ISRAA UNIVERSITY JOURNAL OF APPLIED SCIENCES ISSN : 2523- 0522
Volume 2 : issue 1 , October , 2018

Prediction of Hourly Indoor Carbon Monoxide Concentrations by Using Multivariate Methods with Sensitivity Analysis Technique

Maher Elbayoumi*[1] Suheir Elbayoumi Harb[2]

Received Feb 2018; accepted May 2018

Abstract: *Precise site prediction of indoor hourly carbon monoxide (CO) concentrations in school buildings is a key issue in air quality research nowadays due to its impact on children's health. In this study multivariate statistical methods, multiple linear regression (MLR) and principle component analysis (PCA), were employed to predict hourly indoor CO concentration in Gaza Strip, Palestine. Measurements were carried in 12 schools from October 2012 to May 2013 (one academic year). The results suggested that the selected models are effective forecasting tools and hence can be applicable for short-term forecasting of indoor CO level. The predicted indoor CO concentration values agree strongly well with the measured data with high coefficients of determination (R^2) 0.869, 0.870 for MLR and PCA-MLR (PCR) respectively. Overall, results showed that PCA model combined with MLR improved MLR model of predicting indoor CO concentration, with reduced errors by as much as 7.14%.*

Keywords: *Natural Ventilation; Indoor Air Quality; Data Driven Models; Air Quality Prediction.*

1-Maher Elbayoumi, Energy and Sustainable Environment Center, School of Engineering, Israa University, Gaza, Palestine. elbayoumimaher@gmail.com
2-Suheir Elbayoumi Harb, Palestine Technical College, Dier Elbalah, Gaza Strip, Palestine,suheirma@gmail.com

1. INTRODUCTION

CO is one of the most characteristic traffic pollutants in urban areas and produce as a primary pollutant during the incomplete combustion of fossil fuels and biomass in fumes produced by portable generators, stoves, and gas ranges (USEPA, 2013). Carbon Monoxide (CO) is colorless and odorless pollutant that arises from both natural and anthropogenic sources. CO exhibits toxicity characteristics due to its higher affinity with hemoglobin and as a result reduces the delivering oxygen to the body's tissues. Sustained exposure to CO has long been associated with effects on increasing adverse cardiovascular outcomes, asthma symptoms, hospital admission rates, and heart rate among children. (Slaughter et al., 2003; Liao et al., 2004; Cakmak et al., 2006; ATSDR, 2012). There are also adverse impacts for exposure to low concentrations of CO for a long period among children on learning ability, manual dexterity, attention level, headaches, dizziness, nausea (feeling sick) and tiredness (HPA, 2009; USEPA, 2013) (Raub and Benignus, 2002; Goniewicz et al., 2009; HPA, 2009).

Several studies have investigated diurnal and seasonal CO concentration in different type's buildings. However; these studies have mainly focused on the monitoring, but not on the prediction of IAQ inside buildings (Chaloulakou et al., 2003; Currie et al., 2009; Elbayoumi et al., 2014a). In addition to that, indoor CO concentrations heavily depend outdoor CO concentrations, indoor sources, infiltration, ventilation, and air mixing between and within rooms and on local conditions, such as weather changes,such as differences in temperature, humidity, pressure, atmospheric stability, and wind speed.Therefore; direct and long-term measurements of CO concentrations are not practical. In the absence of effective and efficient means to directly measure indoor CO from school buildings, development of mathematical prediction models might be a good alternative to provide reasonably accurate estimates.

Multiple linear regression (MLR) is one of the most popular methodology to express response of a dependent variable of several independent variables (predictor). Several studies used MLR to correlate the outdoor CO, PM, ozone, NO_2 and meteorological variables with indoor concentration of such pollutant (Chaloulakou et al., 2001; Adar et al., 2008; Braniš and Šafránek, 2011; Elbayoumi et al., 2014b). In spite of its success, MLR presents problems in identifying the most important contributors when multicollinearity, or high correlation between the independent variables in regression equation are present (Abdul-Wahab et al., 2005). The most common methods for removing such multicollinearity are principal component analysis (PCA) which has been proven to be effective tools to study the relationship between voluminous data such as air pollution and meteorological records (Yeniay and Goktas, 2002; Poupard et al., 2005; Ul-Saufie et al., 2010). PCA is used to reduce the number of predictive variables and transform into new variables that are mutually orthogonal, or uncorrelated, as well as to determine dominant multivariate relationships (Abdul-Wahab et al., 2005). However, one of the main drawback is that MLR and principal component regression (PCR) cannot adequately model the non-linear relationships (Al-Alawi et al., 2008).

In the literature, little attention has been paid to forecasting indoor air quality within buildings. Thus, the overarching goal of this project is to present the results of the application of multivariate regression analysis (MLR and PCR) in predicting indoor CO concentration as the function of meteorological parameters and other pollution concentration from natural ventilated school buildings.

