

Research Article

Application of Multiple Linear Regression Models and Artificial Neural Networks on the Surface Ozone Forecast in the Greater Athens Area, Greece

K. P. Moustris,¹ P. T. Nastos,² I. K. Larissi,³ and A. G. Paliatsos⁴

¹Department of Mechanical Engineering, Technological Education Institute of Piraeus, 250 Thivon and P. Ralli, 12244 Athens, Greece

²Laboratory of Climatology and Atmospheric Environment, Faculty of Geology and Geoenvironment, University of Athens, University Campus, Panepistimiopolis GR 157 84, Athens, Greece

³Laboratory of Environmental Technology, Department of Electronic Computer Systems Engineering, Technological Education Institute of Piraeus, 250 Thivon and P. Ralli, 12244 Athens, Greece

⁴General Department of Mathematics, Technological Education Institute of Piraeus, 250 Thivon and P. Ralli, 12244 Athens, Greece

Correspondence should be addressed to P. T. Nastos, nastos@geol.uoa.gr

Received 15 February 2012; Revised 17 May 2012; Accepted 23 May 2012

Academic Editor: Christer Johansson

Copyright © 2012 K. P. Moustris et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An attempt is made to forecast the daily maximum surface ozone concentration for the next 24 hours, within the greater Athens area (GAA). For this purpose, we applied Multiple Linear Regression (MLR) models against a forecasting model based on Artificial Neural Network (ANN) approach. The availability of basic meteorological parameters is of great importance in order to forecast the ozone's concentration levels. Modelling was based on recorded meteorological and air pollution data from thirteen monitoring sites within the GAA (network of the Hellenic Ministry of the Environment, Energy and Climate Change) over five years from 2001 to 2005. The evaluation of the performance of the constructed models, using appropriate statistical indices, shows clearly that in every aspect, the prognostic model by far is the ANN model. This suggests that the ANN model can be used to issue warnings for the general population and mainly sensitive groups.

1. Introduction

Air quality has emerged as a major factor affecting the quality of living in urban areas, especially in densely populated and industrialized areas. Air pollution control is necessary in order to prevent worsening of air quality in the long run. At the same time, short-term forecasting of air quality is required in order to take preventive and clear action during episodes of atmospheric air pollution [1]. The increasing surface ozone concentrations, in recent decades, have become a major concern worldwide due to the adverse effects on human health [2–6].

Long-term human exposure to ozone causes chest pains, persistent coughing, respiratory irritation, impaired lung function, sense of dry throat, worsening of previous respiratory diseases like asthma, severe inflammation of the lungs, and even irreversible damage to the lungs. All these will result in premature aging and chronic human respiratory

diseases. These concern more certain sensitive groups in the population, such as children, asthmatics, and elderly people [7–11], and therefore, it is clear that ozone's forecast is very important for humans.

In the past, there have been many attempts by scientists to forecast ozone concentrations as well as the concentrations of other pollutants in both urban and nonurban areas [12–24]. Specifically, Ziomas et al. [16] presented analytical models relating maximum ozone concentrations in Athens area with various meteorological variables. For this purpose, 54 pollution periods, during the years 1987–1990, were selected and analysed. The evaluation of the developed relationships showed that their degree of successfulness is rather promising. Spellman [18] developed ANN models so as to be able to forecast surface ozone concentration in five different locations in London, during the warm period of the year. In any case, the evaluation of the results by the developed ANN models, indicate that these models surpassed in

comparison with the regression models. Gómez-Sanchis et al. [19] developed a prognostic model in order to predict tropospheric ozone by using ANN models. For this purpose, ambient ozone concentrations were estimated using surface meteorological variables and vehicle emission variables as predictors. Corani [20] applied feed forward neural networks, pruned neural networks, and lazy learning in order to predict at 9 A.M. the concentration estimated for the current day for ozone and PM₁₀ in Milan, Italy. Heo and Kim [21] in their paper described the method of forecasting daily maximum ozone concentrations at four monitoring sites in Seoul, Korea. The forecasting tools developed were fuzzy expert and neural network systems. The hourly data for air pollutants and meteorological variables, obtained both at the surface and at the high-elevation (500 hPa) stations of Seoul City for the period of 1989–1999, were analyzed. Two types of forecast models were developed. The first model, Part I, uses a fuzzy expert system and forecasts the possibility of high ozone levels (equal to or above 80 ppb) occurring on the next day. The second model, Part II, uses a neural network system to forecast the daily maximum concentration of ozone on the following day. The forecasting system includes a correction function so that the existing model can be updated whenever a new ozone episode appears. The accuracy of the forecasting system has been improved continuously through verification and augmentation.

Furthermore, Elkamel et al. [22], in order to forecast the concentration of surface ozone in Kuwait, developed and trained ANN models by using meteorological parameters along with other pollutants concentrations. Finally the developed ANN models were compared with linear and not linear regression models resulting in the supremacy of ANN models in every case. In addition, Slini et al. [25] by using meteorological parameters, such as air temperature and wind speed along with surface ozone concentration of the previous day, proceeded in forecasting the surface ozone concentration the next 24 hours, applying the method of multiple regression analysis. They reached to the conclusion that cooperation of statisticians and environmentalists, leads to rather significant outcomes in matter of air quality forecasting. Finally, Moustris et al. [23] developed ANN models for predicting air quality in many different areas within GAA, one, two, and three days ahead. They used meteorological data, air pollution data, and air quality indices. The results were encouraging and showed that ANN models could give more reliable predictions concerning the quality of ambient environment in the future.

