International Journal of Computational and Applied Mathematics.
ISSN 1819-4966 Volume 2 Number 2 (2007), pp. 125–138

http://www.ripublication.com/ijcam.htm

Artificial Neural Network Modelling of the surface Ozone Concentration

L. S. Iliadis[1], S. I. Spartalis[2], A. K. Paschalidou[3] and P. Kassomenos[4]

[1]*Department of Forestry & Management of the Environment & Natural Resources, Democritus University of Thrace, 193 Padazidou st., 68200, Nea Orestiada, Greece*
[2]*Department of Production Engineering & Management, School of Engineering, Democritus University of Thrace, University Library Building, 67100, Xanthi, Greece*
[3]*Department of Physics, Laboratory of Meteorology, University of Ioannina, 45110, Ioannina, Greece*
E-mail: [1]liliadis@fmenr.duth.gr, [2]sspart@pme.duth.gr

Abstract

This manuscript presents the design and the development of an Artificial Neural Network (ANN) model that estimates the surface ozone concentrations when the values of other pollutant and meteorological parameters are already known for the case of Athens suburb Lykovryssi. This is the characteristic suburb which is very close to a major city and at the same time far from the city center. Thus the developed ANN can constitute a potential model for all suburbs with similar characteristics. The large amount of data records used as input and the good testing results show the generalization ability of the developed ANN. The Training has been performed for different numbers of iteration cycles in order to avoid over-Training. Principal component analysis and stepwise regression analysis were performed for the same area in the past. Comparing the results of the statistical analysis to the output of the designed optimal ANN, we have discovered that the Neural Network performs more accurately.

Key words: Artificial Neural Networks, Pollution of the Atmosphere, Tropospheric Ozone

Introduction

Artificial Neural Networks are a special approach in the creation of Intelligent Systems as they neither use knowledge representation nor they adopt specially designed search algorithms. ANN are based on biological models as they use structures and processes similar to the ones of the human brain. The computing power of ANN is achieved through their massively parallel distributed structure and their ability to learn and therefore generalize (Haykin, 1999). The application areas of ANN are expanded but not limited to various Engineering, Financial and Environmental domains to provide smart solutions towards forecasting or clustering.

The Tropospheric Ozone O_3 is considered to be one of the most pervasive and potentially harmful air pollutants. It is a critical atmospheric species, which drives much of the tropospheric photochemistry. It is also considered responsible for regulating the tropospheric oxidation capacity and it is the main ingredient of the photochemical smog. It is well-documented that its presence in the lower atmosphere degrades the air quality and has negative effects on human health, plant-life, constructions and materials. Various medical studies have revealed that it can be blamed for inflammation and irritation of the respiratory tract, particularly during heavy physical activity, as well as ocular diseases (Brauer and Brook, 1997), (Burnett et al. 1997), (Smith et al. 2000).

The ambient ozone has both natural and anthropogenic sources. The main natural source is the downward flux from the stratosphere. On the other hand, the main anthropogenic source is the photochemical production from precursors emitted primarily during combustion of fossil fuels by industry and transportation, i.e. O_3 is a secondary pollutant.

More specifically, the tropospheric ozone is formed when volatile organic compounds (VOCs), nitric oxide (NO) and nitrogen dioxide (NO_2) react chemically under the influence of heat and sunlight (Paschalidou and Kassomenos, 2004). Hence, NO_x and VOCs are referred to as ozone precursors. The most dominant reaction for ozone formation is the photolysis of nitrogen dioxide. On the other hand, in the absence of other oxidizing agents, the major destruction mechanism for O_3 is the oxidation of nitric oxide (NO) to form NO_2. Other destruction mechanisms include surface deposition and oxidation of SO_2.

The photochemical nature and the temperature dependence of the reactions involved in ozone formation/destruction indicate that the whole mechanism is controlled by meteorological parameters.