2. METHODOLOGY

2.1 Study Area

The Gaza Strip (365 km^2) is located on the eastern coast of the Mediterranean Sea between longitudes 34° 15' and 35° 40' east, and latitude 29° 30' and 23° 15' north. Climatically, the average daily temperature fluctuates from 24°C in summer to 15 °C in winter. Meanwhile, the daily relative humidity varies between 62.5% in the daytime and 83.4% at night in the summer, and between 51.6% in the daytime and 81.3% at night in winter, and the monthly average wind speed is 3 m/s. (Koçak et al. 2010, PMD 2012). The major source of CO in the Gaza Strip is the exhaust of about 61,000 motor vehicles as for 2012, most of which are more than 15 years old and are out-dated (PCBS 2012). Exhausts contain large quantities of CO, CO_2, $PM_{2.5}$, and hydrocarbons.

In addition, during the frequent power outages, many people and institutions use portable electrical generators. Most of the generators used in the Gaza Strip were placed outside but were very close to the buildings to allow the generators to connect to the central electric panel (Elbayoumi et al. 2014). CO from these sources can build up in enclosed or partially enclosed spaces.

Description of Sampling Locations

The concentrations of pollutants were monitored at 12 schools located in north, middle, and south of the Gaza Strip from October 2012 to May 2013 for one academic year. The sampling schools were purposely selected to reflect the diverse natures of human and vehicular activities. The sampling locations are shown in Figure 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 studying activities.

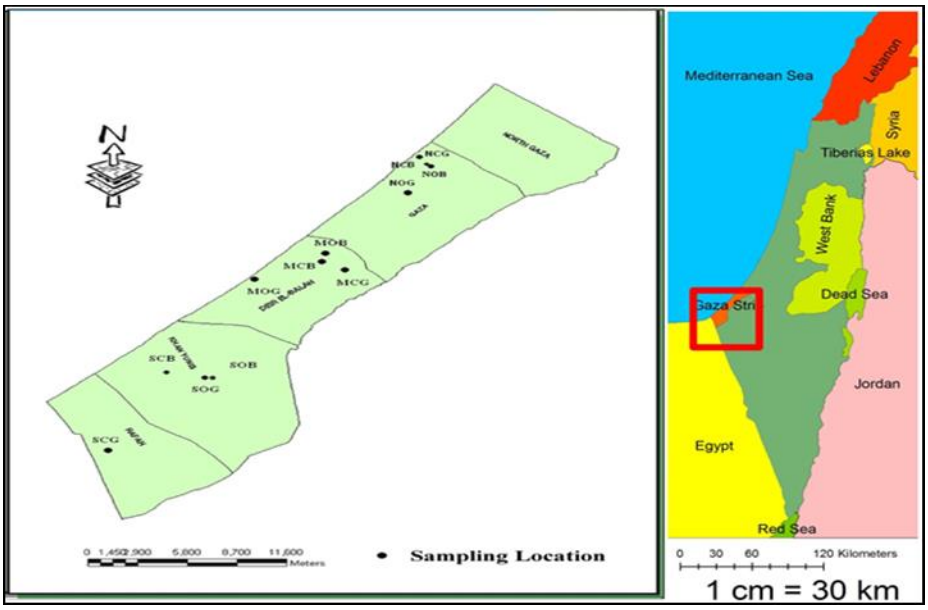

Figure 1: Map of the sampling locations (monitoring schools) in the Gaza Strip (Left) as part of Historical Palestine (Right).

2.2 Measurements and Instrumentation

The measurements were taken place in each site during school hours for three consecutive days. The samplers were placed inside the classrooms opposite the blackboards at least 1 m from the wall and at least 1.5 m height from the floor, as recommended by Blondeau (2005) and WHO (2011). For outdoor sampling, the samplers were placed at the front side of the building, usually near the playground area. A Kanomax IAQ Monitor was used for measuring CO and carbon dioxide (CO_2) concentrations. Meanwhile, the ventilation rate (VR) was calculated using the indoor concentration of CO_2 as a surrogate of the ventilation levels per occupant (Kulshreshtha and Khare 2011, WHO 2011). The mass concentration of particles matters ($PM_{2.5}$ and PM_{10}) has been monitored using handheld optical particle counter (HAL-HPC300). The monitor performs particulate size measurements by using laser light scattering. Air with multiple particle sizes passes through a flat laser beam produced by an ultra-low maintenance laser diode. A 3–channel pulse height analyzer for size classification detects the scattering signals.

Meteorological Data

The surface wind speed (WS), ambient temperature (Temp), relative humidity (RH) and dew point temperature (TDP) in each site were simultaneously measured at the same time with pollutants measurement. A Kanomax IAQ Monitor was used for the temperature and relative humidity measurements. Smart Sensor Electronic Anemometer was used for wind speed.

