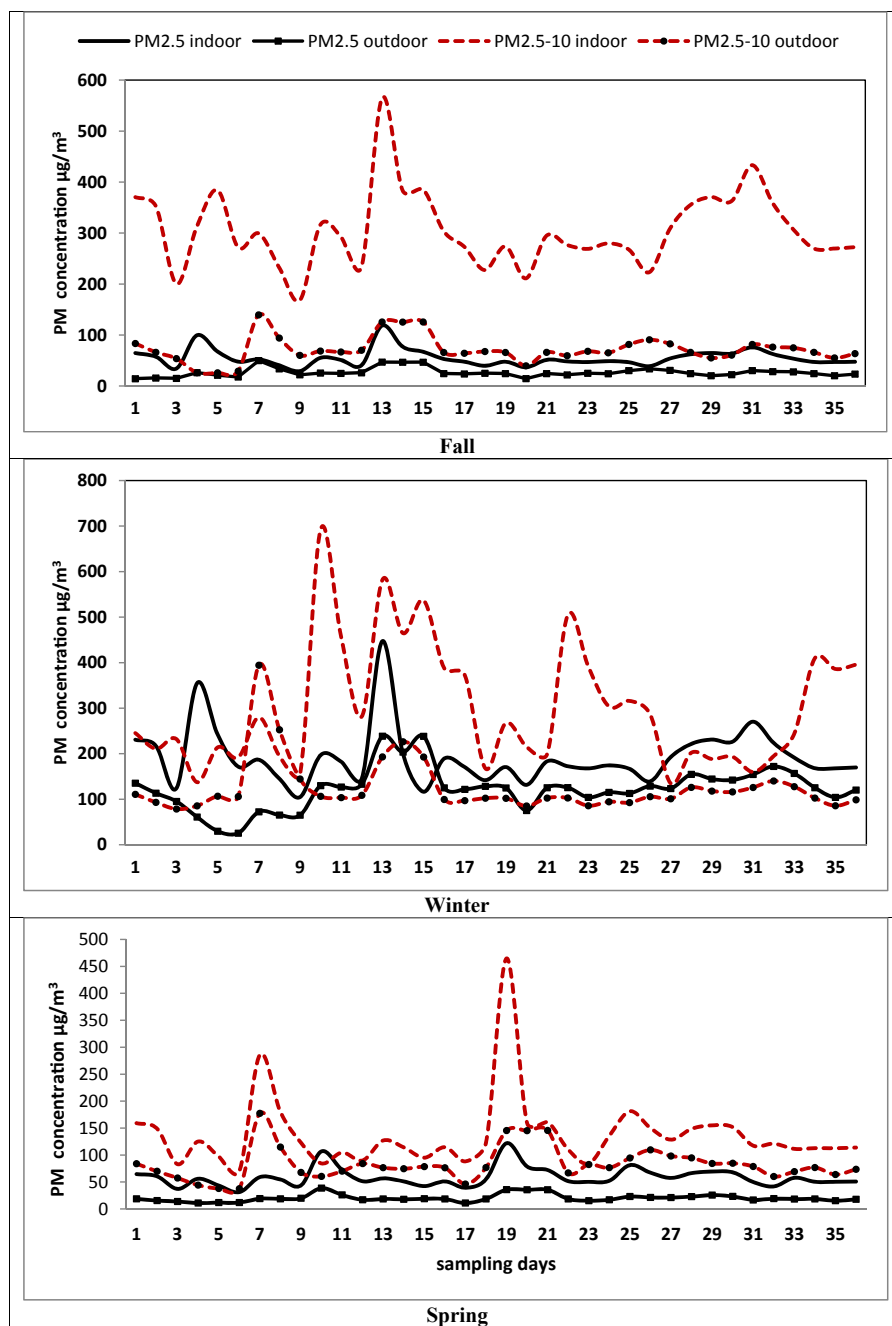


Fig. 3. Seasonal variations of $PM_{2.5-10}$ and $PM_{2.5}$.

winter may be due to influence of meteorological factors. Different studies show that average particle matter concentration tend to be higher in cold seasons, the seasons with the lowest ventilation capability, due to lower temperatures, high relative humidity, available organics emitted from vehicles and decreased atmospheric mixing height (Arkouli et al., 2010; Somuri, 2011). These variations indicated that the pollutants concentrations and emissions during different seasons must be obtained in IAQ models.

3.2. $PM_{2.5-10}$ and $PM_{2.5}$ seasonal forecast

IAQ can be influenced by many factors such as indoor source, outdoor concentration, air exchange rate, particle penetration factor, particle deposition rate, local conditions, such as weather

changes and seasonal variations (Chen and Zhao, 2011). Therefore, bivariate correlation was used to identify the factors that may influence the seasonal indoor PM concentrations.