Numerous modeling scientific studies examining the relationships between meteorological conditions, air pollutant parameters and ozone concentrations have been published. Most of the times, the O_3 modeling 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 or they estimate the ozone health effects (Sousa et al. 2007), (Schlink et al., 2006), (Zolghadri et al., 2004), (Ambroise and Grandvalet, 2001), (Flaum et al., 1996), (Paschalidou and Kassomenos, 2004), (Statheropoulos et al., 1998),(Thompson et al., 2001), (Vukovich and Sherwell,2003), (Wise and Comrie, 2005).

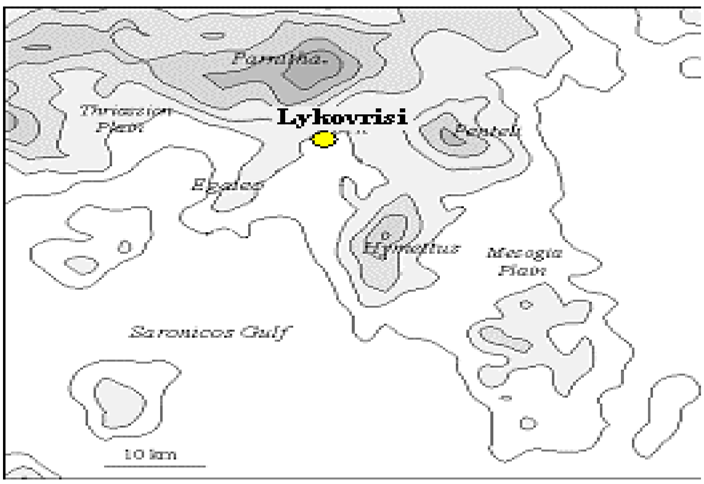
Artificial Neural Networks have also been developed in order to model the ground-level O_3 concentrations or the peak levels in major cities of France, Spain and Greece (Abdul et al., 2002), (Agirre-Basurko et al. 2006), (Balaguer et al., 2002), (Chaloulakou et al., 2003), (Dutot et al. 2006), (Melas et al., 2000), (Soja and Soja, 1999), (Yi, Prybutok, 1996).

The present work aims to design and develop the optimal ANN for the ozone concentration estimation in the site of Lykovryssi, Athens. Lykovryssi is a suburb very close to a major city of four million people and at the same time far away from the city center. For this reason the area of Lykovryssi appears to be very interesting. This research effort will reveal the contribution of a major urban center to the ozone concentrations in a nearby suburb. From this point of view and due to the fact that suburbs very close to Athens have not been investigated before for the levels of ozone concentrations, this research effort can be considered as a very useful one. The developed ANN can be considered as a prototype for areas close to major city centers with similar morphological and climate characteristics. It is the continuity of a research effort which started by constructing the best linear equation for ozone modeling 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 modeling procedure.

Materials and methods

Area of research

This research effort focuses on the site of *Lykovryssi* which is located away from the city centre of Athens, on the north, and it is rather a residential-background station. Thus primary traffic pollutants display low concentrations, while ozone concentrations display high levels, in accordance with the photostationary equilibrium, which dictates that low concentrations of NO lead to high concentrations of O_3 and vice-versa. For the above reason the specific area was chosen.



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

ANN description

The ANN technology is rooted in many different disciplines like mathematics, physics, neurosciences, statistics, and several sectors of engineering. Successful applications of ANN can be found in so many diverse fields due to their ability to learn from input data either in supervised or in unsupervised mode.

Although an ANN consists of units that have a very limited computing capability, still the complete network is capable of performing a very complicated task, as many of the above units are connected to each other. The actual power of ANN is due to their massively parallel distributed structure. ANN are networks of interconnected elements, inspired from studies of biological nervous systems. Specifically, an ANN is a collection of units (the so-called Neurons or Nodes or Processing Elements) that are connected in some pattern which allows communication among them. Neurons are simple processors whose computing ability is typically restricted to a rule for combining input signals and an activation rule that processes the combined input signal and calculates the output signal (Callan, 1999). Output signals are transmitted among Neurons through connections known as weights. The weights excite or inhibit the signal according to the case and the desired result.