Data Interpretation

A vital step in the development of a forecast indoor air model is the choice of input parameters (Jef et al. 2005). Sensitivity analysis is a very useful method for ranking the importance of input

ISRAA UNIVERSITY JOURNAL OF APPLIED SCIENCES ISSN : 2523-0522
Volume 2 : issue 1 , October , 2018

variables by assessing their contribution (percentage) to the variability of the model output. In order to choose the most appropriate set of inputs parameters for FFBP, a number of statistical methods can be applied such as stepwise regression (SR), principal component analysis (PCA), and cluster analysis (Wilks 2011). The importance of these methods is to reduce the number of input variables into the models, and, thus, considerably diminish redundant information, instabilities and over-fitting. Here, the selection of variables for the model was made independently for each monitoring school, through a forward stepwise regression (FSR). During this procedure, which starts with the variable most correlated with the target, additional variables are added which, together with the previously selected variables, most accurately predict the target (Wilks 2011). 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% and by using the standardized regression determination coefficient (R^2) values (Wilks 2011). All 15 potential predictors for indoor CO were first considered. The use of the FSR for each monitoring school has reduced the complexity by retaining substantially less variables. The analysis of the data was carried out using the statistical software, SPSS (Statistical Package for Social Science, version 22) and MATLAB, version10. The data had been classified randomly into two sets using MATLAB software. Data set 1, which consists of 70% of the original data, was used for model formulation; and data set 2 (30%) was used for model validation.

Multiple Linear Regression (MLR)

Stepwise multiple regression was carried for CO and the result was checked for multicollinearity by examining the variance inflation factors (VIF) of the predictor variables. Durbin Watson statistic used to check if the model does not have any first order autocorrelation problem. MLR can be expressed according to the following Equation (1):

$$y = b_0 + b_1 x_{1i} + b_2 x_{2i} + \cdots + b_k x_{ki} + \varepsilon \quad (1)$$

where, b_k is the regression coefficients, x_k is the explanatory variables, i =1,2,……k and ε is stochastic error associated with the regression (Agirre-Basurko *et al.*, 2006). The residuals (or error) were checked to evaluate if they were normally distributed with zero mean and constant variance to verify the adequacy of the statistical model (Al-Alawi *et al.*, 2008).

Principle Component Analysis (PCA)

PCA is a multivariate technique that is widely used in dealing with a large amount of data in monitoring studies, such as air pollution studies. In this study, this technique is applied for variables reduction and to provide the most relevant variables in CO variations (Dominick *et al.*, 2012). The PCs were extracted so that the first principal component (PC1) accounted for the largest amount of total variation in the data set, whereas the following components accounted for the remaining variations that were not considered in PC1 (Kovač-Andrić *et al.*, 2009). In general, the PCs are expressed in Equation 2 as follows:

$$PC_i = a_{1i}\, v_1 + a_{2i}\, v_2 + \cdots + a_{ni}\, v_n \qquad (2)$$

where PC_i is the principal component i and a_{ni} is the loading (correlation coefficient) of the original variable V_1 (Özbay *et al.*, 2011).

Each PC represents a linear combination of data (variables) at specific coordinates at different values of chosen parameters (Elbayoumi *et al.*, 2014b). PCs are computed by calculating eigenvalues and eigenvectors. Eigenvalues will determine the eigenvectors and PCs; for each PC, only eigenvalues larger or equal to 1 are considered significant (Ul-Saufie *et al.*, 2013). Rotated PCs using varimax rotation to maximize the relationship between the PCs and original variables (Abdul-Wahab *et al.*, 2005). Dominick *et al.* (2012) reported that varimax rotation ensures that each variable maximally correlated with only one component and has minimal association with other components. The significant variables for each component are determined based on the loading factor where greater than 0.5 is considered strong, 0.4 is moderate, and 0.30 is weak.

Hybrid Models

Hybrid models are models combine MLR technique with PCA. PCR is a combination of MLR and PCA. The use of PCs as input in MLR is intended to reduce the complexity and multicollinearity problems of the models. The selected variables with high loading from PCA ensured that the majority of the original variances were included in the models, and they were ideal for use as independent variables in MLR (Gervasi, 2008; Gvozdić *et al.*, 2011). Figure 2 shows the architecture of a PCR model for prediction of indoor CO concentrations.

