Neural Modelling of the Tropospheric Ozone Concentrations in an Urban Site

A. K. Paschalidou[1*], L. S. Iliadis [2], P. Kassomenos[1], C. Bezirtzoglou[2]
[1] *Department of Physics, Laboratory of Meteorology, University of Ioannina, 45110, Ioannina, Greece*
[2] *Democritus University of Thrace, 193 Padazidou st., 68200, Nea Orestiada, Greece*
Email: me00760@cc.uoi.gr*

Abstract

The objective of the present study is to design and develop an Artificial Neural Network (ANN) model for estimations of the ambient ozone concentrations based on meteorological and pollutant parameters. The study focuses on an urban site in the metropolitan area of Athens. The research proves that the optimal ANN is a Modular one that uses the Back Propagation Optimization Algorithm. This ANN includes a Gating Network and it has a single Hidden Layer. Two other Back Propagation ANNs with a simpler architecture reveal a good performance as well. The large amount of data records combined with the good testing results prove the generalization ability of the developed ANN. Statistical analysis techniques, such as combinations of Principal Component and Stepwise Regression Analysis, have been used for the same area in a previous study. Comparing the results of the statistical analysis to the output of the designed optimal ANN reveals that the Neural Network performs more accurately.

1. Introduction

The Tropospheric Ozone O_3 is an air pollutant of particular interest, as its presence in the lower atmosphere degrades the air quality and affects negatively not only human beings but also plant life, constructions and materials. Medical studies have revealed that in high concentrations it can be blamed for inflammation and irritation of the respiratory system particularly during heavy physical activity, reduced lung activity, aggravation of asthma as well as ocular diseases [5], [6], [24].

It is produced near the ground through a series of reactions between Volatile Organic Compounds (VOCs), nitric oxide (NO) and nitrogen dioxide (NO_2) under the influence of sunlight. In other words it is a secondary pollutant, whose main anthropogenic source is the photochemical reaction between precursors emitted mainly during combustion of fossil fuels in industry and transportation. It is well documented that the production of ozone in an urban area is highly depended on the solar radiation and the air temperature. In a recent work [18] it was found that in the metropolitan area of Athens radiation levels greater than 600W/m^2 combined with temperature observations greater than 28^{o}C and wind speeds lower than 3m/sec imply favorable meteorology to photochemical ozone production.

On the other hand, in the absence of other oxidizing agents, the major destruction mechanism is the oxidation of nitric oxide (NO) to form (NO_2). Other destruction procedures include surface deposition and oxidation of sulphur dioxide (SO_2).

Due to the fundamental role of the tropospheric ozone on the air quality, numerous scientific studies examining the relationships between meteorological conditions, air pollutant parameters and ozone concentrations have been published. In most of these research efforts the O_3 modelling relies, either on the statistical analysis of current and previous meteorological conditions and pollutant precursors, or on theories related to physical and chemical processes in the atmosphere; see e.g. [10], [27], [3], [28], [30], [18], [33], [31], [23]. Artificial Neural Networks have also been developed to meet the need for accurate ozone concentrations forecasting. These ANN models provide a new and sophisticated technique in order to model the ground-level O_3 concentrations or the peak levels in major cities; see e.g. [32], [25], [15], [1], [4], [8], [2], [9], [26].

This manuscript presents a research effort that has been contacted towards the design and development of an Artificial Neural Network capable of estimating the ozone concentrations in the center of the city of Athens. More specifically the research focuses on the Patission Street monitoring station, where the headquarters of the Air Pollution and Noise Control Division of the Greek Ministry of Environment are situated. Although the data of this station are not representative of the whole city air quality, as they are typical of the city centre only, still they are used here for 2 main reasons. Firstly, they consist of a high-reliable continuous record, and, secondly, most decisions about air pollution temporal abatement measures in Athens are based on the data of this station. Therefore, the site of the Patission Street appears to be of extreme interest. This work is the continuity of a research effort which started by constructing the best linear equation for ozone modelling through various combinations of Regression Analysis and Multivariate Methods such as Principal Component Analysis. Hence, another objective of the study is to compare the statistical techniques with the ANN modelling procedures. The developed ANN model can then be used for the estimation of the ozone concentration in other major cities with similar climatic, topographic and traffic characteristics.

2. **Materials and methods**

2.1. Research area

In the Greater Area of Athens (GAA) more than 2 millions of vehicles are registered, while the industrial activities are centered in the southwestern and western part. The main air pollution sources are automobiles, industrial activities and central heating systems (during the cold period of the year). Some of the main air pollutants routinely recorded are carbon monoxide, nitric oxide, nitrogen dioxide and ozone. All the pollutants present an almost single intra-annual variation. In general, the primary pollutants display lower concentrations during summer time, when central heating systems are not operating and the traffic is reduced due to the Athenians' summer vacations. On the other hand, ozone displays maximum values in accordance with the annual variation of the solar radiation and temperature; see e.g. [19], [13], [12].