The objective of this study is to perform and compare the MLR models against a forecasting model based on Artificial Neural Network (ANN) approach, in order to forecast the ozone's concentration levels towards the next 24 hours, within the GAA.

2. Data and Methodology

2.1. Area and Data. Multiple linear regression models and ANN models were used in the present study to forecast the maximum daily surface ozone concentration of the

forthcoming 24 hours, within the GAA. The study was based on maximum daily values of surface ozone concentrations ($\mu\text{g}/\text{m}^3$) that were recorded for a five-year period (2001–2005). Hourly surface ozone concentration data and meteorological variables were acquired by the network of the Hellenic Ministry of Environment, Energy and Climate Change (HMEECC). The meteorological variables of air temperature ($^{\circ}\text{C}$) and wind speed (m/s) were obtained by eight different sites stations of the HMEECC monitoring network. Since surface ozone levels can vary considerably within GAA [16], eight suburban and five urban monitoring stations were selected for the analysis. A detailed description of the HMEECC network is found in relevant publications (e.g., [26]).

In the present study, the data set collection consists of 5-year long daily observations of maximum mean hourly ozone concentrations ($\mu\text{g}/\text{m}^3$), from the thirteen stations of HMEECC monitoring network, maximum hourly air temperature ($^{\circ}\text{C}$), and mean wind speed (m/s), from the eight meteorological observation stations of the HMEECC for the same time period. According to scientific literature [16, 25, 27] the choice of the appropriate meteorological data is based on their availability and their impact on shaping the surface ozone concentrations. Air temperature plays an important role, because an increase in air temperature is linked, with rise of the surface ozone concentration. Additionally, another basic meteorological parameter determining the horizontal transport and dispersion of air pollutants is the mean wind speed.

The final constructed data set, from 2001 to 2005, was divided into two subsets. The first concerns the four-year period, from 2001 to 2004, and the second the one-year period for 2005. The data from the first subset were used for the applied models training. The second data subset was not used for the training of the applied models but only for the evaluation of their predictive ability.

2.2. Methodology. The first step followed in the present analysis was the adjustment of daily observations of maximum mean hourly ozone values to the requirements involved by the application of multiple linear regression analysis. Because of the fact that this method requires the data to be normally distributed, the natural logarithm transformation may be applied in ozone values [28], in order to fulfill this prerequisite assumption. The distribution of original (Figure 1(a)) and transformed (Figure 1(b)) ozone values, indicated by bars, compared to the normal distribution. It is obvious that the transformed ozone values, fit to the normal distribution.

A variety of statistical methods have been utilized in order to develop techniques, which will enable qualitative or quantitative short-term forecasts. The most common alternative is to employ a multivariate statistical approach that is widely used in operational ozone forecasting and research oriented statistical modelling (e.g., [14, 25]). Multiple (multivariate) linear regression analysis is the most popular of these techniques, and it has the general form:

$$y = c + \sum_{i=1}^n a_i x_i + \varepsilon, \quad (1)$$

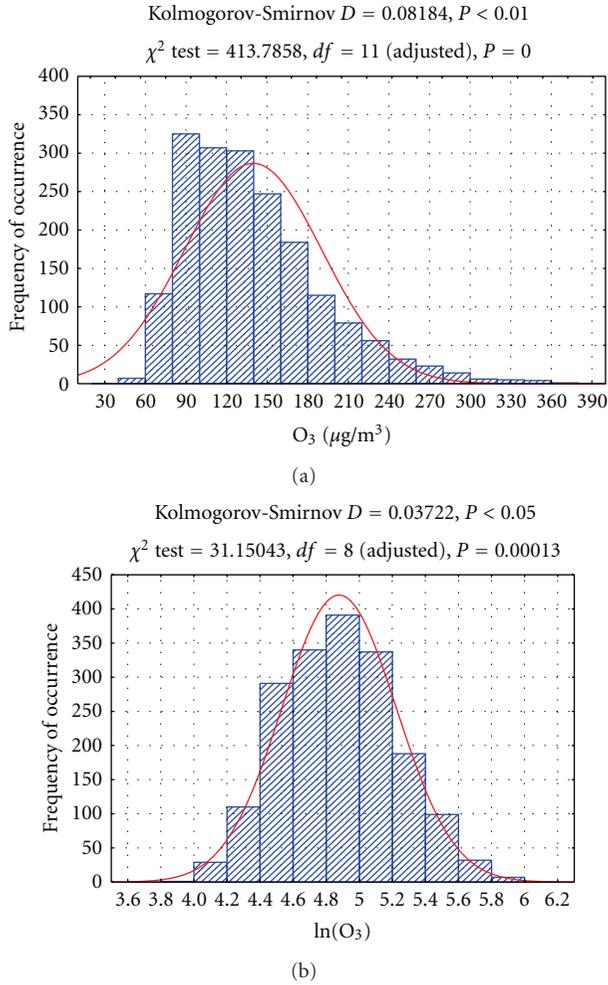


FIGURE 1: Normality test results for ozone (O_3) concentration values (a) and the natural logarithm transformation ($\ln O_3$) of ozone concentrations (b).

where c is the intercept term. a_i , $i = 1, 2, \dots, n$, are the regression coefficients, y is the response variable (surface ozone concentration), x_i , $i = 1, 2, \dots, n$, are the independent predictor variables, and ε is a residual error. When the regression equation is used in predictive mode, e (the difference between actual and predicted values not accounted for by the model) is omitted because its expected value is zero [29].