ISRAA UNIVERSITY JOURNAL OF APPLIED SCIENCES
ISSN: 2523-0522
Volume 2 : issue 1 , October , 2018

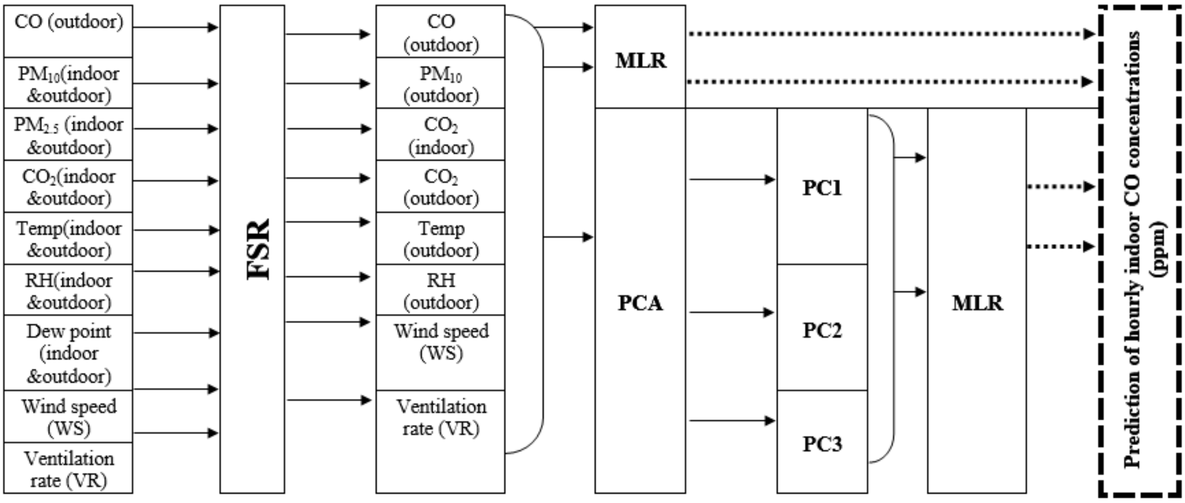

Figure 2. Architecture of a PC-MLR model for the prediction of indoor CO concentrations.

2.3 Performance Indicators

The analysis of prediction performance typically involves calculation of errors between observed y and predicted y values. In this study five performance indicators were used which are normalized absolute error (NAE), root mean square error (RMSE), predication 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 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 (Gervasi 2008, Karppinen et al. 2000).

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

Figure 3 shows the box plot and descriptive statistics of daily indoor and outdoor CO concentrations from 2012 to 2013. In December 2012, February and April 2013, both indoor and outdoor CO concentrations were higher than the remaining months. The outdoor concentrations were 4.80 ppm, 6.07 ppm, and 5.35 ppm for December, February and April, respectively. CO is considered an urban-scale pollutant that is generated by road traffic and tends to be present at high concentrations throughout the city and at significantly reduced concentrations in adjacent rural areas (WHO 2005). The school locations displayed in Figure 1 are very close to street intersections, and most of the schools located in overpopulated areas are characterized by congested traffic. Thus, frequent traffic jams resulting from poorly maintained roads, high traffic density, and very low wind speed are considered the main factors that contribute to high emission, accumulation, and low dilution of generated CO. Moreover, during the December and February months (December-January-Febraury are the coldest months of winter in Palestine), the catalytic converters of vehicles take time to reach the operating temperature when the engine is cold, thereby resulting in increased CO production (Marković et al. 2008). In addition, the CO production in overcrowded residential areas increases when cars move slowly near schools. Thus, the location of the investiged schools may influence indoor CO concentrations. Most of time, the maximum daily concentrations of CO were below WHO’s guidelines of 9 ppm, except for some exceedences that were observed during November, December, February, March and April. The total number of exceedences that were recorded on outdoor concentrations is 82 exceedences, and 22 exceedences on indoor concentrations. The indoor /outdoor (I/O) ratio were less than 1.0 during the monitoring period. Similar results were obtained by Chaloulakou and MaBVRoidis (2002), who revealed that air pollutants, such as CO, that are non-reactive and cannot be absorbed strongly on walls have an I/O ratio close to 1.0 in the absence of indoor sources. Thus, the buildings’ envelope provides little protection from outdoor CO pollution, and peaks in indoor concentrations reached the extremes of outdoor concentrations, regardless of the airtightness in these buildings. Airtightness is the fundamental building’s property that impacts infiltration and exfiltration. In other words, it is the

ISRAA UNIVERSITY JOURNAL OF APPLIED SCIENCES
ISSN: 2523-0522
Volume 2 : issue 1 , October , 2018

uncontrolled inward and outward leakage of outdoor air through cracks, interstices or other unintentional openings of a building, caused by pressure effects of the wind and/or stack effect.