The present study focuses on the Patission Street monitoring station. The position of the station can be considered as the centre of the Athens Metropolitan Area and its sampling inlet is about 10m above the street level. The area is a typical street canyon with buildings close to 30m high. The ventilation is usually weak and the traffic volume high. As a result, the primary traffic pollutants display high concentrations, while ozone concentrations display low levels in accordance with the photostationary equilibrium, which dictates that high concentrations of NO lead to low concentrations of O_3 and vice-versa. The following figure 1 shows the topographic details of the research area.

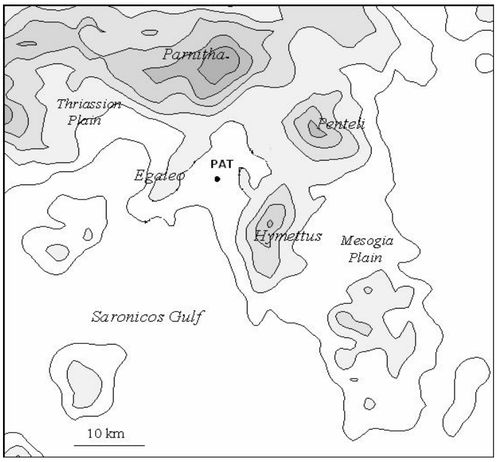

Figure 1. Topography map of the Attica Peninsula. Contours are drawn every 200m

2.2. Basic ANN keypoints

Artificial Neural Networks (ANN) are a special kind of Intelligent Systems whose computing power is achieved through their massively parallel distributed structure and their ability to learn and therefore generalize [11]. The ANN technology is rooted in many different disciplines such as engineering, mathematics, physics, neurosciences and statistics. This is due to their ability to learn from input data either in supervised or in unsupervised mode.

A typical ANN consists of several units (called neurons) that have a very limited computing capability. However the neuron combination that forms the complete network is capable of performing very complicated tasks. The neurons use various rules that combine the input signals and an activation rule that processes the combined signal and calculates the output [7]. The output signals are transmitted among the neurons through the connections known as weights. The weights excite or inhibit the signal according to the case and the desired result. During the training/learning and testing/recall phases weight adjustments take place aiming to the determination of the optimal ANN.

It is very important that an ANN is considered successful only when it proves its ability to generalize [11]. Generalization is a measure of an ANN's ability to produce reasonable output for inputs that are not encountered during the training phase [11].

2.3. Determining the input vector

The *Input Layer* of the Artificial Neural Network (ANN) developed in this study consists of ten neurons corresponding to the ten independent parameters. More specifically, mean hourly values gathered only in the day-light period, concerning seven meteorological and three pollutant parameters for the high summer season (June-August) for a 4-year period 2001-2004 have been gathered. The selection of the above months was based on the results of a previous study [18], which indicated that these months display favorable meteorology (in terms of temperature, solar radiation and wind speed) to ozone production. The pollution parameters that were used as input in the Neural Network are carbon monoxide (CO in mgm^{-3}), nitric oxide (NO in μgm^{-3}) and nitrogen dioxide (NO_2 in μgm^{-3}). The meteorological parameters that were used as input are the mean air temperature (T in $°C$), the total solar radiation (Q in Wm^{-2}), the mean pressure at sea level (P in hPa), the relative humidity (RH in %), the mean wind speed (WS in $m\,s^{-1}$), the NW-SE direction wind component (u' in $m\,s^{-1}$) and the SW-NE direction wind component (v' in $m\,s^{-1}$), normal to u'.

The selection of the u' and v' components instead of the conventional ones, u (W-E) and v (S-N), was considered necessary as u' is almost parallel to the Saronic Gulf coast and v' to the direction of the sea breeze circulation and the main axis of the Athens Basin (see Fig. 1). Only the day light time period, i.e. from 7:00 to 19:00 LST, was used for all the parameters, since this is documented as the most important photochemical production period [18]. Thus, in total, 4611 data records were used; 66% of them were used for the training procedure, while 34% of them were used for the testing procedure. Missing values were excluded. It is noted that the separation of the data set in the above two groups was performed randomly so that each piece of data could have equal chances to be picked.

3. Experimenting for the determination of the optimal ANN

Various experiments with different architectures, optimization algorithms and learn or transfer functions were performed in order to determine the optimal ANN. Each experiment in its initial stage comprised a large number of training cycles.

The Input Vector consisted of ten parameter values, while the output included only one neuron, which corresponds to the Ozone concentration. Several Artificial Neural Network models revealed a good performance during the training phase. However, the good performance of each ANN was confirmed in the final testing phase, in terms of its ability to generalize.