The value of the correlation coefficient (r) between the indoor and outdoor data can be used as an indicator of the degree to which PM measured indoors is attributed to the infiltration from the outdoors. A strong relationship exists between indoor and outdoor $PM_{2.5}$ concentrations during fall, winter, and spring ($r = 0.79, 0.78$, and 0.89 for fall, winter, and spring, respectively), as shown in Table S1. In addition, a strong relationship exists between indoor and outdoor $PM_{2.5-10}$ concentrations during fall, winter, and spring ($r = 0.64, 0.73$, and 0.71 for fall, winter, and spring, respectively). Outdoor particle concentrations were identified as an important factor that influences the concentration levels of indoor particles;

several studies have detected a positive correlation between ambient and indoor particle concentrations (Branis et al., 2005; Poupard et al., 2005).

Air quality varies from season to season because atmospheric dynamics and meteorological conditions play an important role in governing the fate of air pollutants. Thus, the inter-correlation between the PM concentration, CO, CO₂, ventilation rate, temperature, relative humidity, and wind speed was explored, as shown in Table S1. Indoor PM_{2.5} concentration is negatively correlated with ventilation rate and positively correlated with wind speed during fall. The increment of air exchange rates may increase the penetration of pollutants and increase the dilution process, the ex-filtration, and the deposition rate. This finding is similar with Chao and Wong (2002) and Alshitawi et al. (2009) studies who investigated the influence of air change rate on the PM_{2.5–10} indoor/outdoor (I/O) ratios in different microenvironments. As found during fall, PM_{2.5} correlates positively with wind speed during winter, but the trend of ventilation rate during winter is different from that of the rate during fall.

As seen from Table S1, PM_{2.5–10} and PM_{2.5} concentrations were negatively correlated with temperature during winter. The inverse relationship between indoor PM and temperature is due to the temperature differences between indoor and outdoor spaces. When the indoor temperature is higher than the outdoor temperature, air is forced out of the building and dilutes the indoor concentration (Milner et al., 2004). Decreasing and increasing outdoor temperature influence weather stability and thus disturb outdoor PM concentration. High temperatures increase evaporation processes as well as increase the mixing height. This situation increases the dispersion rate of surface particles towards the upper part of the atmosphere which reduces the PM concentration. However, low temperature creates a low mixing height, and prevents the dispersion of pollutants which in turns increases the PM concentrations. Different studies reported an inverse relationship between ambient temperature and outdoor PM species concentration (Monn et al., 1995; Chan, 2002; Onat and Stakeeva, 2013). Relative humidity was also positively correlated with PM_{2.5} and PM_{2.5–10} during the three seasons. Fromme et al. (2007) reported a significant increase in indoor PM_{2.5} by 1.7 µg/m³ per increase in humidity by 10%. The same result is presented by several studies (Branis et al., 2005; Razali et al., 2015).

3.3. MLR modeling development

The results from the previous sections indicated that seasonal variations were observed in the levels of PM, CO, CO₂, and meteorological variables. Thus, a seasonal model may better capture the characteristics of seasonal variation.

MLR modeling (stepwise method) of indoor PM_{2.5–10} and PM_{2.5} concentrations was conducted to find a predictive equation by using outdoor pollutant variables, such as PM_{2.5–10}, PM_{2.5}, CO, CO₂, and meteorological factors with consideration for regression assumptions approximately satisfied. The coefficient of determination R² provides the proportion of the variation in the PM_{2.5–10} and

PM_{2.5} concentrations, as explained by the independent variables in the models. Table 3 shows that when the two best variables are fitted to the fall PM_{2.5} data, the values of the R² are approximately 0.72. Thus, approximately 72% of the variation in the PM_{2.5} concentrations can be explained by the independent variables, as listed in Table 3. For PM_{2.5} winter and spring data, the R² values are approximately 0.59 and 0.54, respectively, as reported in Table 3. Therefore, approximately 59% of the variation in the PM_{2.5} concentrations can be explained by the five independent variables listed in the table for winter data.

Table 3 shows that when the three best variables are fitted to the fall PM_{2.5–10} data, the values of the R² are approximately 0.47, which is attributed to the weak correlation between indoor and outdoor PM_{2.5–10}. The winter and spring models showed better R² compared with those in the fall season. The R² values are 0.59 and 0.54 for winter and spring seasons, respectively, as reported in Table 3. The common variables in seasonal PM_{2.5} and PM_{2.5–10} other than outdoor PM species are temperature, relative humidity, wind speed, and ventilation rate. The common variables in MLR models are outdoor PM_{2.5–10}, outdoor PM_{2.5}, RH, and temperature and ventilation rate. These aforementioned variables vary by season and the result is in concurrence with the findings of Gold et al. (1996) and Diapoulis et al. (2007).