In multiple regression analysis, an important assumption is that the explanatory variables are independent from each other [30]. However, in some applications of regression, the explanatory variables are related to each other. This problem is called the multicollinearity problem [31]. The non multicollinearity term means that in any case of multiple regression analysis, any correlation between the independent variables should not exist. For this purpose, a multicollinearity index known as variance inflation factor (VIF) is used [30]. The VIF index is calculated by the following equation:

$$\text{VIF} = \frac{1}{1 - R_i^2}, \quad i = 1, 2, 3, \dots, n, \quad (2)$$

where n is the number of predictor variables and R_i^2 is the square of the multiple correlation coefficient of the i th variable with the remaining $(n - 1)$ variables where, according to Hossain et al. [30], if $0 < \text{VIF} < 5$, there is no evidence of multicollinearity problem. If $5 \leq \text{VIF} \leq 10$, there is a moderate multicollinearity problem, and finally if $\text{VIF} > 10$, there is seriously multicollinearity problem of variables.

ANN models are an attempt to imitate and simulate the function of neurons in the human's brain through mathematical functions. In late 1940, Hebb [32, 33] made one of the world's first theories about the mechanism and the ability of human's neurons in the brain, such as their ability to learn. The branch of ANN begins to grow when McCulloch and Pitts [34] create the first ANN model. In 1958, Rosenblatt [35] invented the model of artificial perception-cognition (Perceptron model) which has created more interest in the scientific community for ANN models and their ability to solve at least some simple problems. It is obvious that the simplest ANN model is composed by a single neuron. In this case, of course, the term "network" is unfair because there are no other artificial neurons connected to each other in order to eventually form a network. The only connections are those between inputs $x_1, x_2, x_3, \dots, x_n$ and the artificial neuron. A Perceptron model that is essentially a single artificial neuron is very limited. Such a network, for example, can only represent flat surfaces. This limitation no longer exists when more than one artificial neuron are used. This has led to multilayer perception (MLP). MLP consists of an input layer on which there are the artificial neurons that correspond to the input data (inputs). After the input layer there are one or more hidden layers with one or more artificial neurons. Each artificial neuron on each one of the hidden layers is connected and interchanges information with all the other neurons of both the previous and the next layer. Finally, the output layer, on which there are the "targets" artificial neurons, follows. Since data flow within the artificial neural network from a layer to the next one without any return path, such kind of ANN models are defined as feedforward ANN models. The structure of a feedforward MLP ANN can be found in relevant publications (e.g., [23]).

A weighting factor known as "synaptic weight" corresponds in each connection between neurons. The synaptic weight of an input is a number, which when multiplied with the input gives the weighted input. These weighted inputs are then added together and if they exceed a preset threshold value, the neuron is activated and gives a response-result. When the ANN transfers information from layer to layer the values of these synaptic weights change continuously. After the result reaches the output layer, it is compared with the target value and the error is calculated. In the process, the ANN returns back suitably modifying the values of synaptic weights. When the error is between acceptable limits, the values of synaptic weights are freezing. The ANN model has now been trained properly and is able to replicate and apply its "knowledge" or "experience" on new data and information, which was unknown to it previously. This training ANN process-algorithm is known as the backpropagation error algorithm.

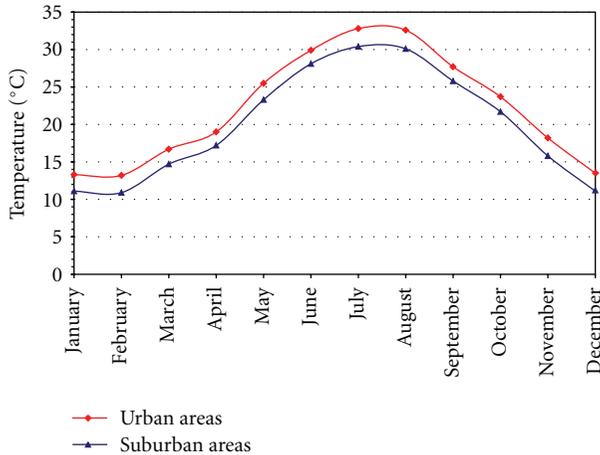


FIGURE 2: Annual variation of mean maximum monthly temperature for the urban and suburban areas within the GAA. Period 2001–2005.

For the evaluation of the results and the predicting performance of the developed model, appropriate statistical indices such as the coefficient of determination (R^2), the mean bias error (MBE), the root mean square error (RMSE) and the index of agreement (IA) were used [29, 36–40]. The RMSE is a commonly used measure of the differences between the predicted values by a predictable model and the real-observed values. The RMSE was used as a single measure that indicates the ability of the model to predict and has the same units as the predicted value. The RMSE is always positive and a zero value is ideal. The MBE provides information on the long-term performance. A low MBE is desirable. Ideally a zero value of MBE should be obtained. A positive value gives the average amount of overestimation in the calculated and negative underestimate. The coefficient of determination is used in cases of statistical models, whose main purpose is the forecast of future outcomes on the basis of other related information. It is the proportion of the variability in a dataset that is accounted for by the statistical model. It provides a measure of how well future outcomes are likely to be predicted by the model. The coefficient values range from zero to one ($0 \leq R^2 \leq 1$). The closer the value is to one, the better and more accurate is the prediction. The index of agreement is a dimensionless measure with values between zero and one ($0 \leq IA \leq 1$). When $IA = 0$, there is no agreement between prediction and observation, while $IA = 1$ denotes a perfect agreement between prediction and observation [37].