Dozens of supervised ANN types were tried including *Back Propagation* ANN (BP), *Modular* ANN, *General Regression* ANN and *Radial Basis Function* Neural Networks (RBFNN). Hence, numerous different topologies were applied. As it is shown in the following Table 1, training and testing experiments included iterations with the use of the *Tangent Hyperbolic (TanH)*, the *Sigmoid*, the *DNNA* and the *Sine* transfer functions. The Learning Rule applied in the experiments was the Extended Delta Bar Delta (ExtDBD), the Quick Propagation (Quick Prop), the Norm-Cum-Delta, the Delta and finally the Max Prop [17].

The TanH is a smooth version of a [-1,1] step function and it can be considered as a bipolar version of the Sigmoid function, which is a smooth version of a [0,1] step function. Generally, the TanH is given by the following equation 1.

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \qquad (1)$$

The BP Algorithm is the most popular and effective local algorithm for adjusting the weights of a multi-layer neural network [21], [22]. Back Propagation ANN are very common and they have been used in numerous prediction modelling applications [29]. For many years there were no available rules for the weight-update in multilayered ANN undergoing supervised training. In the 1970s Werbos developed a technique for adapting the weights but it was Rummelhart who defined a weight adoption rule called *Back-Propagation*. In a network where the Processing Elements (PE) of a layer are connected to every sngle PE in the upper layer, the BP algorithm performs at first a forward sweep, from the input to the output layer, and then a backward sweep from the output towards the input layer. In the backward sweep, error values are propagated back through the ANN, in order to determine the way in which the weights will be changed during the training.

Modular Neural Networks (MODANN) are Adaptive Mixtures of “local experts” [7]. They were introduced by Jacobs, Jordan, Nowlan and Hinton [17]. In fact MODANN contain several separate ANN models. They comprise a group of BP networks, referred to as “local experts”, which compete to each other in learning different aspects of a problem. A “Gating Network” plays the role of the referee and it controls the competition. It learns to assign different parts of the data vector to the different networks. It is the Gating Network that suggests the optimal local expert for a given problem. According to Jacobs, Jordan, Nowlan and Hinton Modular ANN can be used for System Modelling, Prediction, Classification and Filtering [16].

RBFNN are networks having an internal representation of radially symmetric hidden neurons. For a neuron to be radially symmetric it needs to have the following three constituents: a) A center which is a vector in the input space. b) A distance measure to determine how far an input vector stands from the center. In this case standard Euclidean distance is used. c) A transfer function (using a single variable) which determines the output of the Neuron by mapping the output of the distance function. Usually a Gaussian function is applied, producing stronger values when the distance is small. The output of a pattern unit is a function of the distance between an input vector x and the stored center c exclusively. This is shown in the following equation 2.

$$f(x) = \phi(\|x - c\|) \qquad (2)$$

In the RBFNN, the hidden layer is fully connected to a linear output layer. The pattern units l_k are defined by using the Euclidean summation as follows in equation 3 (X is the input vector and c is the stored center), whereas a Gaussian transfer function is applied as shown in equation 4 [20], [14].

$$l_k = \|X - c_k\| = \sqrt{\sum_{i=1}^{N} (x_i - c_{ki})^2} \qquad (3)$$

$$U_k = \exp\left(-0.5 * \frac{l_k^2}{\sigma_\kappa^2}\right) \qquad (4)$$

3.1. Evaluation instruments applied

In all of the prediction modelling efforts the target is the minimization of the differences between the predicted values and the actual experimental data. All of the experiments performed here included 1000, 2000, 3000, 4000 and 5000 iterations. To avoid over-Training the iterations were terminated every time that the performance started to drop.

Two ANN instruments, the Root Mean Square Error (RMS Error) and the Confusion Matrix (CM) were used to check the ANN's validity. The RMS Error adds up the squares of the errors for each PE in the output layer, divides by the number of PEs in order to obtain an average and finally estimates the square root of that average.

The CM is a graphical way of measuring and displaying the performance of an ANN and it can be used in both training and testing processes. The CM correlates the actual results of the ANN to the desired results in a visual display and it consists of a matrix containing a number of small cells called bins [16]. The ANN with the optimal configuration must have the bins on the diagonal from the lower left to the upper right. The value of the vertical axis in the produced histogram is the Common Mean Correlation (CMC) coefficient of the desired (d) and the actual (predicted) output (y) across the Epoch. The CMC is calculated by the following equation 5.

$$CMC = \frac{\sum \left(d_i - \bar{d}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum \left(d_i - \bar{d}\right)^2 \sum \left(y_i - \bar{y}\right)^2}}, \text{ where}$$

$$\bar{d} = \frac{1}{E}\sum_1^n d_i \quad \text{and} \quad \bar{y} = \frac{1}{E}\sum_1^n y_i \qquad (5)$$

It is noted that *d* stands for the desired values, *y* for the predicted values where *i* ranges from 1 to *n* (the number of cases in the data training set) and *E* for the Epoch size. The epoch size is the number of training data sets presented in the ANN learning cycles among weight updates.