The coefficients of the regressions for all models were highly and statistically significant ($P < 0.01$). The residual distributions were approximately normal, with zero means and no detectable serial correlation; these features indicate adequate model fit. Furthermore, collinearity problems in MLR might be diagnosed by using the VIF and tolerance indicator. The VIF values for both MLR model PM_{2.5–10} and PM_{2.5} ranged between 1 and 1.8, which indicate several associations between predictor variables. However, these factors are insufficient to cause problems. In addition, the tolerance values for the variables in both MLR models are higher than 0.3. In accordance with the findings of Field (2009), the tolerance value must be smaller than 0.1 to indicate a multicollinearity problem. The Durbin–Watson test statistic is used to test autocorrelations between errors. Table 3 shows the values of Durbin–Watson test statistics, which are close to 2 and indicate that the assumption is satisfied.

3.4. MLR models validation

The relationship between seasonally measured and predicted PM_{2.5–10} and PM_{2.5} concentrations is illustrated in Table 4. The calculated R² for PM_{2.5} during the three seasons ranged from 0.58 to 0.70. The R² for the three season results shows that independent variables are fitted to PM_{2.5} data and can explain approximately 58%–70% of the variation in the indoor PM_{2.5}. In addition, the values of RMSE ranged from 1.3% for winter model to 1.7% for the fall model. A comparison among the performances of the three PM_{2.5} models indicates that the spring MLR model produced the highest R² values. The calculated R² for PM_{2.5–10} during the three seasons ranged from 0.44 to 0.57. Thus, approximately 44%–57% of the variation in the indoor PM_{2.5–10} concentrations can be

Table 3
Results of the MLR analysis for indoor PM_{2.5} and PM_{2.5–10}.

Pollutant	Season	R ²	Equation	Range of VIF	Durbin–Watson
PM _{2.5}	Fall	0.72	PM _{2.5} (in) = 0.02 + 0.98 [PM _{2.5} (out)] – 0.08RH (in)	1.046–1.046	1.247
	Winter	0.59	PM _{2.5} (in) = –0.01 + 0.58 [PM _{2.5} (out)] + 0.15 [PM _{2.5–10} (out)] + 0.21VR + 0.16WS – 0.17 Temp (In)	1.029–1.839	1.084
	Spring	0.54	PM _{2.5} (in) = 0.05 + 0.72 [PM _{2.5} (out)] + 0.14 [PM _{2.5–10} (out)] – 0.01 [CO ₂ (out)]	1.065–1.313	1.206
PM _{2.5–10}	Fall	0.475	PM _{2.5–10} (in) = 0.09 – 0.49 [PM _{2.5} (out)] + 0.03RH (In) + 1.35 [PM _{2.5–10} (out)]	1.240–1.791	1.271
	Winter	0.595	PM _{2.5–10} (in) = 0.01 + 0.75 [PM _{2.5–10} (out)] + 0.13RH (in) – 0.01Temp (out)	1.111–1.241	1.542
	Spring	0.548	PM _{2.5–10} (in) = –0.09 + 0.49 [PM _{2.5–10} (out)] + 0.34 [PM _{2.5} (out)] + 0.12RH (out)	1.014–1.248	1.157

Table 4
Performance indicators for the best selected validation MLR models.

Pollutant	Seasonal model	NAE	RMSE ($\mu\text{g}/\text{m}^3$)	PA	R ²	IA
PM _{2.5}	Fall	0.33	1.7	0.77	0.58	0.86
	Winter	0.38	1.3	0.85	0.69	0.87
	Spring	0.23	1.4	0.85	0.70	0.88
PM _{2.5–10}	Fall	0.26	2.1	0.66	0.44	0.75
	Winter	0.38	1.4	0.75	0.56	0.83
	Spring	0.29	2.1	0.75	0.57	0.83

explained by the aforementioned independent variables. The RMSE values were 1.4%, 2.1%, and 2.1% for winter, fall and spring, respectively. A comparison among the performances of the three PM_{2.5–10} models indicates that the winter MLR model produced the lowest RMSE value and comparable R² values as spring model. Comparing both MLR models, the MLR model for PM_{2.5} concentration follows the changes in the observed indoor concentration satisfactory and predicts concentration levels with reasonable accuracy better than MLR model of PM_{2.5–10}. PM_{2.5–10} model shows the tendency to under estimated indoor PM_{2.5–10} concentrations as it does not take into account the occupant's activities which highly affect the indoor concentrations during the class hours. This is may be due to the fact the fine particles have higher infiltration rate than coarse particles. On other hand, indoor concentration of the fine particles are least affected by occupant activities and movements. Different studies worldwide reported that the occupant's activities/movements which may cause either resuspension of deposited coarse particle or delayed the deposition process is the major source of indoor PM (Goyal and Khare, 2011; Elbayoumi et al., 2013; Habil et al., 2013).