ANN are considered to be successful only when they are able to generalize (Haykin, 1999). Generalization refers to their ability to produce reasonable output for inputs not encountered during the training phase (Haykin, 1999).

Determining the nature and structure of input data

The i*nput layer* of the Artificial Neural Network (ANN) developed in this study consists of eleven Neurons corresponding to eleven independent parameters. More specifically, mean hourly values gathered only in the day-light period, concerning seven meteorological and four 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 (Paschalidou and Kassomenos, 2004) 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 mgr^{-3}), nitric oxide (NO in μgr^{-3}), nitrogen dioxide (NO_2 in μgr^{-3}) and the particulate matter (PM_{10} in μgr^{-3}). The meteorological parameters that were also used as input are the mean air temperature (T in $°C$), the total solar radiation (Q in $W\ m^{-2}$), 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 (Paschalidou and Kassomenos, 2004). Thus, in total, *4232* data records were used; 74% of them were used for the training procedure, while 26% 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.

Designing the optimal ANN

The Training-Testing Processes

The two most important iterative phases in the ANN design are the Training (also called Learning) and the testing. Learning is the process of adapting or modifying the connection weights according to the stimuli being presented at the input buffer or optionally at the output buffer. A stimulus presented at the output buffer corresponds to a desired response to a given input. This desired response is provided by a knowledge “teacher”. In such a case the learning is called “supervised learning”. The requirements of a supervised learning strategy are a suitable weight adjustment mechanism and suitable error functions.

The classical architecture of an ANN consists of the input layer, the hidden layer and the output layer. Other existing variation will be described briefly in this paper. The hidden layer is the place where the data is being processed and it may consist of one or more sub-layers depending on the designer’s view.

The selection of the proper ANN model always requires the performance of several training experiments. The Feed Forward network structure with input, output and hidden layers varying from 1 to 3 applying Optimization algorithms was tried in the training phase of this project.

The good performance of an ANN in the Training phase is a strong indication of its ability to produce reliable results. Obviously, the reliability of the ANN is measured in terms of its ability to generalize.

ANN design

Applied Architectures

The first step in the ANN development is always the determination of its optimal Topology and structure through various experiments comprising several training cycles.

The Input Vector consists of eleven parameter values. Several dozens of Artificial Neural Network models and topologies were tried in both training and testing phases before the determination of the optimal one, while thousands of iterations were performed.

More specifically, *Back Propagation* (BP) ANN, *Modular* ANN, *General Regression* ANN, and *Radial Basis Function* Neural Networks (RBFNN) have been developed and trained. Hence, numerous different topologies were applied, while several combinations of Optimization and Transfer functions were used. Thus, the *Tangent Hyperbolic (TanH),* the *Sigmoid*, the *DNNA* and the *Sine* transfer functions with the Extended Delta Bar Delta (ExtDBD) learning rule were adopted during training and testing (Neuralworks, 2001). 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.

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

The BP Algorithm is the most popular and effective local algorithm for adjusting the weights of a multi-layer neural network (Rummelhart et al., 1985), (Rummelhart et al., 1986). For many years there were no available rules for the weight-update in multilayered ANN undergoing supervised training. Rummelhart first defined a weight adoption rule called *Back-Propagation* (Rummelhart et al., 1985). In a network where the Processing Elements (PE) of a layer are connected to every single 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.

The term Modular Neural Network refers to the Adaptive Mixtures of “local experts” as proposed by Jacobs, Jordan, Nowlan and Hinton (Neuralworks, 2001). It consists of a group of BP networks referred to as “local experts” competing to learn different aspects of a problem. A “Gating Network” controls the competition and learns to assign different parts of the data space to different networks. When only one ANN is appropriate for a given problem, the Gating Network tends to favor just one of the local experts. According to Jacobs, Jordan, Nowlan and Hinton Modular ANN can be used for System Modelling, Prediction, Classification and Filtering (Neuralware, 2001). Back Propagation ANN have been developed recently by many researchers (Van Looy et al., 2005).