4. Training and Testing results

The training vector consisted of 3442 actual data records and the testing vector included 1169 data records for all of the summer months. In all of the experiments the candidate optimal ANN had a maximum number of two Hidden sub-layers in an effort to keep the optimal ANN as simple as possible.

All of the above evaluation instruments were applied in both phases (training and testing). The Epoch value was kept stable to the value of 16 in all of the iterations.

Two ANN were characterized as having the best fit from the Back Propagation ones. They both had ten (10) neurons in the Input Layer corresponding to the ten independent parameters, a single Hidden Layer consisting of eleven (11) neurons and an output layer of one (1) neuron corresponding to the ozone concentration. The first BP best fit ANN used the Sine transfer function and the ExtDBD learning rule, whereas the second ANN used the Tangent Hyperbolic (TanH) transfer function and the ExtDBD learning rule. The values of the Correlation R^2 and the RMS Error for the first ANN were 0.9286 and 0.1323 respectively in the training phase, whereas in testing the R^2 = 0.9153 and the RMS Error=0.1358. The second best fit BP ANN had R^2= 0.9676 and RMS Error=0.1404 in the training phase, R^2= *0.9144* and RMS Error=*0.1253* in the testing. The following Table

1 shows that apart from the optimal BP ANN several other ANN had a very good performance as well.

Table 1. Training and Testing results at 5000 iterations

LEARNING RULE	OPTIMIZATION ALGORITHM	TRANSFER FUNCTION	Neurons in the Input Layer	Neurons in the 1st Hidden Sub-Layer	Neurons in the 2nd Hidden Sub-Layer	Neurons in the Output Layer	R^2	RMS Error
1. *ExtDBD*	*Back Propagation*	*TanH*	*10*	*11*	*0*	*1*	*Training 0.9676 Testing 0.9144*	*Training 0.1404 Testing 0.1253*
2. *ExtDBD*	*Back Propagation*	*Sigmoid*	*10*	*11*	*0*	*1*	*Training 0.8897 Testing 0.8419*	*Training 0.0641 Testing 0.0630*
3. *ExtDBD*	*Back Propagation*	*Sine*	*10*	*11*	*0*	*1*	*Training 0.9286 Testing 0.9153*	*Training 0.1323 Testing 0.1358*
4. *Quick Prop*	*Back Propagation*	*TanH*	*10*	*11*	*0*	*1*	*Training 0.9327 Testing 0.8834*	*Training 0.1709 Testing 0.1452*
5. *Norm-Cum-Delta*	*Back Propagation*	*TanH*	*10*	*11*	*0*	*1*	*Training 0.9486 Testing 0.9058*	*Training 0.1316 Testing 0.1317*
6. ***Delta-Rule***	*Back Propagation*	*TanH*	*10*	*11*	*0*	*1*	*Training 0.9091 Testing 0.8834*	*Training 0.1728 Testing 0.1473*
7. *ExtDBD*	*Modular ANN*	*TanH*	*10*	*11*	*0*	*1*	*Training 0.9602 Testing 0.9290*	*Training 0.1546 Testing 0.1186*
Gating Network of the Modular ANN (Connect Prior)			-	*4*	-	*1*		
8.	RBF ANN *(ExtDBD TanH)*		*10*	*Pattern 50*	11	*1*	*Training 0.9321 Testing 0..8901*	*Training 0.1090 Testing 0.1411*

The following figure 2 shows the architecture of the first BP Ext DBD ANN, whereas figures 3 and 4 show the output of the evaluation instruments for the same ANN in training and testing respectively. Furthermore, figure 4 presents a diagram that estimates the degree of input contribution of each independent parameter for the determination of the ozone concentration. The Input Contribution diagram reveals that the NO concentration plays the most important role in the determination of the ozone concentration, whereas CO is second with a much lower (almost negligible) contribution degree.

However, it is clearly shown in the above Table 1 that the optimal ANN has proven to be a Modular one. It had ten (10) neurons in the Input Layer corresponding to the ten independent parameters, a single Hidden Layer consisting of eleven (11) neurons and an output layer of one (1) neuron corresponding to the Ozone concentration. The Gating Network had a Hidden Layer with four (4) neurons and an Output Layer with three (3) neurons. The optimal Modular ANN had R^2= 0.9602 and RMS Error=0.1546 in the Training phase, R^2= *0.9290* and RMS Error=*0.1186* in the testing. Figure 5 presents the architecture of the Modular optimal ANN, while figure 6 displays the output of the evaluation instruments for the same ANN in the testing phase.