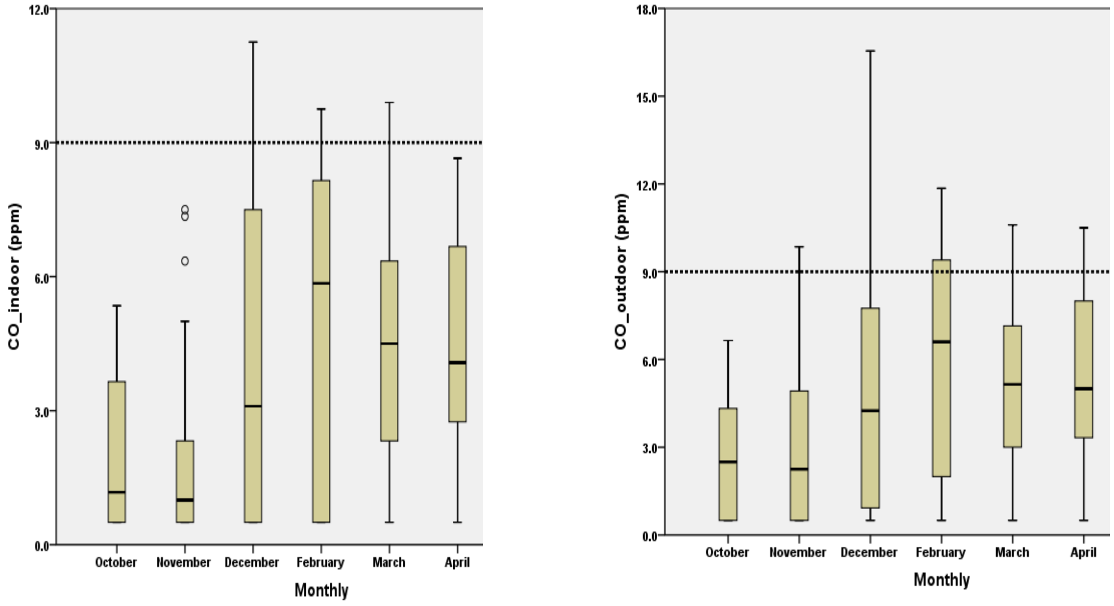

Figure 3: Box plots and descriptive statistics for monthly CO indoor and outdoor for the investigated schools in the Gaza Strip, Palesti

3.2 Bivariate Correlation Analysis

Several studies confirmed that the IAQ is dependent on outdoor concentrations and local conditions, such as weather changes and seasonal variations (Kam et al. 2011). Therefore, bivariate correlation was used to identify the factors that may influence indoor CO concentrations presented in Table 1. A weak relationship exists between indoor and outdoor CO concentrations (r = 0.47), which is generally normal. The value of the correlation coefficient (r) between the indoor and outdoor data can be used as indicator of the degree to which CO measured indoors is attributed to the infiltration from outdoors. Chaloulakou and MaBVRoidis (2002) showed that the indoor peak-concentrations of CO are slightly dampened and lag behind outdoor peaks, thus suggesting that indoor CO concentrations are not immediately affected by outdoor concentration changes, due to changes in air exchange. Moreover, a positive correlation exists between indoor and outdoor CO concentrations and PM_{10}, $PM_{2.5}$ and CO_2 due to the same emission source. The PM–CO correlations observed in this study are consistent with the findings of Dominick et al. (2012). The indoor CO concentration was found to be negatively correlated with indoor temperature and relative humidity. In naturally ventilated buildings, the high building ventilation rate promptly brings indoor humidity to the same level encountered outside. Thus, a negative correlation between humidity and CO infiltration and/or build-up inside the building is expected. Furthermore, a negative correlation exists between indoor and outdoor CO concentration and both ventilation rate and wind speed because low wind speeds favor the accumulation of pollutants (low wind speeds are also related to stable atmospheric conditions).

Table 1 Correlation coefficients (dimensionless) between indoor and outdoor CO concentrations and meteorological Parameters.

Parameters	**CO (Indoor)**	**CO (Outdoor)**
CO (indoor)	1	0.47*
CO (outdoor)	0.47*	1
$PM_{2.5}$(indoor)	0.38*	0.23*
$PM_{2.5}$ (outdoor)	0.34*	0.19*
PM_{10} (indoor)	0.15*	0.13*
PM_{10} (outdoor)	0.29*	0.37*
CO_2 (indoor)	0.26*	0.31*
CO_2 (outdoor)	0.42*	0.40*
Temp (indoor)	-0.41*	-0.39*
Temp (outdoor)	-0.42*	-.34*
RH (indoor)	-0.08*	-0.02
RH (outdoor)	-0.06	-0.01
VR	-0.21*	-0.16*
WS	-0.10*	-0.09*
**Correlation is significant at the 0.01 level (2-tailed).*		