3. Results and Discussion

Figure 2 depicts the annual variation of the mean maximum monthly temperature for the urban and suburban examined areas, respectively.

According to Figure 2, there is a permanent difference of about 2°C , between urban and suburban areas. The result is expected and is consistent with the urban heat island (UHI) effect, resulting in higher temperatures within the

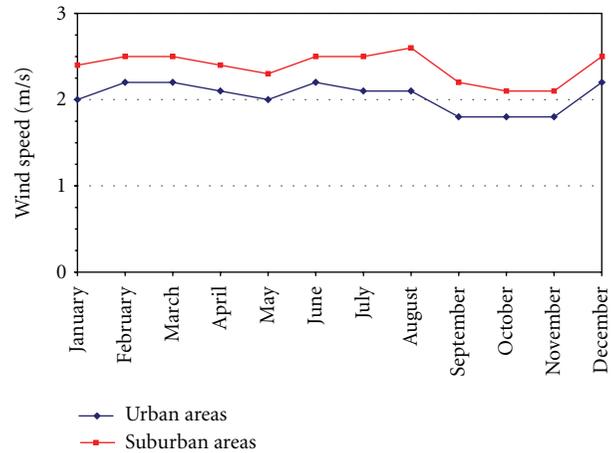


FIGURE 3: Annual variation of mean monthly wind speed for the urban and suburban areas within the GAA. Period 2001–2005.

city's center than the periphery. The mean maximum annual air temperature is about 22.1°C concerning the examined urban areas and 20.1°C for the suburban areas, respectively. During the warm period of the year (May to September), the mean maximum air temperature is about 29.6°C concerning the examined urban areas and 27.5°C for the suburban areas, respectively. Finally, during the cold period of the year (October to April), the mean maximum air temperature is about 16.8°C concerning the examined urban areas and 14.7°C for the suburban areas, respectively.

In a similar way to the above analysis, a brief analysis for the wind speed was done. Figure 3 presents the annual variation of the mean monthly wind speed for the urban and suburban examined areas, respectively.

It is shown that the wind speed for both urban and suburban areas within the GAA is low enough. There is a relatively difference for the wind speed of about 0.3 m/s , between urban and suburban areas. The mean annual wind speed is about 2.0 m/s concerning the examined urban areas and 2.4 m/s for the suburban areas, respectively. During the warm period of the year (May to September) the mean wind speed is about 2.1 m/s concerning the examined urban areas and 2.4 m/s for the suburban areas, respectively. Finally, during the cold period of the year (October to April) the mean wind speed is about 2.0 m/s concerning the examined urban areas and 2.3 m/s for the suburban areas, respectively.

The characteristics and the differences of daily maximum surface ozone concentration in suburban and urban areas were classified and analyzed separately during the examined period 2001 to 2005. Figure 4 shows the annual variation of the mean maximum monthly ozone concentrations. There is a permanent difference of about $15\text{ }\mu\text{g}/\text{m}^3$ between urban and suburban areas. The mean maximum annual ozone concentration is about $89\text{ }\mu\text{g}/\text{m}^3$ for the urban areas and $105\text{ }\mu\text{g}/\text{m}^3$ for the suburban areas, respectively. During the warm period of the year the ozone concentration is about $115\text{ }\mu\text{g}/\text{m}^3$ for the urban areas and $132\text{ }\mu\text{g}/\text{m}^3$ for the suburban areas, respectively. Finally, during the cold period of the year the ozone concentration is about $70\text{ }\mu\text{g}/\text{m}^3$

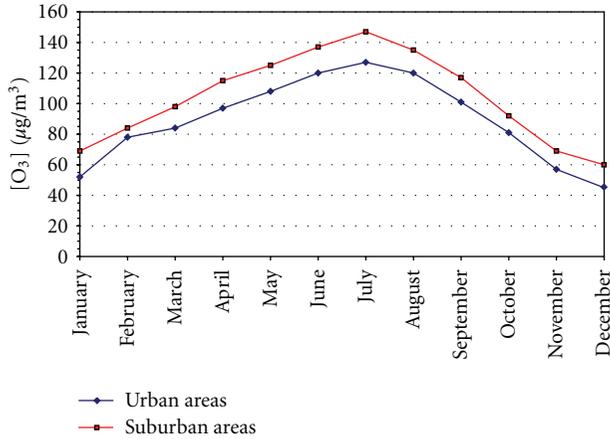


FIGURE 4: Annual variation of mean maximum monthly ozone concentrations for the urban and suburban areas within the GAA. Period 2001–2005.

for the urban areas and $84 \mu\text{g}/\text{m}^3$ for the suburban areas, respectively.

According to the HMEECC, the ozone concentration threshold (maximum hourly values during the day) for the public awareness is $180 \mu\text{g}/\text{m}^3$ and the alert threshold is $240 \mu\text{g}/\text{m}^3$, respectively. For both the urban and suburban examined areas within the GAA, the percentage of days exceeding the above thresholds were analyzed. In an annual base, the results showed that at least in one of the examined urban areas 12.3% of the days exceed the threshold for the public awareness and 95.5% of these days appear during the warm period of the year and only 4.5% during the cold period. During the warm period exceedances days, the mean maximum daily temperature is about 31.9°C and the mean daily wind speed is about 1.7 m/s. Taking into account the suburban areas, 15.8% of the days within the whole year exceed the threshold for the public awareness. From these days, a percentage of 85.5% appears during the warm period of the year and the rest 14.5% during the cold period. Within the warm period exceedances days in suburban areas, the mean maximum daily temperature is 30.0°C and the mean daily wind speed is about 1.9 m/s against 21.5°C and 1.9 m/s, respectively, for the cold period exceedances days.