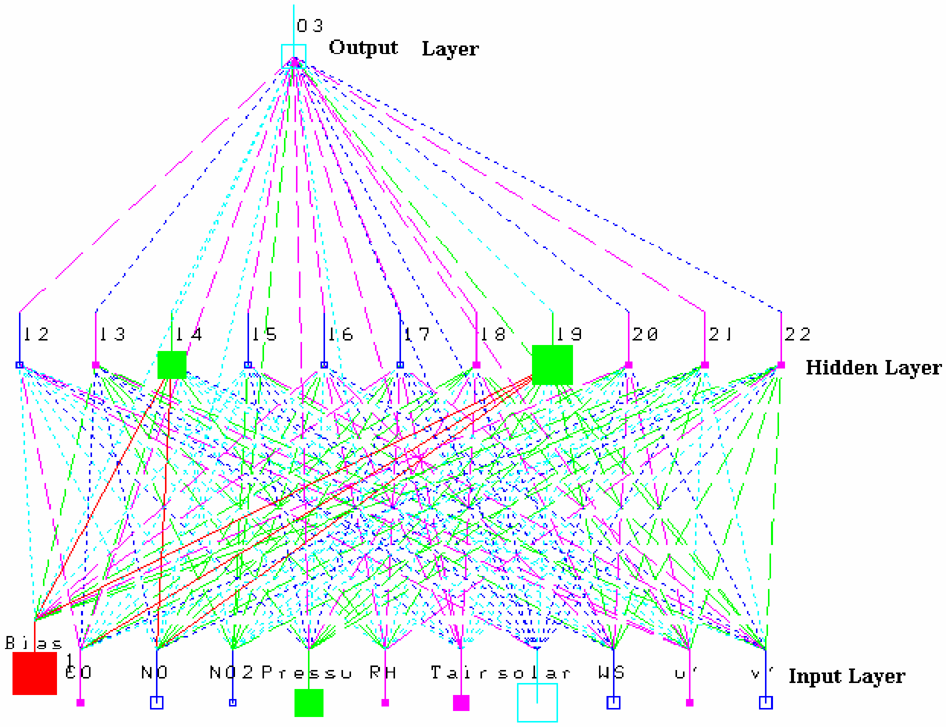

Figure 2. Architecture of the BP Ext-DBD Sine ANN

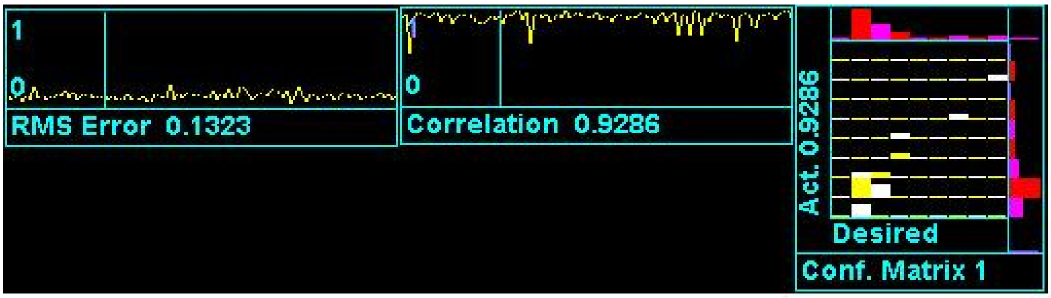

Figure 3. Evaluation instruments' values in training for the Ext-DBD-Sine BP ANN

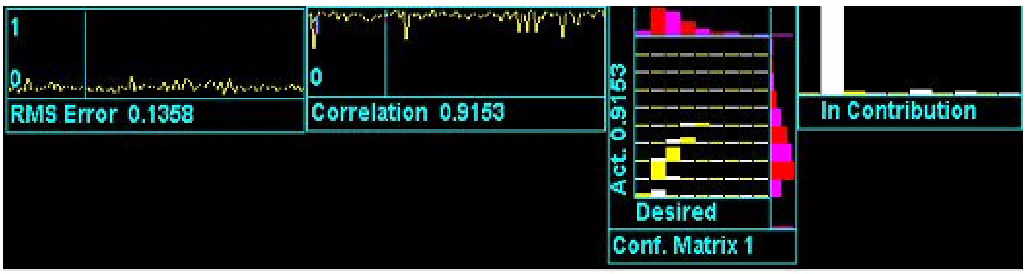

Figure 4. Evaluation instruments-input contribution in testing for the Ext-DBD-Sine BP ANN