3.3 Principal component analysis

Sensitivity analysis, using FSR technique, was undertaken so as to examine the level of importance for 15 variables i.e. outdoor CO, indoor and outdoor PM_{10}, indoor and outdoor $PM_{2.5}$, indoor and outdoor CO_2, temperature, relative humidity, dew point temperature, ventilation rate ,and wind speed. The standardised regression determination coefficient (R^2) values were used to estimate the relationship between indoor CO and these variables. The results show that 8 variables were identified before the extraction using PCA as shown in Figure 2. After the extraction was applied, three factors were considered as the principal component based on eigenvalues of more than 1. PCA procedures was followed by a varimax rotation to maximize the loading of a predictor variable and the higher loading variable with absolute values greater than 40% were selected for the principal component interpretation (Abdul-Wahab *et al.*, 2005). The factor loading values after rotation are very important in interpretation of PCA results. The factor loadings correlate the factors and the variables, the higher the factor loading, the more variable contributes to the variation of the PC (Ul-Saufie *et al.*, 2013). The eigenvalues for all linear components after rotation are shown in Table 2. The cumulative variance of the principal components is 74.06%. The first PC explains 39. 98% of the total variation in the data set, which indicates a heavy load on relative humidity, temperature dew point, ventilation rate, wind speed and outdoor CO. The second PC, which accounts for approximately 20.96% of the total variation, loads heavily on indoor and outdoor PM_{10} and $PM_{2.5}$. Among the principal components, the third account for approximately 13.12% and load heavily on outdoor and indoor temperature.

3.4 Multiple lin ear regression (MLR)

The stepwise MLR models of indoor CO prediction using the original parameters and PCs as the inputs were conducted with regression assumptions approximately satisfied (Table 3). The four goodness of fit measures showed that the residual distributions were approximately normal, with zero means and no detectable serial and the correlation coefficients of the regressions were all highly statistically significant (P< 0.01) (Abdul-Wahab et al., 2005).

Table 2 Rotated principal component loadings matrix

	Component		
	1	2	3
CO(outdoor)	-0.60		
RH (indoor)	-0.86		
RH(outdoor)	-0.80		
VR	0.83		
WS	0.82		
CO_2(indoor)	0.62		
TDP(indoor)	0.90		
TDP(outdoor)	0.90		
$PM_{2.5}$(indoor)		0.77	
PM_{10}(indoor)		0.77	
PM_{10}(outdoor)		0.75	
$PM_{2.5}$(outdoor)		0.73	
Temp(indoor)			0.93
Temp(outdoor)			0.93
Eigenvalue	6.00	3.14	1.97
% of Variance	39.98	20.96	13.12
Cumulative %	39.98	60.94	74.06

The result showed that the developed models did not encounter multicollinearity problems as the VIF was less than 3.0. In addition, the tolerance values for the variables in MLR model are higher than 0.3. In accordance with the findings of Field et al. (2009), the tolerance value must be smaller than 0.1 to indicate a multicollinearity problem. However, DW may indicate slightly positive auto-correlation problems in the model because these values ranged from 1.966 to 1.961. The developed MLR and PCR models were also assessed using the coefficient of determination (R^2), which was used as an indicator of the ability of the selected variables to explain the variations in indoor CO concentration (Abdul-Wahab et al., 2005).

Table 3 Summary models for indoor CO concentration predictions based on original parameters and PCA as inputs.

Method	Models	R^2	Range of VIF	Durbin-Watson
MLR	CO (indoor) = 0.07 +1.07CO (outdoor) + 0.07CO_2(outdoor) +0.06PM_{10}(outdoor)-0.10Temp(outdoor)	0.72	1.07-2.32	1.96
PCR	CO (indoor) = 0.08+0.88PC1 +0.19PC2-0.03PC3	0.73	1.00	1.96

As presented in Table 3 when the four best variables are fitted to the CO data, the value of the R^2 is approximately (0.72). Thus, approximately 72% of the variation in the indoor CO concentrations can be explained by the four variables, as listed in the table. Meanwhile, the usage of PCs as the inputs in MLR could improve the efficiency of the model to explain the variations in CO concentrations. During these time periods, R^2 values for PCR was approximately (0.73), as reported in Table 4. Therefore, approximately 73% of the variation in the indoor CO concentrations can be explained by the three independent components.

In this study, the applicability of the developed models for predicting the indoor CO concentration variations was assessed using hourly average monitoring records from validation data set. The performance levels of the validation MLR and PCR models indicated that relatively strong relationships were obtained between the observed and predicted values and R^2 for these models was almost 0.87. Thus, R^2 can explain approximately 87% of the variation in the indoor CO by using both models. In addition to that, the values of RMSE ranged from 0.027 % for PCR model to 0.029% for

MLR model. By comparing the performance of the two models, even the R^2 values for the both models were comparable PCR model produced the lowest RMSE comparing with MLR.

3.5 Comparisons

Five performance indicators were used to evaluate and compare between the two models (MLR and PCR) used to predict indoor CO as shown in Table 4. For a good model NAE and RMSE value should approach zero, while PA, R^2 and IA should be closer to one. The results suggest that the four models are effective forecasting tools and hence can be applicable for short-term forecasting of indoor CO level. PCA didn't improve the accuracy measures of MLR model comparing with PCR model where the results of PA, R^2 and IA were almost comparable with MLR but PCA reduced the prediction error in PCR comparing with MLR as calculated by the percentage differences as much as 13.0% and 7.14% for NAE and RMSE, respectively. Thus, it can be recommended that PCR could be a new promising methodology instead of a MLR to predict IAQ in naturally ventilated buildings.