Concerning the alert threshold in an annual base, 1.6% of the days exceed the threshold at least in one of the examined urban areas. All of these exceedances days (100%) are presented during the warm period of the year. Furthermore, during these days the mean maximum daily temperature is about 33.6°C and the mean daily wind speed is about 1.6 m/s. For the suburban areas, it is found that 4.1% of the days during the whole year exceed the alert threshold. From these days, 86.7% appear during the warm period of the year and the rest 13.3% within the cold period. During the warm period alert exceedances days in suburban areas, the mean maximum daily temperature is 31.3°C and the mean daily wind speed is about 1.7 m/s against 23.7°C 1.5 m/s, respectively, for the cold period. In general, the observed high air temperatures and low wind speeds promote air pollution and especially high ozone concentrations.

For the maximum hourly surface ozone concentration forecast of the next 24 hours within the GAA, prognostic models were developed using multiple linear regression and ANN models. Thus in this work, daily values of the maximum hourly air temperature ($^\circ\text{C}$) and the mean daily wind speed (m/s) within the GAA as well as the daily values of maximum hourly ozone concentrations ($\mu\text{g}/\text{m}^3$) of the previous day were used.

Essentially, three independent variables have been chosen (Table 1) while the dependent variable is the concentration of surface ozone 24 hours ahead (Table 1). Concretely, the independent variable x_1 is the natural logarithm of the maximum observed daily ozone concentration of the previous day, from the thirteen examined areas. The independent variable x_2 is the maximum observed daily air temperature of the previous day, from the eight examined meteorological stations. The independent variable x_3 is the mean daily observed wind speed of the previous day, from the eight examined meteorological stations. Finally, the dependent variable y is the natural logarithm of the maximum daily surface ozone concentration for the next 24 hours, and concerns the maximum ozone concentration from the thirteen examined regions within the GAA one day ahead.

The transformation of ozone concentrations using their natural logarithm was applied, in order to satisfy the linear multiple regression analysis requirement, namely, the values of the independent variable follow the normal distribution. Thus, the dependent variable is now on the natural logarithm of surface ozone concentration 24 hours ahead. The “normality” of both the values of ozone concentrations and the values of the natural logarithm of ozone concentration was checked by the application of the Kolmogorov-Smirnov test (statistic D) and the results are shown in Figure 1. Based on the values of Kolmogorov-Smirnov statistical index D, it is obvious that in the case of the original ozone concentrations ($D = 0.08184$) the residuals do not follow the normal distribution (Figure 1(a)). In the case of the natural logarithm of ozone concentrations ($D = 0.03722$) the residuals seem to follow the normal distribution (Figure 1(b)).

In general terms, we can say that in any case there are three independent variables and one dependent variable. Table 1 presents the variables of all applied forecasting models, which are the same for each developed model as well as their abbreviations.

The nonmulticollinearity between the three independent variables was investigated by the application of the multicollinearity index VIF. Results are presented in Table 2. According to Table 2, the term of nonmulticollinearity between the three independent variables x_1 , x_2 and x_3 is fully fulfilled.

The estimated multiple linear regression (MLR) model was

$$y = 1.4271 + 0.6562x_1 + 0.0101x_2 + 0.0076x_3. \quad (3)$$

As far as the ANN forecasting model is concerned, a multilayer perception neural network with the architecture structure of a time lagged recurrent network (TLRN) was properly developed and trained. Most of the real-world data

TABLE 1: Variables of the developed models.

Independent variables	Description
x_1	The natural logarithm of the maximum daily ozone concentration of the previous day
x_2	The maximum daily air temperature of the previous day
x_3	The mean daily wind speed of the previous day
Dependent variable	Description
y	The natural logarithm of the maximum daily ozone concentration 24 hours ahead