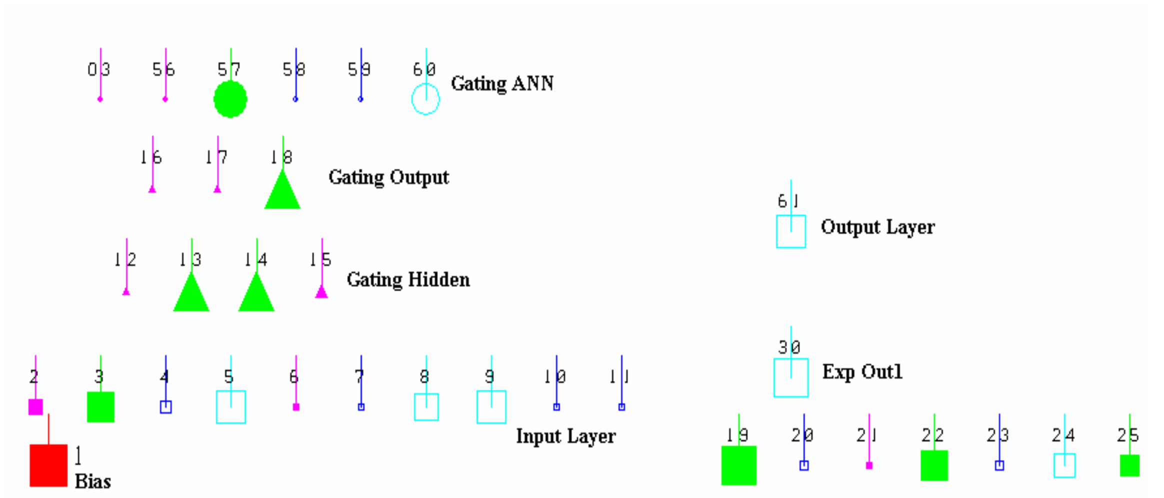

Figure 5. Architecture of the Modular optimal ANN

4.1. Reliability of the ANN

To avoid over-Training, the network was trained by performing initially 1000 iterations. The number of iterations was increased gradually (with a step equal to 1000), while the process stopped when the performance started to drop. In this way the optimal ANN was identified in 5000 iterations. The large amount of actual data used in the ANN development process (3442 in training and 1169 in testing) combined with the very good results of the evaluation instruments and the simple structure of the optimal network confirms its ability to generalize and therefore its reliability.

The following figure 7, presents a comparison between the actual and the ANN predicted ozone concentration values for 1167 cases in the testing phase with first time seen data. It can easily be seen that the compatibility is quite high.

4.2. Comparison with Statistical Analysis

In the context of the statistical approach, a Stepwise Regression Analysis was applied in the values of the independent variables (CO, NO, NO_2, pressure, solar radiation, relative humidity, air temperature, wind speed, u' and v' components), in order to produce prediction models for the logarithmic transformation of the ozone concentrations $\ln[O_3]$. It is noted that the logarithmic transformation of ozone $\ln[O_3]$ was used instead of O_3 because its frequency distribution is closer to the normal and it is well-known that the regression analysis works better with normal variables. The coefficient of determination R^2 can be interpreted as the percentage of the variation of the predictand $\ln[O_3]$ that is accounted for by the regression model. Even though, the R^2 value was found to be 0.86, a deeper look in the results revealed strong multicollinearity evidence. The low Tolerances (down to 0.15) and the high Variance Inflation Factors (up to 6.54) indicated that the predictors were strongly intercorrelated to