3.5. Feedforward backpropagation (FFBP)

The optimum trained FFBP structures were selected according to the minimum root mean square error (RMSE), normalized absolute error (NAE), and maximum prediction accuracy (PA), R², and index of agreement (IA) of the test set. The optimal parameters of FFBP neural model for the seasonal indoor PM_{2.5–10} and PM_{2.5} concentrations based on the results of several performance indicators from different and repeated models are summarized in Table 5.

Table 5
Optimal parameters of seasonal FFBP models.

		Input variables	Network structure Input:Neurons:Output	Transfer function for hidden layer/Transfer function for output layer	NAE	RMSE ($\mu\text{g}/\text{m}^3$)	PA	R ²	IA
PM _{2.5}	Fall	PM _{2.5} (out), PM _{2.5–10} (out), CO ₂ (in), RH (in)	04:12:01	Tansig/Purlin	0.18	1.4	0.90	0.75	0.91
	Winter	PM _{2.5} (out), PM _{2.5–10} (out), CO ₂ (in), Temp (out), WS	05:13:01	Tansig/Purlin	0.16	1.4	0.89	0.78	0.92
	Spring	PM _{2.5} (out), PM _{2.5–10} (out), CO ₂ (in), RH (out) Temp (out), VR, WS	07:15:01	Tansig/Purlin	0.16	1.3	0.90	0.79	0.95
PM _{2.5–10}	Fall	PM _{2.5} (out), PM _{2.5–10} (out) VR, RH (in)	04:13:01	Tansig/Purlin	0.18	2.9	0.81	0.65	0.89
	Winter	PM _{2.5–10} (out), CO (in), CO ₂ (out) RH (in), Temp (out), WS	06:14:01	Tansig/Purlin	0.18	1.3	0.87	0.73	0.89
	Spring	PM _{2.5} (out), PM _{2.5–10} (out), CO ₂ (in), RH (in) Temp (out), WS	06:13:01	Tansig/Purlin	0.19	2.9	0.88	0.78	0.90

The performance levels of the validation models are illustrated in Table 5. The R² for the three seasons results shows that independent variables are fitted to PM_{2.5} data and can explain approximately 75%–79% of the variation in the indoor PM_{2.5}. In addition, the values of RMSE ranged from 1.3% for fall and spring models to 1.4% for the winter model. A comparison among the performances of the three PM_{2.5} models indicates that the spring FFBP-NN model produced the lowest RMSE value and highest R² values. The calculated R² for PM_{2.5–10} during the three seasons ranged from 0.65 to 0.78. Thus, approximately 65%–78% of the variation in the indoor PM_{2.5–10} concentrations can be explained by the aforementioned independent variables. The RMSE values were 1.4%, 2.9%, and 2.9% for fall, winter, and spring, respectively. Although spring FFBP-NN produced higher RMSE value comparing with winter model, it performed better than the winter model at both ends of the indoor PM_{2.5–10} range.

3.6. Comparison between models

Performance indicators were used to compare between seasonal PM_{2.5} and PM_{2.5–10} MLR and FFBP models to indicate the models' performance and to test the best model, as shown in Tables 4 and 5. NAE and RMSE were used to find the error of the model, where a value closer to 0 indicated a better model. The other three performance indicators, namely, IA, PA, and R², were used to check the accuracy of the model result, where a higher accuracy is given by a value closer to 1. The R² of FFBP models show strong correlations between predicted and observed concentrations during the three seasons compared with the MLR models. The remaining accuracy measures (PA and IA) for FFBP models are higher than those for MLR models. In addition, the values of the error measures, namely, NAE and RMSE, are smaller for FFBP models compared with those of the MLR models. Furthermore, FFBP improved the accuracy of the models (R²) as calculated by the percentage differences compared with MLR by as much as 25.56%, 12.24%, and 12.08% during fall, winter, and spring for prediction of indoor PM_{2.5}, and by 38.53%, 26.36%, and 31.11% during fall, winter, and spring, respectively for prediction of indoor PM_{2.5–10}. In addition, FFBP improved the accuracy of the models with reduced error (RMSE) as calculated by the percentage differences compared with MLR by as much as