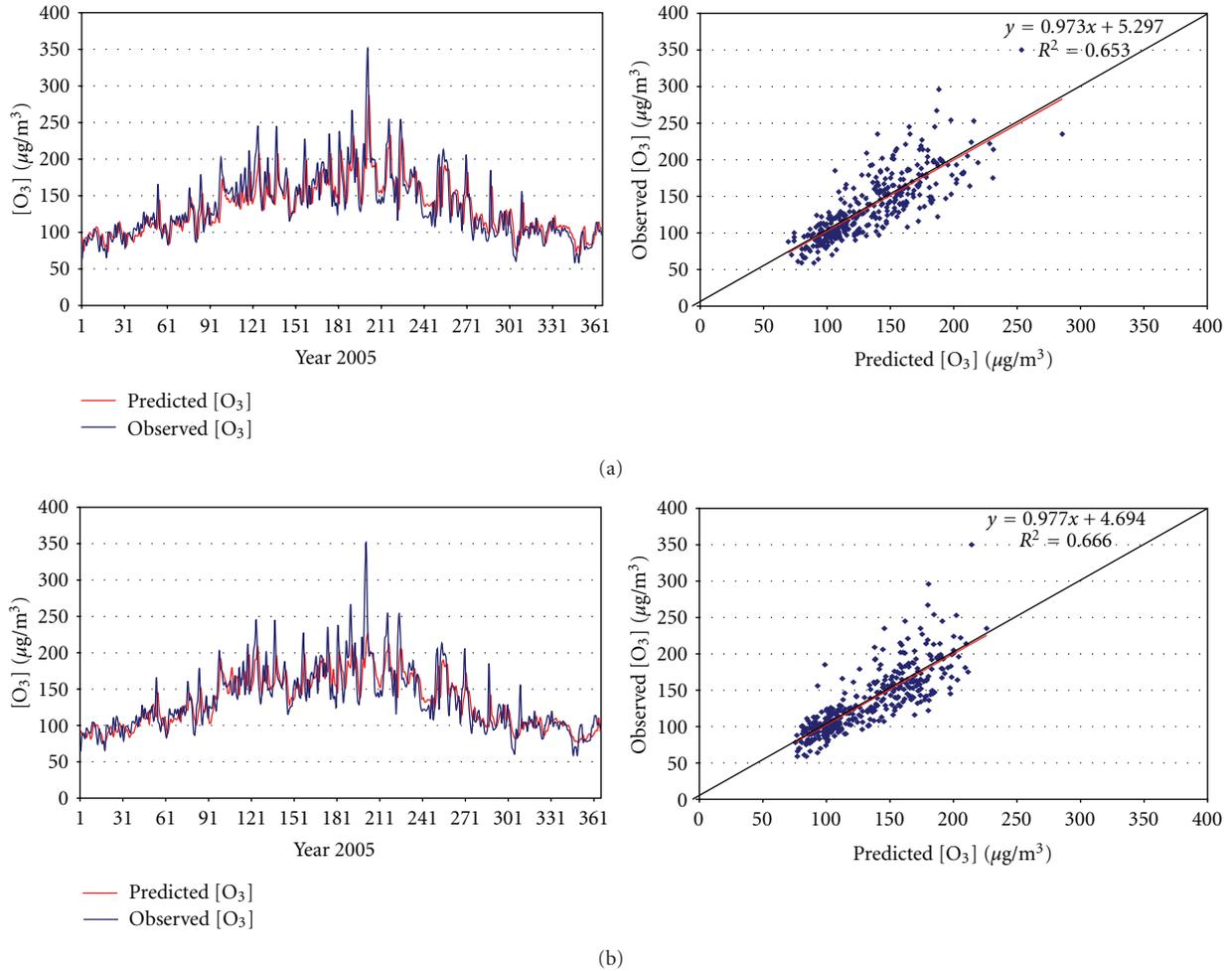


FIGURE 5: Observed (blue line) and predicted (red line) ozone concentration values 24 hours ahead for MLR (a) and ANN (b) model, year 2005.

TABLE 2: Nonmulticollinearity test results between the three independent variables.

Combination of the independent variables	Multiple correlation coefficient	VIF
$x_1 = f(x_2, x_3)$	0.516	2.1
$x_2 = f(x_1, x_3)$	0.509	2.0
$x_3 = f(x_1, x_2)$	0.028	1.0

contain information according the structure of time, that is, how the data changes with time. However, most neural

TABLE 3: Performance statistics for the validation of the developed models.

Forecasting model	R^2	MBE ($\mu\text{g}/\text{m}^3$)	RMSE ($\mu\text{g}/\text{m}^3$)	IA
MLR	0.653	-1.6	25.5	0.887
ANN	0.666	-1.5	25.1	0.892

networks are purely static classifiers. The TLRN neural architecture structure is the evolution of technology in linear time series prediction, system identification, and temporal pattern classification [41, 42].

The developed ANN model consists of one hidden layer with four artificial neurons. The choice of both the type and the architecture of the structure of the developed ANN prognostic model was done by applying the trial-and-error method. In order to train the ANN model, the independent variables x_1 , x_2 , and x_3 (Table 1) were used as input data. The independent variables x_1 , x_2 , and x_3 were chosen as input training data in order that the results of the developed ANN model are comparable with those of the MLR developed model (3). The ANN output is the dependent variable y , the natural logarithm of the maximum daily surface ozone concentration for the next 24 hours, within the GAA. In other words, the input training data are the independent variables x_1 , x_2 , and x_3 of the developed MLR forecasting model, as well as the output of the ANN model is the dependent variable y of the developed MLR forecasting model.

As mentioned before, the used data set was divided into two subsets. The first was included data from the period 2001–2004 and used for the ANN training, and the second subset was the year 2005, which was absolutely unknown to the model and used for the evaluation of the forecasting ability of the developed ANN. The same procedure was followed in the case of all developed MLR models.

Table 3 presents the values of the evaluation statistical indices for both the MLR and the ANN model. At this point, it should be noted that the values of these evaluation statistical indices concern the correlation between the real-observed and the predicted ozone concentrations for the validation year 2005 and they have been emerged after taking back the predicted values of the natural logarithms of surface ozone concentration values of each model.

Figure 5 shows the observed values of maximum daily ozone concentration against the corresponding predicted models values, within GAA for the year 2005. According to Table 3 and Figure 5, it seems that the prediction of all the developed models is in a very satisfactory level ($P < 0.01$). In conclusion, the daily maximum surface ozone concentration forecasting 24 hours ahead is most reliable by using the ANN approach compared to the forecasting by using MLR models. Very close to the ANN prediction is the prediction of the polynomial regression model.