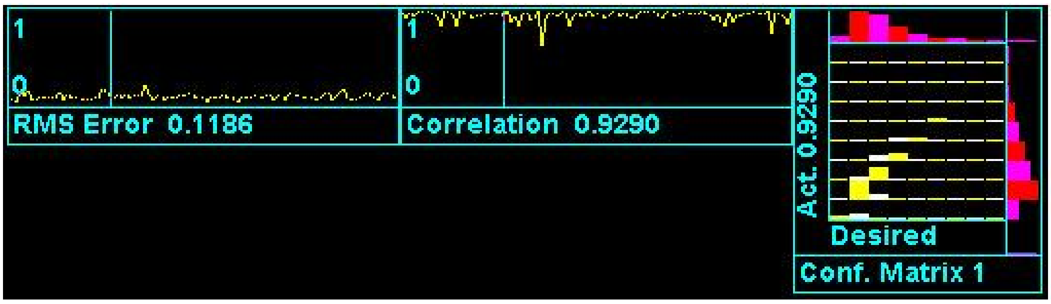

Figure 6. Evaluation instruments' values in testing for the Modular ANN

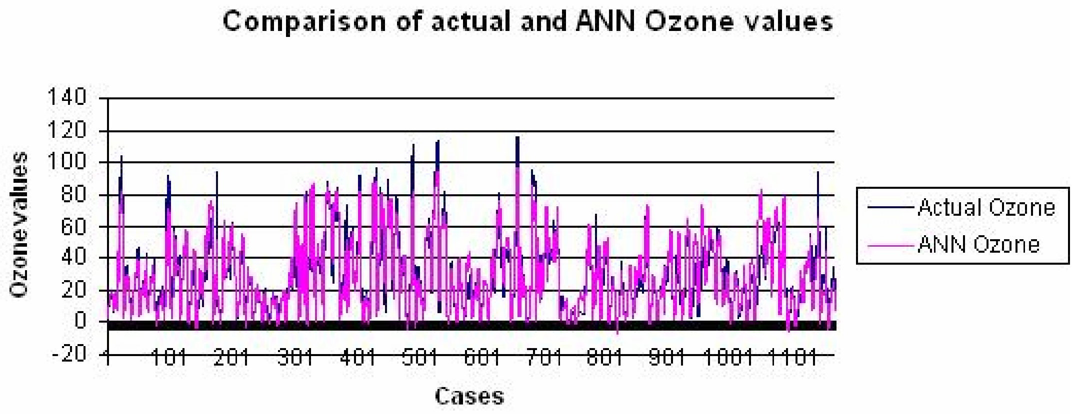

Figure 7. Comparison between actual and forecasted ozone for the Modular ANN

each other, so that small changes in the data values could lead to large changes in the estimates of the coefficients.

To overcome the problem of multicollinearity, Principal Component Analysis with a varimax rotation was applied. in the initial data matrix, in order to reduce the number of the original intercorrelated variables. At first, the analysis was carried out keeping all the PCs and then only the strongest (in terms of loadings) were retained. Thus, by excluding the variables that were not highly correlated to a PC (loadings < 0.70) the remaining variables were more-or-less uncorrelated to each other.

Next, free of the multicollinearity problem, a Stepwise Regression Analysis was applied in the remaining original variables. The coefficient of determination was then found to be 0.8356. This means that approximately 84% of the variation in O_3 values was explained by the produced equation. However, in the case of the optimal ANN the coefficient of determination R^2 was found to be 0.9602 in the training phase and 0.9290 in the testing phase. Consequently, almost 93% of the variation in ozone values has been accounted for. It is now clear that in the case of the Patission Street the use of ANN offers greater reliability in the issue of ozone modelling on its precursors.

5. Conclusions

The objective of the study was to design and develop a series of ANN models for the estimation of the tropospheric ozone concentrations in a heavy traffic road in the center of Athens, Greece. For this reason a big number of ANN types were tested. Several of them revealed high accuracy in terms of high levels of R^2. However, the optimal ANN was proven to be a Modular one, which has ten (10) neurons in the input layer, a single hidden layer consisting of 11 neurons and an output layer consisting of one (1) neuron. This Modular ANN revealed R^2=0.9602 in the training phase and R^2=0.9290 in the testing phase. The high levels of R^2 combined with the simple structure guarantees its reliability and generalization ability. Therefore, it can serve as prototype and can be tested and applied in other major urban centers with similar climatic, topographic and traffic conditions. It is among the authors' intentions to cooperate in the near future with other major European cities authorities in order to investigate the potential usage of these ANN models in them.

Finally, compared to our previous statistical analysis performed on the same data, most of the designed ANN models have proven to perform more efficiently. This can be considered as a strong motivation for our research team to continue this effort in other areas in the near future.

6. References

[1] S. Abdul, A. Wahab, and S.M. Al-Alawi, Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. Environmental Modelling & Software, 17 (Issue 3), 2002, pp. 219-228.

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

[3] C. Ambroise and Y. Grandvalet, Prediction of ozone peaks by mixture models. Ecological Modelling, 145, 2001, pp. 275-289.

[4] E. Balaguer Ballester, G. Camps i Valls, J.L.Carrasco – Rodriguez, E. Soria Olivas and S. Del Valle – Tascon, Effective 1-day ahead prediction of hourly surface ozone concentrations in eastern Spain using linear models and neural networks. Ecological Modelling, 156, 2002, pp. 27-41.

[5] M Brauer and J.R. Brook, Ozone personal exposures and health effects for selected groups residing in the Fraser Valley. Atmospheric Environment, 31, 1997, pp. 2113-2121.

[6] R.T. Burnett., J.R. Brook, W.T.Yung, R.E. Dales and D. Krewski, Association between Ozone and Hospitalization for Respiratory Diseases in 16 Canadian Cities, Environmental Research, 72, 1997, pp. 24-31.

[7] R. Callan, The Essence of Neural Networks. Prentice Hall, UK, 1999.

[8] A. Chaloulakou, M. Saisana and N. Spyrellis, Comparative assessment of neural networks and regression models for forecasting summertime ozone in Athens. The Science of the Total Environment, 313, 2003, pp. 1-13.

[9] A.L. Dutot, J. Rynkiewicz, F.E. Steiner and J. Rude, 24-h forecast of ozone peaks and exceedance levels using neural classifiers and weather predictions. Environmental Modelling & Software, In Press, Corrected Proof, Available online 3 November 2006.

[10] J.B. Flaum, S.T. Rao and L.G. Zubenko, Moderating the influence of meteorological conditions on ambient ozone concentrations. Journal of the Air and Waste Management Association, 46, 1996, pp. 35-46.

[11] S. Haykin, Neural Networks: A comprehensive foundation. Mcmillan College Publishing Company, New York, 1999.