19.35%, 7.41%, and 7.41% during fall, winter, and spring for prediction of indoor $PM_{2.5}$, and by 32.00%, 7.41%, and 32.00% during fall, winter, and spring, respectively for prediction of indoor $PM_{2.5-10}$. Comparison statistics between linear (MLR) and nonlinear models (FFBP) presented by several studies indicate that the NN approach has an edge over linear models, prediction error and prediction accuracy, demonstrating that NN models, if properly trained and formed, can provide adequate solutions to particulate pollution predictive demands (Ul-Saufie et al., 2013; Mishra and Goyal, 2015; Russo et al., 2015). A comparison with other studies (Table S2) shows that the predicted indoor average concentrations of $PM_{2.5-10}$ recorded at the three seasons in this study were far higher compared with results obtained from other studies such as: Braniš and Šafránek (2011) in Prague, Czech Republic; Goyal and Khare (2011) in India and Krasnov et al. (2015) in Israel. The predicted indoor concentrations of $PM_{2.5}$ in this study can be considered comparable compared to the concentrations of indoor $PM_{2.5}$ recorded in study by Goyal and Khare (2011) in India and Tippayawong and Khuntong (2007) in Thailand. Thus, the results obtained from the FFBP models were more encouraging than those of the constructed MLR models.

4. Conclusion

Previously, MLR and FFBP neural network methods were used effectively to study air pollution and meteorological records. This study used the capability of these two techniques to predict seasonal indoor $PM_{2.5}$ and $PM_{2.5-10}$ concentrations in natural ventilated schools located in Mediterranean climate. To improve the efficiency of FFBP neural network prediction, FSR method was used to select the key input variables for the optimal structure of the models. The obtained results of seasonal MLR and FFBP prediction models for indoor $PM_{2.5}$ were in good agreement with the actual measurements, with R^2 values that range between 58% and 70% for MLR and between 75% and 79% for FFBP. The result (R^2) for seasonal indoor $PM_{2.5-10}$ of both MLR and FFBP models ranged between 44% to 57% and 65%–78%, respectively. An assessment of the model's performance verified that the FFBP models provided higher-quality prediction with lower error (NAE, RMSE) and with greater accuracy (IA, PA, and R^2) compared with the MLR models. FFBP neural network could be a new promising methodology instead of MLR for predicting IAQ in naturally ventilated buildings.

Acknowledgment

The authors wish to acknowledge the Universiti Sains Malaysia (USM) and School of Civil Engineering for USM Post-Doctoral Research Fellow program, RUI 814183 grant and RUI 811206 grant.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.apr.2015.09.001>.

References

Adar, S.D., Davey, M., Sullivan, J.R., Compher, M., Szpiro, A., Sally Liu, L.-J., 2008. Predicting airborne particle levels aboard Washington state school buses. *Atmos. Environ.* 42, 7590–7599.

Agirre-Basurko, E., Ibarra-Berastegi, G., Madariaga, I., 2006. Regression and multi-layer perceptron-based models to forecast hourly O_3 and NO_2 levels in the bilbao area. *Environ. Model. Softw.* 21, 430–446.

Al-Alawi, S.M., Abdul-Wahab, S.A., Bakheit, C.S., 2008. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. *Environ. Model. Softw.* 23, 396–403.

Almeida, S.M., Canha, N., Silva, A., Freitas, M.C., Pegas, P., Alves, C., Evtugina, M., Pio, C.A., 2011. Children exposure to atmospheric particles in indoor of Lisbon primary schools. *Atmos. Environ.* 45, 7594–7599.

Alshittawi, M., Awbi, H., Mahyuddin, N., 2009. Particulate matter mass concentration (PM_{10}) under different ventilation methods in classrooms. *Int. J. Vent.* 8, 93–108.

Arkouli, M., Ulke, A., Endlicher, W., Baumbach, G., Schultz, E., Vogt, U., Muller, M., Dawidowski, L., Faggi, A., Wolf-Benning, U., 2010. Distribution and temporal behavior of particulate matter over the urban area of Buenos Aires. *Atmos. Pollut. Res.* 1, 1–8.

Ashmore, M., Dimitroulopoulou, C., 2009. Personal exposure of children to air pollution. *Atmos. Environ.* 43, 128–141.

Blondeau, P., Iordache, V., Poupard, O., Genin, D., Allard, F., 2005. Relationship between outdoor and indoor air quality in eight French schools. *Indoor Air* 15, 2–12.

Branis, M., Rezacová, P., Domasová, M., 2005. The effect of outdoor air and indoor human activity on mass concentrations of PM_{10} , $PM_{2.5}$, and PM_1 in a classroom. *Environ. Res.* 99, 143–149.

Branis, M., Šafránek, J., 2011. Characterization of coarse particulate matter in school gyms. *Environ. Res.* 111, 485–491.

Buonanno, G., Fuoco, F., Morawska, L., Stabile, L., 2013. Airborne particle concentrations at schools measured at different spatial scales. *Atmos. Environ.* 67, 38–45.