RBFNN are networks having an internal representation of hidden neurons which are radially symmetric. 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.

$$f(x)=\phi(\|x-c\|) \tag{2}$$

According to the architecture of a RBFNN, the hidden layer of pattern units 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. For a more detailed description see (Platt, 1991) (Kohonen, 1988).

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

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

ANN Evaluation instruments

In this research effort the choice of the optimal configuration has been based on minimizing the difference between the ANN predicted values and the actual experimental data after 5000 iterations and this was expressed by the R^2 measure. The number of iterations has been determined after several experiments in order to avoid over-Training. This is discussed in the following paragraphs. Additionally, two ANN instruments, the Root Mean Square Error (RMS Error) and the Confusion Matrix (CM) were used to check the ANN 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. Hence the name root square.

The CM provides an advanced way of measuring the ANN's performance during the "learn" and "recall" phases. It allows the correlation of the actual results of the ANN to the desired results in a visual display (Neuralware, 2001). This is achieved by providing the user with a visual indication of how well the ANN is performing. The network with the optimal configuration must have the bins (the cells in each matrix) on the diagonal from the lower left to the upper right. An important aspect of the CM is that 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.

$$CMC = \frac{\sum\left(d_i - \bar{d}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum\left(d_i - \bar{d}\right)^2 \sum\left(y_i - \bar{y}\right)^2}}, \text{ where } \bar{d} = \frac{1}{E}\sum_1^n d_i \quad \text{and} \quad \bar{y} = \frac{1}{E}\sum_1^n y_i \quad (5)$$

It should be clarified 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, which is the number of training data sets presented in the ANN learning cycles among weight updates.

Results

The training process used 3114 data records (out of the 4232 totally available) for all of the summer months. The rest of the records were kept unseen for the testing. The following Table 1 presents the training and the testing results, where numerous ANN architectures were used and various Transfer functions and Optimization algorithms were applied. It is noted that the candidate ANN had a maximum number of two Hidden sub-layers. This was done in an effort to keep the optimal ANN as simple as possible. As it will be discussed later the results were promising.

Table 1: Training and Testing results at 5000 iterations

TRAINING AND TESTING RESULTS Epoch=16								
LEARNING RULE	OPTIMIZATION ALGORITHM	TRANSFER FUNCTION	Number of Neurons in the Input Layer	Number of Neurons in the 1st Hidden Sub-Layer	Number of Neurons in the 2nd Hidden Sub-Layer	Number of Neurons in the Output Layer	R^2	RMS Error
1. ExtDBD	Back Propagation	TanH	11	15	0	1	Training 0.859	Training 0.123
							Testing 0.601	Testing 0.188
2. ExtDBD	Back Propagation	Sigmoid	11	15	0	1	Training 0.734	Training 0.046
							Testing 0.649	Testing 0.066
3. ExtDBD	Back Propagation	DNNA	11	5	5	1	Training 0.84	Training 0.057
							Testing 0.563	Testing 0.075
4. ExtDBD	Back Propagation	TanH	11	11	0	1	Training 0.773	Training 0.139
							Testing 0.698	Testing 0.167
5. ExtDBD	Back Propagation	Sine	11	15	0	1	Training 0.83	Training 0.122
							Testing 0.565	Testing 0.196
6. ExtDBD	Back Propagation	TanH	11	4	4	1	Training 0.704	Training 0.158
							Testing 0.656	Testing 0.176
7. ExtDBD	Modular ANN	TanH	11	9	0	1	Training 0.888	Training 0.119
							Testing 0.741	Testing 0.156
8. RBF ANN			11	Pattern 50		1	Training 0.678	Training 0.205
							Testing 0.639	Testing 0.182

Various evaluation instruments both graphical and numerical were also applied in the testing process which used 1118 data records. The Epoch value was kept stable to the value of 16 in all of the iterations.