Table 4 Performance indicator for ANN models vs. MLR models for indoor CO

Model	NAE	RMSE	R^2	PA	IA
MLR	0.213	0.029	0.869	0.937	0.965
PCR	0.187	0.027	0.870	0.937	0.961

4. CONCLUSION

Previously, MLR and PCR methods were used effectively to study air pollution and meteorological records. In this study the capability of these techniques to predict indoor CO concentrations in natural ventilated schools located in Gaza Strip, Palestine was employed. To raise the efficiency of MLR, FSR method was used to select the key input variables for the optimal structure of the models. PCA didn't improve the accuracy measures of MLR model comparing with PCR model where the results of PA, R^2 and IA were almost comparable with MLR but PCA reduced the prediction error in PCR comparing with MLR as calculated by the percentage differences as much as 13.0% and 7.14% for NAE and RMSE, respectively. Overall, it was found that the two models i.e. MLR and PCR are effective forecasting tools and hence can be applicable for short-term forecasting of indoor CO level.

ISRAA UNIVERSITY JOURNAL OF APPLIED SCIENCES

ISSN: 2523-0522

Volume 2 : issue 1 , October , 2018

References

2. Abdul-Wahab, S. A., Bakheit, C. S. and Al-Alawi, S. M. (2005). Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environmental Modelling and Software*, 20(10), 1263-1271.

3. Adar, S. D., Davey, M., Sullivan, J. R., Compher, M., Szpiro, A. and Sally Liu, L.-J. (2008). Predicting airborne particle levels aboard Washington State school buses. *Atmospheric Environment*, 42(33), 7590-7599.

4. Agirre-Basurko, E., Ibarra-Berastegi, G. and Madariaga, I. (2006). Regression and multilayer perceptron-based models to forecast hourly O_3 and NO_2 levels in the Bilbao area. *Environmental Modelling and Software,* 21(4), 430-446.

5. Al-Alawi, S. M., Abdul-Wahab, S. A. and Bakheit, C. S. (2008). Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. *Environmental Modelling & Software,* 23(4), 396-403.

6. ATSDR (2012). Toxicological profile for CO. 1st ed.: Agency for Toxic Substances and Disease Registry: U.S. Department of Health and Human Services. Atlanta

7. Blondeau, P., Iordache, V., Poupard, O., Genin, D. and Allard, F. (2005). Relationship between outdoor and indoor air quality in eight French schools. *Indoor Air,* 15(1), 2-12.

8. Braniš, M. and Šafránek, J. (2011). Characterization of coarse particulate matter in school gyms. *Environmental Research,* 111(4), 485-491.

9. Cakmak, S., Dales, R. E. and Judek, S. (2006). Do gender, education, and income modify the effect of air pollution gases on cardiac disease? *Journal of Occupational and Environmental Medicine,* 48(1), 89-94.

10. Chaloulakou, A., Mavroidis, I. and Duci, A. (2003). Indoor and outdoor carbon monoxide concentration relationships at different microenvironments in the Athens area. *Chemosphere,* 52(6), 1007-1019.

11. Chaloulakou, A., MaBVRoidis, I., 2002. "Comparison of indoor and outdoor concentrations of co at a public school. Evaluation of an indoor air quality. Atmospheric Environment, 36 (2002) 1769–1781.

12. Chaloulakou, A., Mavroidis, I., Spyrellis, N. and (2001) In *7th International Conference on Environmental Science and Technology* Ermoupolis, Syros island, Greece.

13. Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M. and Rivkin, S. G. (2009). Does pollution increase school absences? *The Review of Economics and Statistics,* 91(4), 682-694.

14. Dominick, D., Juahir, H., Latif, M. T., Zain, S. M. and Aris, A. Z. (2012). Spatial assessment of air quality patterns in Malaysia using multivariate analysis. *Atmospheric Environment,* 60(1), 172-181.

15. Elbayoumi, M., Ramli, N., Md Yusof, N. and Al Madhoun, W. (2014a). The effect of seasonal variation on indoor and outdoor carbon monoxide concentrations in Eastern Mediterranean climate *Atmospheric Pollution Research* 5(2),315-324.

16. Elbayoumi, M., Ramli, N. A., Md Yusof, N. F. F., Yahaya, A. S. B., Al Madhoun, W. and Ul-Saufie, A. Z. (2014b). Multivariate methods for indoor PM_{10} and $PM_{2.5}$ modelling in naturally ventilated schools buildings. *Atmospheric Environment,* 94, 11-21.

17. Gervasi, O. (2008). Computational Sicience and Its Applications-ICCSA 2008. Springer. Italy

18. Goniewicz, M. Ł., Czogała, J., Kośmider, L., Koszowski, B., Zielińska-Danch, W. and Sobczak, A. (2009). Exposure to carbon monoxide from second-hand tobacco smoke in Polish pubs. *Central European Journal of Public Health,* 17(4), 220.