Chaloulakou, A., Mavroidis, I., Spyrellis, N., 2001. A comparison of indoor and outdoor carbon monoxide concentration levels at a public school in Athens. Application of a statistical prediction tool. In: 7th International Conference on Environmental Science and Technology, Ermoupolis, Syros Island, Greece.

Chan, A.T., 2002. Indoor-outdoor relationships of particulate matter and nitrogen oxides under different outdoor meteorological conditions. *Atmos. Environ.* 36, 1543–1551.

Chao, C.Y., Wong, K.K., 2002. Residential indoor PM_{10} and $PM_{2.5}$ in Hong Kong and the elemental composition. *Atmos. Environ.* 36, 265–277.

Chen, C., Zhao, B., 2011. Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmos. Environ.* 45, 275–288.

Chithra, V., Nagendra, S., 2012. Indoor air quality investigations in a naturally ventilated school building located close to an urban roadway in Chennai, India. *Build. Environ.* 54, 159–167.

Dayan, U., Heffter, J., Miller, J., Gutman, G., 1991. Dust intrusion events into the Mediterranean basin. *J. Appl. Meteorology* 30, 1185–1199.

Diapouli, E., Chaloulakou, A., Spyrellis, N., 2007. Indoor and outdoor particulate matter concentrations at schools in the Athens area. *Indoor Built Environ.* 16, 55–61.

Elangasinghe, M., Dirks, K., Singhal, N., Costello, S., Longley, I., Salmond, J., 2014. A simple semi-empirical technique for apportioning the impact of roadways on air quality in an urban neighbourhood. *Atmos. Environ.* 83, 99–108.

Elbayoumi, M., Ramli, N., Md Yusof, N., Al Madhoun, W., 2013. Spatial and seasonal variation of particulate matter (PM_{10} and $PM_{2.5}$) in middle eastern classrooms. *Atmos. Environ.* 80, 389–397.

Elbayoumi, M., Ramli, N., Md Yusof, N., Al Madhoun, W., 2014a. Seasonal variation in indoor air schools environments and health symptoms among students in an eastern Mediterranean climate. *Hum. Ecol. Risk Assess.* 20, 1–21.

Elbayoumi, M., Ramli, N., Md Yusof, N., Yahaya, A., Al Madhoun, W., Ul-Saufie, A., 2014b. Multivariate methods for indoor PM_{10} and $PM_{2.5}$ modelling in naturally ventilated schools buildings. *Atmos. Environ.* 94, 11–21.

Field, A., 2009. *Discovering Statistics Using Spss*. Sage Publications, p. 214.

Fromme, H., Twardella, D., Dietrich, S., Heitmann, D., Schierl, R., Liebl, B., Rüden, H., 2007. Particulate matter in the indoor air of classrooms—exploratory results from Munich and surrounding area. *Atmos. Environ.* 41, 854–866.

Gardner, M., Dorling, S., 1999. Neural network modelling and prediction of hourly NO and NO_2 concentrations in urban air in London. *Atmos. Environ.* 33, 709–719.

Gervasi, O., 2008. *Computational Science and its Applications-iccsa 2008*. Springer, Italy.

Gold, D.R., Allen, G., Damokosh, A., Serrano, P., Hayes, C., Castillejos, M., 1996. Comparison of outdoor and classroom ozone exposures for school children in Mexico city. *J. Air Waste Manag. Assoc.* 46, 335–342.

Goyal, R., Khare, M., 2009. Indoor-outdoor concentrations of RSPM in classroom of a naturally ventilated school building near an urban traffic roadway. *Atmos. Environ.* 43, 6026–6038.

Goyal, R., Khare, M., 2011. Indoor air quality modeling for PM_{10} , $PM_{2.5}$, and $PM_{1.0}$ in naturally ventilated classrooms of an urban Indian school building. *Environ. Monit. Assess.* 176, 501–516.

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.

Habil, M., Massey, D.D., Taneja, A., 2013. Exposure of children studying in schools of India to PM levels and metal contamination: sources and their identification. *Ind. Qual. Atmos. Health* 6, 575–587.

HAL, 2012. <http://www.haltechnologies.com/Docs/HAL-HPC300%20User%20Manual.pdf>, (accessed in 2014).

Hassanvand, M.S., Naddafi, K., Faridi, S., Arhami, M., Nabizadeh, R., Sowlat, M.H., Pourpak, Z., Rastkari, N., Momeni, F., Kashani, H., 2014. Indoor/outdoor relationships of PM_{10} , $PM_{2.5}$, and PM_1 mass concentrations and their water-soluble ions in a retirement home and a school dormitory. *Atmos. Environ.* 82, 375–382.