[12] C.G. Helmis, D.N. Asimakopoulos, K.H. Papadopoulos, P. Kassomenos, J.A.Kalogiros, P.G. Papageorgas, and S. Blikas, Air mass exchange between the Athens Basin and the Messogia Plain of Attica, Greece. Atmospheric Environment, 31, 1997, pp.3833-3849.

[13] P. Kassomenos, V. Kotroni and G. Kallos, Analysis of climatological and air quality observations from Greater Athens Area. Atmospheric Environment, 29B, 1995, pp. 3671-3688.

[14] T. Kohonen, Self-Organization and Associative Memory. Second edition, Springer-Verlag. New York, 1988.

[15] D. Melas, I Kioutsioukis and I Ziomas, Neural Network Model for predicting peak photochemical pollutant levels. Journal of the Air and Waste Management *Association*, 50, 2000, pp. 495-501.

[16] Neuralware, Getting started. A tutorial for Neuralworks Professional II/PLUS. Carnegie, PA, USA, 2001.

[17] Neuralworks Reference Guide, pp. RF-40-46, NeuralWare, Carnegie, PA, USA, 2001.

[18] A.K Paschalidou and P.A. Kassomenos, Comparison of air pollutant concentrations between weekdays and weekends in Athens, Greece for various meteorological conditions**.** Environmental Technology, 25, 2004, pp. 1241-1255.

[19] C. Pilinis, P. Kassomenos and G. Kallos, Modelling of photochemical pollution in Athens, Greece: Application of the RAMS-CALGRID modelling system. Atmospheric Environment, 27B, 1993, pp. 353-370.

[20] J Platt, A Resource Allocating Network for Function Interpolation. Neural Computation, 3, 1991, pp. 213-225.

[21] D.E. Rummelhart, G.E. Hinton and R.J. Wiliams, Learning Internal Representations by Error Propagation. Institute for Cognitive Science Report 8506. San Diego, University of California, 1985.

[22] D.E. Rummelhart, G.E. Hinton and R.J. Wiliams, Learning Internal Representations by Error Propagation. Parallel Distributed Processing, Explorations in the Microstructure of Cognition. Vol 1, Foundations. Cambridge, MA: MIT Press, 1986.

[23] U. Schlink, O. Herbarth, M. Richter., S. Dorling, G. Nunnari, G. Cawley and E. Pelikan, Statistical models to assess the health effects and to forecast ground-level ozone. Environmental Modelling & Software, 21 (Issue 4), 2006, pp. 547-558.

[24] M. Smith-Doron, D. Stieb, M. Raizenne, J. Brook, R. Dales, J. Leech, S. Cakmak and D. Krewski, Association between ozone and hospitalisation for acute respiratory diseases in children less than 2 years of age. American Journal of Epidemiology, 153, 2000, pp. 444-452.

[25] G. Soja and A.M. Soja, Ozone indices based on simple meteorological parameters: potential and limitations of regression and neural network models. Atmospheric Environment, 33, 1999, pp. 4229-4307.

[26] S.I.V. Sousa, F.G. Martins, M.C.M. Alvim-Ferraz and M.C. Pereira, Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environmental Modelling & Software, 22 (Issue 1), 2007, pp. 97-103.

[27] M. Statheropoulos, N. Vassiliadis and A. Pappa, Principal component and canonical correlation analysis for examining air pollution and meteorological data. Atmospheric Environment, 32, 1998, pp. 1087-1095.

[28] M.L. Thompson, J. Reynolds, L.H. Cox, P. Guttorp and P.D. Sampson, A review of statistical methods for the meteorological adjustment of tropospheric ozone. Atmospheric Environment, 35, 2001, pp. 617-630.

[29] S. Van Looy, J. Meeus, B. Wyns, B. Vander Cruyssen, L. Boullart and F. Keyser, Comparison of Machine Learning models for prediction of dose increase in patients with rheumatoid arthritis. Proceedings of the 9th International Conference of Engineering Applications of Neural Networks, Lille France, 2005, pp. 189-196.

[30] F.M. Vukovich and J. Sherwell, An examination of the relationship between certain meteorological parameters and surface ozone variations in the Baltimore – Washington corridor, Atmospheric Environment, 37, 2003, pp. 971-981.

[31] E.K. Wise and A.C. Comrie, Meteorologically adjusted urban air quality trends in the Southwestern United States. Atmospheric Environment, 39, 2005, pp. 2969-2980.

[32] J. Yi and R. Pybutok, A neural network model forecasting for prediction of daily maximum ozone concentrations in an industrialized urban area. Environmental Pollution, 92, 1996, pp. 349-357.

[33] A. Zolghadri, M. Monsion, D. Henry, C. Marchionini and O. Petrique, Development of an operational model-based warning system for tropospheric ozone concentrations in Bordeaux, France. Environmental Modelling & Software, 19 (Issue 4), 2004, pp. 369-382.