The Modular ANN using the Tangent Hyperbolic (TanH) transfer function and the ExtDBD learning rule, with eleven PE in the input layer, one hidden sub-layer with nine PE and an output layer with a single PE, has proven to be the optimal one. Its Gating Network consisted of four neurons in the hidden Layer and one in the output one. The $R^2 = 0.8827$ and the RMS Error=0.1164 in the training phase, whereas in the testing process the $R^2 = 0.742$ and the RMS Error=0.156. Furthermore, several other ANN had a very good performance as well.

The following figure 1 shows clearly the architecture of the optimal Modular ANN, while figure 2 shows the actual output of the optimal ANN.

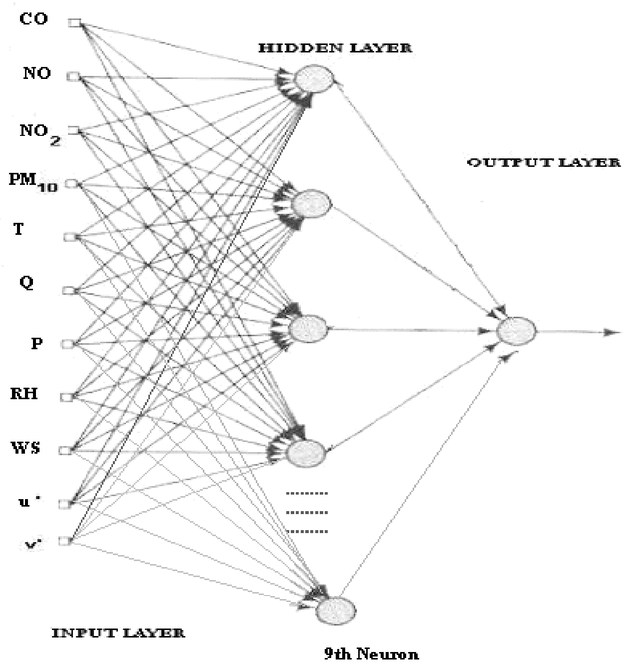


Figure 1: Architecture of the optimal ANN

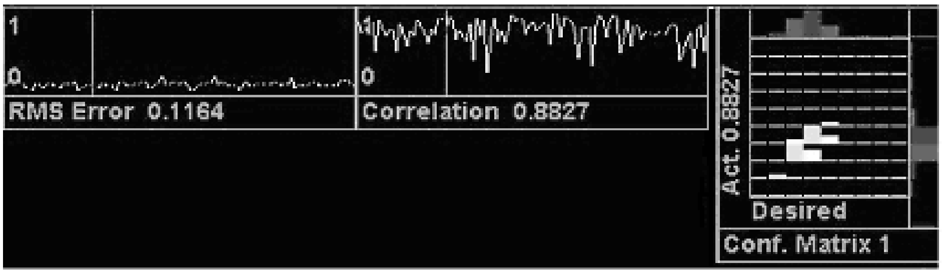


Figure 2: Correlation, RMS Error and the Confusion Matrix in Training

Figure 2 clearly presents the values of the ANN evaluation instruments for the optimal Modular ANN in the training phase. As long as the Confusion Matrix concerns the bins are located very close to the main diagonal of the matrix. The ANN also produces a diagram that estimates the degree of input contribution of each parameter.

Figure 3: Input Contribution of the independent parameters

The above figure 3, presents the input contribution of the eleven input parameters as it was estimated by the ANN. Of course the Input Contribution is produced during the testing phase by the Neuralworks II Plus software. The actual values of the columns on the Y-axis are not important. What really matters is the comparison between the contributions of the input independent parameters. It is clearly shown in figure 3 that the CO and NO play the most important role in the formulation of Ozone.

Reliability of the ANN

Training was performed for different numbers of iteration cycles. To avoid over-Training we have trained the network by performing initially 1000 iterations, increasing gradually with a step equal to 500 and the process has stopped when the performance has started dropping. Thus we have produced the optimal ANN at 5000 cycles. The good performance of the ANN with 1118 first time seen data records in the testing process, combined with the application of a very large data set of totally 4232 cases and its simple structure shows its ability to generalize and therefore its reliability.