- Heudorf, U., Neitzert, V., Spark, J., 2009. Particulate matter and carbon dioxide in classrooms—the impact of cleaning and ventilation. *Int. J. Hyg. Environ. Health* 212, 45–55.
- Janssen, N., Hoek, G., Brunekreef, B., Harssema, H., 1999. Mass concentration and elemental composition of PM₁₀ in classrooms. *Occup. Environ. Med.* 56, 482–487.
- Jedrychowski, W.A., Perera, F.P., Spengler, J.D., Mroz, E., Stigter, L., Flak, E., Majewska, R., Klimaszewska-Rembiasz, M., Jacek, R., 2013. Intrauterine exposure to fine particulate matter as a risk factor for increased susceptibility to acute broncho-pulmonary infections in early childhood. *Int. J. Hyg. Environ. Health* 216, 395–401.
- Jef, H., Clemens, M., Gerwinb, D., Fransb, F., Olivierc, B., 2005. A neural network forecast for daily average PM₁₀ concentrations in Belgium. *Atmos. Environ.* 39, 3279–3289.
- Kam, W., Cheung, K., Daher, N., Sioutas, C., 2011. Particulate matter (PM) concentrations in underground and ground-level rail systems of the Los Angeles metro. *Atmos. Environ.* 45, 1506–1516.
- Karppinen, A., Kukkonen, J., Elolähde, T., Kontinen, M., Koskentalo, T., Rantakrans, E., 2000. A modelling system for predicting urban air pollution: model description and applications in the Helsinki metropolitan area. *Atmos. Environ.* 34, 3723–3733.
- Koçak, M., Kubilay, N., Tugrul, S., Mihalopoulos, N., 2010. Atmospheric nutrient inputs to the northern levantine basin from a long-term observation: sources and comparison with riverine inputs. *Biogeosciences* 7, 4037–4050.
- Krasnov, H., Katra, I., Novack, V., Vodonos, A., Friger, M.D., 2015. Increased indoor PM concentrations controlled by atmospheric dust events and urban factors. *Build. Environ.* 87, 169–176.
- Kriesel, D., 2007. A Brief Introduction to Neural Networks, first ed.
- Krom, M., Herut, B., Mantoura, R., 2004. Nutrient budget for the eastern mediterranean: Implications for phosphorus limitation. *Limnol. Oceanogr.* 49, 1582–1592.
- Kulshreshtha, P., Khare, M., 2011. Indoor exploratory analysis of gaseous pollutants and respirable particulate matter at residential homes of Delhi, India. *Atmos. Pollut. Res.* 2, 337–350.
- Lal, B., Tripathy, S.S., 2012. Prediction of dust concentration in open cast coal mine using artificial neural network. *Atmos. Pollut. Res.* 2, 211–218.
- Matvev, V., Dayan, U., Tass, I., Peleg, M., 2002. Atmospheric sulfur flux rates to and from Israel. *Sci. Total Environ.* 291, 143–154.
- Milner, J., Dimitroulopoulou, C., ApSimon, H., 2004. Indoor Concentrations in Buildings from Sources Outdoors. UK Atmospheric Dispersion Modeling Liaison Committee. ADMLC/2004/2.
- Mishra, D., Goyal, P., 2015. Development of artificial intelligence based NO₂ forecasting models at Taj Mahal, Agra. *Atmos. Pollut. Res.* 6, 99–106.
- Mishra, D., Goyal, P., Upadhyay, A., 2015. Artificial intelligence based approach to forecast PM_{2.5} during haze episodes: a case study of Delhi, India. *Atmos. Environ.* 102, 239–248.
- Monn, C., Braendli, O., Schaeppi, G., Schindler, C., Ackermann-Liebrich, U., Leuenberger, P., 1995. Particulate matter <10 µm (PM₁₀) and total suspended particulates (TSP) in urban, rural and Alpine air in Switzerland. *Atmos. Environ.* 29, 2565–2573.
- Nejadkoorki, F., Baroutian, S., 2011. Forecasting extreme PM₁₀ concentrations using artificial neural networks. *Int. J. Environ. Res.* 6, 277–284.
- Onat, B., Stakeeva, B., 2013. Personal exposure of commuters in public transport to PM_{2.5} and fine particle counts. *Atmos. Pollut. Res.* 3, 329–335.
- Ozkaynak, H., Xue, J., Spengler, J., Wallace, L., Pellizzari, E., Jenkins, P., 1996. Personal exposure to airborne particles and metals: results from the particle team study in riverside, California. *J. Expo. Analysis Environ. Epidemiol.* 6, 57.
- Pegas, P.N., Evtyugina, M.G., Alves, C.A., Nunes, T., Cerqueira, M., Franchi, M., Pio, C., Almeida, S.M., Freitas, M.C., 2010. Outdoor/indoor air quality in primary schools in Lisbon: a preliminary study. *Quim. Nova* 33, 1145–1149.
- PMD, (The Palestinian Meteorological Department), 2012. <http://www.pmd.ps/ar/mna5palestine.php>, (accessed in 2014).