4. Conclusions

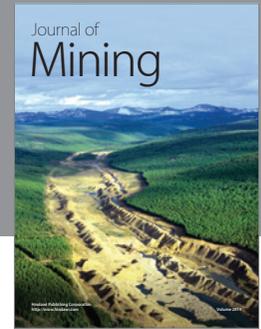
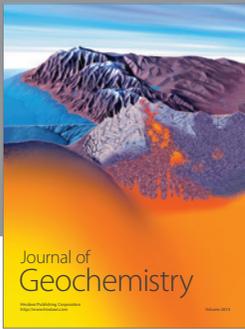
The objective of this study was to examine the one-day forecast of the daily maximum surface ozone concentration within GAA, by developing predictive models based on the method of multiple regression analysis against artificial neural network approach. The performed analysis has indicated that the coefficient of determination (MLR: 0.653, ANN: 0.666) and the index of agreement (MLR: 0.887, ANN: 0.892), between the observed and the predicted ozone concentrations for the year 2005, with respect to MLR and ANN models, are statistically significant ($P < 0.01$). Finally, it should be mentioned that ANN models forecasting ability does present a limited precedence against MLR models. Furthermore, the improvement of the predictive ability of the constructed ANN model could be reached

by the use of several others parameters such as nitrogen oxides concentration, the intensity of solar radiation, and the sunlight duration, based on the disposability of those data over space and time.

References

- [1] W. Z. Lu, H. Y. Fan, A. Y. T. Leung, and J. C. K. Wong, "Analysis of pollutant levels in central Hong Kong applying neural network method with particle swarm optimization," *Environmental Monitoring and Assessment*, vol. 79, no. 3, pp. 217–230, 2002.
- [2] M. Lippmann, "Health effects of tropospheric ozone," *Environmental Science and Technology*, vol. 25, no. 12, pp. 1954–1962, 1991.
- [3] M. Brauer and J. R. Brook, "Ozone personal exposures and health effects for selected groups residing in the Fraser Valley," *Atmospheric Environment*, vol. 31, no. 14, pp. 2113–2121, 1997.
- [4] X. Wang, W. Lu, W. Wang, and A. Y. T. Leung, "A study of ozone variation trend within area of affecting human health in Hong Kong," *Chemosphere*, vol. 52, no. 9, pp. 1405–1410, 2003.
- [5] W. Lu, X. Wang, W. Wang, A. Y. T. Leung, and K. Yuen, "A preliminary study of ozone trend and its impact on environment in Hong Kong," *Environment International*, vol. 28, no. 6, pp. 503–512, 2002.
- [6] Y. C. Lee, G. Calori, P. Hills, and G. R. Carmichael, "Ozone episodes in urban Hong Kong 1994–1999," *Atmospheric Environment*, vol. 36, no. 12, pp. 1957–1968, 2002.
- [7] J. Schwartz, "Air pollution and hospital admissions for respiratory disease," *Epidemiology*, vol. 7, no. 1, pp. 20–28, 1996.
- [8] A. G. Paliatsos, K. N. Priftis, I. C. Ziomas et al., "Association between ambient air pollution and childhood asthma in Athens, Greece," *Fresenius Environmental Bulletin*, vol. 15, no. 7, pp. 614–618, 2006.
- [9] P. T. Nastos, A. G. Paliatsos, M. B. Anthracopoulos, E. S. Roma, and K. N. Priftis, "Outdoor particulate matter and childhood asthma admissions in Athens, Greece: a time-series study," *Environmental Health: A Global Access Science Source*, vol. 9, no. 1, article no. 45, 2010.
- [10] E. G. Kalantzi, D. Makris, M. -N. Duquenne, S. Kاكلamani, E. Stapountzis, and K. I. Gourgoulis, "Air pollutants and morbidity of cardiopulmonary diseases in a semi-urban Greek peninsula," *Atmospheric Environment*, vol. 45, pp. 7121–7126, 2011.
- [11] E. Samoli, P. T. Nastos, A. G. Paliatsos, K. Katsouyanni, and K. N. Priftis, "Acute effects of air pollution on pediatric asthma exacerbation: evidence of association and effect modification," *Environmental Research*, vol. 111, no. 3, pp. 418–424, 2011.
- [12] S. M. Robeson and D. G. Steyn, "Evaluation and comparison of statistical forecast models for daily maximum ozone concentrations," *Atmospheric Environment. Part B*, vol. 24, no. 2, pp. 303–312, 1990.
- [13] P. Zannetti, *Air pollution modeling: Theories, Computational Methods and Available Software*, Van Nostrand Reinhold, New York, NY, USA, 1990.
- [14] A. E. Milionis and T. D. Davies, "Regression and stochastic models for air pollution-I. Review, comments and suggestions," *Atmospheric Environment*, vol. 28, no. 17, pp. 2801–2810, 1994.