19. Gvozdić, V., Kovač-Andrić, E. and Brana, J. (2011). Influence of Meteorological Factors NO2, SO2, CO and PM10 on the Concentration of O3 in the Urban Atmosphere of Eastern Croatia. *Environmental modeling & assessment,* 16(5), 491-501.

20. Hal. (2012). HAL-HPC300 user manual. [Online] Available from: http://www.haltechnologies.com/Docs/HAL-HPC300%20User%20Manual.pdf [Accessed 02/10/2012].

21. HPA. (2009). Carbon monoxide:General information. [Online] Available from:www.hpa.org.uk/webc/hpawebfile/hpaweb_c/1202487036268 [Accessed 29/11/2013].

22. Jef, H., Clemensa, M., Gerwinb, D., Fransb, F. and Olivierc, B. (2005). A neural network forecast for daily average PM10 concentrations in Belgium. *Atmospheric Environment*, 39(18), 3279-3289.

23. 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," *Atmospheric Environmen, 45:1506-1516*.

24. Koçak, M., Kubilay, N., Tugrul, S. and Mihalopoulos, N. (2010). Atmospheric nutrient inputs to the northern levantine basin from a long-term observation: sources and comparison with riverine inputs. *Biogeosciences*, 7(12), 4037-4050.

25. Kovač-Andrić, E., Brana, J. and Gvozdić, V. (2009). Impact of meteorological factors on ozone concentrations modelled by time series analysis and multivariate statistical methods. *Ecological Informatics,* 4(2), 117-122.

26. Kulshreshtha, P. and Khare, M. (2011). Indoor exploratory analysis of gaseous pollutants and respirable particulate matter at residential homes of Delhi, India. *Atmospheric Pollution Research,* 2(3), 337-350.

27. Liao, D., Duan, Y., Whitsel, E. A., Zheng, Z.-j., Heiss, G., Chinchilli, V. M. and Lin, H.-M. (2004). Association of higher levels of ambient criteria pollutants with impaired cardiac autonomic control: a population-based study. *American Journal of Epidemiology*, 159(8), 768-777.

28. Marković, D. M., Marković, D. A., Jovanović, A., Lazić, L. and Mijić, Z. (2008). Determination of O_3, NO_2, SO_2, CO and PM_{10} measured in Belgrade urban area. *Environmental Monitoring and Assessment*, 145(1-3), 349-359.

29. Özbay, B., Keskin, G. A., Doğruparmak, Ş. Ç. and Ayberk, S. (2011). Multivariate methods for ground-level ozone modeling. *Atmospheric Research,* 102(1), 57-65.

30. PCBS (2012). Statistical Yearbook of Palestine Report, Palestinian Central Bureau of Statistics (PCBS), Ramalla- Palestine.

31. PMD. (2012). Climate Bulletine. [Online] Available from: http://www.pmd.ps/ar/mna5palestine.php [Accessed 01/10/2012].

32. Poupard, O., Blondeau, P., Iordache, V. and Allard, F. (2005). Statistical analysis of parameters influencing the relationship between outdoor and indoor air quality in schools. *Atmospheric Environment*, 39(11), 2071-2080.

33. Raub, J. A. and Benignus, V. A. (2002). Carbon monoxide and the nervous system. *Neuroscience & Biobehavioral Reviews*, 26(8), 925-940.

34. Slaughter, J. C., Lumley, T., Sheppard, L., Koenig, J. Q. and Shapiro, G. G. (2003). Effects of ambient air pollution on symptom severity and medication use in children with asthma. *Annals of Allergy, Asthma and Immunology*, 91(4), 346-353.

35. Sofuoglu, S. C., Sofuoglu, A., Birgili, S. and Tayfur, G. (2006). Forecasting ambient air SO2 concentrations using artificial neural networks. *Energy Sources, Part B*, 1(2), 127-136.

36. Ul-Saufie, A., Yahya, A. and Ramli, N. (2010). Improving multiple linear regression model using principal component analysis for predicting PM_{10} concentration in Seberang Prai, Pulau Pinang. *International Journal of Environmental Sciences*, 2(2), 403-408.

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

38. USEPA. (2013). An Introduction to Indoor Air Quality (IAQ), Carbon Monoxide (CO). [Online] Available from: http://www.epa.gov/iaq/co.html [Accessed 12/03/2014].

39. WHO (2005). WHO Air Quality Guidelines Global Update 2005, World Health Organization (WHO), Bonn, Germany.

40. WHO (2011). Methods for monitoring indoor air quality in schools World Health Organization (WHO) Bonn. Germany

41. Wilks, D. S. (2011). Statistical methods in the atmospheric sciences. Academic press

42. Yeniay, O. and Goktas, A. (2002). A comparison of partial least squares regression with other prediction methods. *Hacettepe Journal of Mathematics and Statistics*, 31(99), 99-111.