- Pope, C.A., Young, B., Dockery, D., 2006. Health effects of fine particulate air pollution: Lines that connect. *J. Air & Waste Manag. Assoc.* 56, 709–742.
- Pope III, C.A., Ezzati, M., Dockery, D.W., 2013. Fine particulate air pollution and life expectancies in the United States: the role of influential observations. *J. Air & Waste Manag. Assoc.* 63, 129–132.
- Poupard, O., Blondeau, P., Iordache, V., Allard, F., 2005. Statistical analysis of parameters influencing the relationship between outdoor and indoor air quality in schools. *Atmos. Environ.* 39, 2071–2080.
- Ramírez, N., Cuadras, A., Rovira, E., Borrull, F., Marcé, R.M., 2012. Chronic risk assessment of exposure to volatile organic compounds in the atmosphere near the largest mediterranean industrial site. *Environ. Int.* 39, 200–209.
- Razali, N.Y.Y., Latif, M.T., Dominick, D., Mohamad, N., Sulaiman, F.R., Srithawirat, T., 2015. Concentration of particulate matter, CO and CO₂ in selected schools in Malaysia. *Build. Environ.* 87, 108–116.
- Roberts, S., 2004. Interactions between particulate air pollution and temperature in air pollution mortality time series studies. *Environ. Res.* 96, 328–337.
- Russo, A., Lind, P.G., Raischel, F., Trigo, R., Mendes, M., 2015. Neural network forecast of daily pollution concentration using optimal meteorological data at synoptic and local scales. *Atmos. Pollut. Res.* 6, 540–549.
- Shomar, B., Osenbrück, K., Yahya, A., 2008. Elevated nitrate levels in the ground-water of the Gaza strip: distribution and sources. *Sci. Total Environ.* 398, 164–174.
- Sofuoglu, S.C., Sofuoglu, A., Birgili, S., Tayfur, G., 2006. Forecasting ambient air SO₂ concentrations using artificial neural networks. *Energy Sources, Part B* 1, 127–136.
- Somuri, D.C., 2011. Study of Particulate Number Concentrations in Buses Running with Bio Diesel and Ultra Low Sulfur Diesel. Master Thesis. University of Toledo, Canada.
- Sousa, S., Martins, F., Alvim-Ferraz, M., Pereira, M.C., 2007. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environ. Model. Softw.* 22, 97–103.
- Stafoggia, M., Samoli, E., Alessandrini, E., Cadum, E., Ostro, B., Berti, G., Faustini, A., Jacquemin, B., Linares, C., Pascal, M., 2013. Short-term associations between fine and coarse particulate matter and hospitalizations in southern europe: results from the med-particles project. *Environ. Health Perspect.* 121, 1026–1033.
- Stranger, M., Potgieter-Vermaak, S., Van Grieken, R., 2008. Characterization of indoor air quality in primary schools in Antwerp, Belgium. *Indoor Air* 18, 454–463.
- Tippayawong, N., Khuntong, P., 2007. Model prediction of indoor particle concentrations in a public school classroom. *J. Chin. Inst. Eng.* 30, 1077–1083.
- Ul-Saufie, A., Yahaya, A., Ramli, N., Awang, N., Hamid, H., 2013. Future daily pm₁₀ concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). *Atmos. Environ.* 77, 621–630.
- UNCT (Office of the United Nations Special Coordinator for Middle East Peace Process), 2012. www.unrwa.org/userfiles/file/.../gaza/Gaza%20in%202020.pdf, (accessed in 2015).
- Viotti, P., Liuti, G., Di Genova, P., 2002. Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia. *Ecol. Model.* 148, 27–46.
- Wang, W., Lu, W., Wang, X., Leung, A.Y., 2003. Prediction of maximum daily ozone level using combined neural network and statistical characteristics. *Environ. Int.* 29, 555–562.
- WHO (World Health Organization), 2011. Methods for Monitoring Indoor Air Quality in Schools Bonn (Germany).
- Wilks, D.S., 2011. Statistical Methods in the Atmospheric Sciences. Academic Press.
- Yang, J., Rivard, H., Zmeureanu, R., 2005. On-line building energy prediction using adaptive artificial neural networks. *Energy Build.* 37, 1250–1259.
- Zereini, F., Wiseman, C.L.S., 2010. Urban Airborne Particulate Matter: Origin, Chemistry, Fate and Health Impacts. Springer.
- Zwoździak, A., Sówka, I., Krupińska, B., Zwoździak, J., Nych, A., 2013. Infiltration or indoor sources as determinants of the elemental composition of particulate matter inside a school in Wrocław, Poland? *Build. Environ.* 66, 173–180.