- [15] R. A. Olea and V. Pawlowsky-Glahn, "Kolmogorov-Smirnov test for spatially correlated data," *Stochastic Environmental Research and Risk Assessment*, vol. 23, no. 6, pp. 749–757, 2009.
- [16] I. C. Ziomas, D. Melas, C. S. Zerefos, and A. F. Bais, "On the relationship between peak ozone levels and meteorological variables," *Fresenius Environmental Bulletin*, vol. 4, no. 1, pp. 53–58, 1995.
- [17] J. Yi and V. R. Prybutok, "A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area," *Environmental Pollution*, vol. 92, no. 3, pp. 349–357, 1996.
- [18] G. Spellman, "An application of artificial neural networks to the prediction of surface ozone concentrations in the United Kingdom," *Applied Geography*, vol. 19, no. 2, pp. 123–136, 1999.
- [19] J. Gómez-Sanchis, J. D. Martín-Guerrero, E. Soria-Olivas, J. Vila-Francés, J. L. Carrasco, and S. del Valle-Tascón, "Neural networks for analysing the relevance of input variables in the prediction of tropospheric ozone concentration," *Atmospheric Environment*, vol. 40, no. 32, pp. 6173–6180, 2006.
- [20] G. Corani, "Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning," *Ecological Modelling*, vol. 185, no. 2–4, pp. 513–529, 2005.
- [21] J. S. Heo and D. S. Kim, "A new method of ozone forecasting using fuzzy expert and neural network systems," *Science of the Total Environment*, vol. 325, no. 1–3, pp. 221–237, 2004.
- [22] A. Elkamel, S. Abdul-Wahab, W. Bouhamra, and E. Alper, "Measurement and prediction of ozone levels around a heavily industrialized area: a neural network approach," *Advances in Environmental Research*, vol. 5, no. 1, pp. 47–59, 2001.
- [23] K. P. Moustris, I. C. Ziomas, and A. G. Paliatsos, "3-day-ahead forecasting of regional pollution index for the pollutants NO₂, CO, SO₂, and O₃ using artificial neural networks in Athens, Greece," *Water, Air, and Soil Pollution*, vol. 209, no. 1–4, pp. 29–43, 2010.
- [24] Y. Feng, W. Zhang, D. Sun, and L. Zhang, "Ozone concentration forecast method based on genetic algorithm optimized back propagation neural networks and support vector machine data classification," *Atmospheric Environment*, vol. 45, no. 11, pp. 1979–1985, 2011.
- [25] Th. Slini, K. Karatzas, and A. Papadopoulos, "Regression analysis and urban air quality forecasting: an application for the city of Athens," *Global Nest*, vol. 4, pp. 153–162, 2002.
- [26] A. G. Paliatsos, J. K. Kaldellis, P. S. Koronakis, and J. E. Garofalakis, "Fifteen year air quality trends associated with the vehicle traffic in Athens, Greece," *Fresenius Environmental Bulletin*, vol. 11, no. 12 B, pp. 1119–1126, 2002.
- [27] M. W. Gardner and S. R. Dorling, "Meteorologically adjusted trends in UK daily maximum surface ozone concentrations," *Atmospheric Environment*, vol. 34, no. 2, pp. 171–176, 2000.
- [28] EPA, *A Guideline for Developing an Ozone Forecasting Programme*, EPA-454/R-99-009, Research Triangle Park, Office of Air Quality Planning and Standards, US Environmental Protection Agency, NC, USA, 1999.
- [29] A. C. Comrie, "Comparing neural networks and regression models for ozone forecasting," *Journal of the Air and Waste Management Association*, vol. 47, no. 6, pp. 653–663, 1997.
- [30] M. G. Hossain, M. Sabiruzzaman, S. Islam, F. Ohtsuki, and P. E. Lestrel, "Effect of craniofacial measures on the cephalic index of Japanese adult female students," *Anthropological Science*, vol. 118, no. 2, pp. 117–121, 2010.
- [31] S. Chatterjee and A. S. Hadi, *Regression Analysis by Example*, Wiley & Sons, New York, 2006.
- [32] D. O. Hebb, "The effects of early and late brain injury upon test scores and the nature of normal adult intelligence," *Proceedings of the American Philosophical Society*, vol. 85, pp. 275–292, 1942.
- [33] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Theory*, Wiley & Sons, New York, NY, USA, 1949.
- [34] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.
- [35] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958.
- [36] C. J. Willmott, "Some comments on the evaluation of model performance," *Bulletin American Meteorological Society*, vol. 63, no. 11, pp. 1309–1313, 1982.
- [37] C. J. Willmott, S. G. Ackleson, R. E. Davis et al., "Statistics for the evaluation and comparison of models," *Journal of Geophysical Research*, vol. 90, pp. 8995–9005, 1985.
- [38] S. E. Walker, L. H. Slørdal, C. Guerreiro, F. Gram, and K. E. Grønskei, "Air pollution exposure monitoring and estimation Part II. Model evaluation and population exposure," *Journal of Environmental Monitoring*, vol. 1, no. 4, pp. 321–326, 1999.
- [39] M. Kolehmainen, H. Martikainen, and J. Ruuskanen, "Neural networks and periodic components used in air quality forecasting," *Atmospheric Environment*, vol. 35, no. 5, pp. 815–825, 2001.
- [40] A. L. Dutot, J. Rynkiewicz, F. E. Steiner, and J. Rude, "A 24-h forecast of ozone peaks and exceedance levels using neural classifiers and weather predictions," *Environmental Modelling and Software*, vol. 22, no. 9, pp. 1261–1269, 2007.
- [41] S. M. Kelo and S. V. Dudul, "Short-term Maharashtra state electrical power load prediction with special emphasis on seasonal changes using a novel focused time lagged recurrent neural network based on time delay neural network model," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1554–1564, 2011.
- [42] Y. Liang and A. Kelemen, "Time course gene expression classification with time lagged recurrent neural network," *Computational Intelligence in Bioinformatics*, vol. 94, pp. 149–163, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