The following figure 4, shows a comparative diagram that presents the compatibility between the actual and the ANN predicted Ozone values for 497 cases. It can easily be seen that only in a few cases some overestimations of the Ozone concentrations have been produced by the optimal ANN. However, in most of the cases the actual values are very close to the predicted ones.

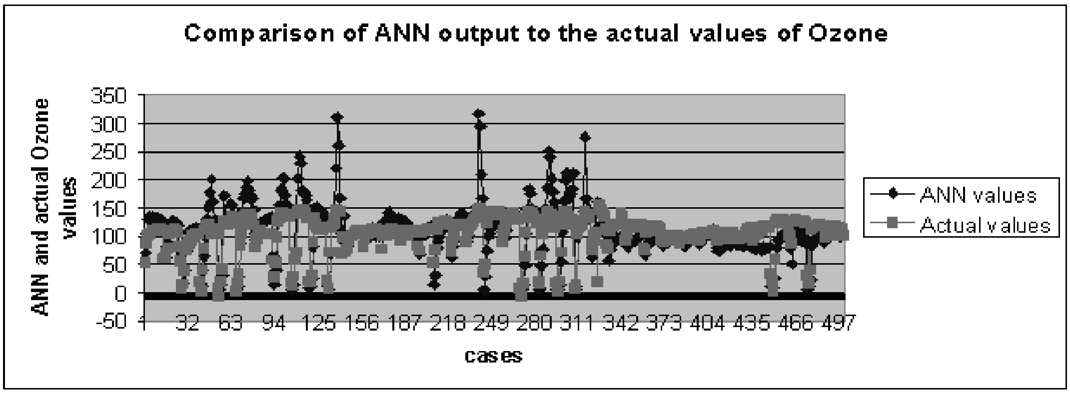


Figure 4: Comparison between the actual and the forecasted Ozone values

Comparison with regression 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, PM_{10}, 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. Even though, the coefficient of determination R^2 was found to be 0.685, a deeper look in the results revealed strong multicollinearity evidence. The low Tolerances (down to 0.258) and the high Variance Inflation Factors (up to 3.880) indicated that the predictors were strongly intercorrelated to each other, so that small changes in the data values may 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 0.672. This means that approximately 67% of the variation in O_3 values was explained by the produced equation. However, in the case of ANN the coefficient of determination R^2 was found to be 0.742. Consequently, almost 74% of the variation in ozone values has been accounted for. It is now clear that the use of ANN offers greater reliability in the issue of ozone modeling on its precursors.

Conclusions

The aim of this research effort has been the development of a reliable ANN model capable of estimating the Tropospheric Ozone concentrations in an area very close to a major city and at the same time very far from the city center. Due to the values of the testing instruments and also due to the vast volume of data used as input especially in the testing process, the generalization ability of the developed ANN model can be considered as high. From this point of view, the final product of this research is much more than a preliminary step, and it can be applied reliably towards the design of environmental protection policy and management.

The ANN can be reliably used mainly for the estimation of the ozone values in the specific area of Lykovryssi and potentially in other similar areas as well. From this point of view further research would be done towards the use of the developed Network as a prototype for Mediterranean suburbs that are not far from a major city and they have the same morphological profile with Lykovryssi. Of course this requires more time and data.

In the near future we intend to divide the area of Attica to characteristic clusters that share common properties and to model the ozone concentrations (using ANN) for the case of each cluster.

In this case, ANN have proven to be powerful tools offering a very reliable approach towards ozone concentration estimation. Compared to our previous statistical analysis, performed on the same data, most of the designed ANN have proven to perform much 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.

References

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

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

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

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

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

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

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

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

[9] Dutot A.L., Rynkiewicz J., Steiner F. E., Rude J., 2006. 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] Flaum J. B., Rao S. T., Zubenko L. G., 1996. Moderating the influence of meteorological conditions on ambient ozone concentrations, Journal of the Air and Waste Management Association, 46, pp. 35-46

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[26] Van Looy S., Meeus J., Wyns B., Vander Cruyssen B., Boullart L., Keyser F. 2005. 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 pp. 189-196

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

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

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

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