A NEW BACK-PROPAGATION ALGORITHM FOR MODELLING AIR QUALITY TIME SERIES

Giuseppe Nunnari * **Flavio Cannavó** *

* *Dipartimento di Ingegneria Elettrica, Elettronica e dei Sistemi, Universitá di Catania, Viale A.Doria 6, 95125 Catania, Italy [gnunnari,fcannavo]@diees.unict.it*

Abstract: In this paper a new Back-propagation algorithm appropriately studied for modelling air pollution time series is proposed. The underlying idea is that of modifying the error definition in order to improve the capability of the model to forecast episodes of poor air quality. In the paper five different expressions of error definition are proposed and their performances are rigorously evaluated in the framework of a real case study which refer to the modelling of 1 hour average daily maximum Ozone concentration recorded in the industrial area of Melilli (Siracusa, Italy). Results indicate that despite the traditional and the proposed version of Back-propagation performs quite similarly in terms of Success Index which gives a cumulative evaluation of the model, this latter algorithm performs better in terms of the percentage of exceedences correctly forecast.

Keywords: Stochastic modelling, Air pollution, Forecasts, Neural Networks, Backpropagation

1. INTRODUCTION

Non linear regression techniques based on Multilayer Perceptron (MLP) neural networks have drawn the attention of several scientists involved in the stochastic modelling of pollutant time series (Arena *et al.*, 1996), (Gardner and Dorling, 1998), (Kolhmainen *et al.*, 2001), (Chelani *et al.*, 2002*a*). Several authors have shown that MLP works better than traditional linear regression techniques (Finzi *et al.*, 1998), (Nunnari *et al.*, 1998), (Volta *et al.*, 1998) and also many other non-linear techniques as short-term predictors of pollutant concentrations at a point (Chelani *et al.*, 2002*b*), (Schlink *et al.*, 2003), (Dorling *et al.*, 2003). The Backpropagation algorithm (Rumelhart *et al.*, 1986), which is the basic approach to training a supervised Multilayer Perceptron (MLP) neural network, is based on the minimisation of the traditional average squared error cost function defined as follows

$$J_0 = \frac{1}{2N}\sum_{p=1}^{N} E_p = \frac{1}{2N}\sum_{p=1}^{N}(Y_p - T_p)^2 \qquad (1)$$

In (1) T_p and Y_p represent the target and actual model output value respectively. However, it is easy to understand that this assumption is not the most appropriate when dealing with pollution time series containing a relatively small number of episodes of poor air quality, as shown for instance in Table 1. This table shows the exceedances of the attention level (180 $\mu g/m^3$) of the 1 hour average of the daily maximum (DMAX) ozone concentration, recorded during 1995-1999 at the station referred to as Melilli (Siracusa, Italy), (see Fig. 1)

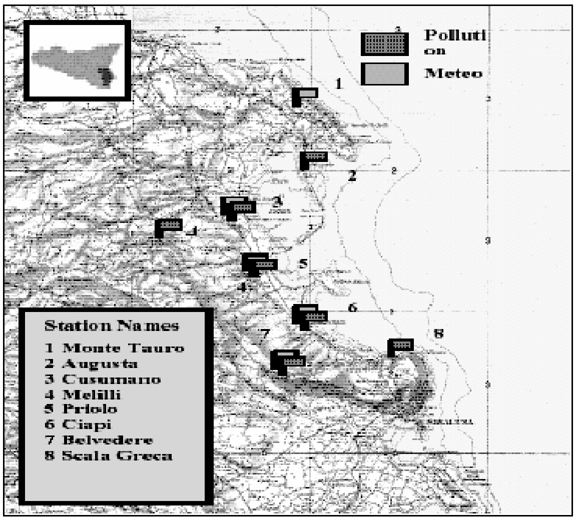

Fig. 1. The air pollution recording network in the industrial area of Siracusa

Table 1. Frequency of exceedances of the 1 hour average daily maximum Ozone concentration during 1995-1999 (attention level threshold 180 $\mu g/m^3$)

Station/Year	1995	1996	1997	1998	1999
Melilli	32	25	29	32	46

It is possible to see, for example, that the DMAX time series recorded at Melilli during 1996 presents 25 episodes exceeding the threshold of 180 $\mu g/m^3$ while the remaining 340 episodes fall below the threshold. The drawback that arises considering the cost function (1) in a similar case is due to the fact that no difference is made between targets above or below a given threshold. Hence the learning algorithm will give the 25 exceeding episodes the same weight as the remaining 340 events. The immediate consequence is that although one of the main targets of the models is the prediction of episodes of poor air quality, these events may not be relevant during the model identification process. The idea underlying this work is to modifying expression (1) in order to weight exceeding events more appropriately.

2. MODIFIED COST FUNCTIONS

In this paper five different cost functions (i.e. error definitions) are considered, as expressed in (2) to (6) respectively

$$J_1 = \frac{\sum_{p=1}^{N} (T_p - Y_p)^2 (T_p - M)^2}{2N} \tag{2}$$

$$J_2 = \frac{\sum_{p=1}^{N} (T_p - Y_p)^2 [(T_p - M)^2 + (Y_p - M)^2]}{2N} \tag{3}$$

$$J_3 = \frac{\sum_{p=1}^{N} (T_p - Y_p)^2 e^{-(\frac{Y_p}{T} - 1)(\frac{T_p}{T} - 1)}}{2N} \tag{4}$$

$$J_4 = \frac{\sum_{p=1}^{N} e^{-(Y_p - T)(T_p - T)(T_p - Y_p)^2}}{N} \tag{5}$$

$$J_5 = \begin{cases} \dfrac{\sum_{p=1}^{N} (T_p - Y_p)^2}{2N} & T_p <= T \\ \dfrac{\sum_{p=1}^{N} 2(T_p - Y_p)^2}{2N} & T_p > T \end{cases} \tag{6}$$

In expressions (2) and (3) M is a constant value and T is the threshold (e.g. 180 $\mu g/m^3$ for ozone daily maximum concentration). The reason for choosing these expressions is the following. Let us assume M to coincide with the average value of the pollutant time series, i.e.

$$M = \frac{1}{N} \sum_{p=1}^{N} T_p \tag{7}$$

In this case the term $(Tp - M)^2$ can be considered as a weight that emphasize the importance of extreme values (i.e. values other than M), such as the exceedances that the model aims to detect. The additive term $(Yp - M)^2$ in expression (3) will drive the model output towards the average values M and hence can be represented as a way to balance the effect of the term $(Tp - M)^2$. In J_3 the mean square error (MSE) is weighted according to an exponential factor. The exponential is greater than 1 when the target and the output of the network are contradictory, i.e. when the target exceedes the threshold T and the network output Y_p is below the threshold (or viceversa). In such a way the discordant cases between the real process and the model are penalized. The same mechanism is at the base of the other two cost functions. In J_4 the effect of penalizing the discordant cases is emphatized by putting the square error into the exponential function. In J_5 error definition takes into account if the values of the target is below or above the threshold. In the latter case the MSE is double weighted with respect to the former case; this gives more emphasis to episodes of poor air quality (exceedences). The properties of the considered cost functions will be illustrated below by a case study, while formulas obtained for the modified backpropagation algorithm are given in the next section. Below we will refer to the traditional back propagation, i.e. the training algorithm based on minimisation of cost function (1) as BP while we will indicate the algorithms corresponding to cost functions (2) to (6) as BP1 to BP5 respectively.

3. THE IMPROVED BACKPROPAGATION ALGORITHM

As is known the Backpropagation algorithm is a recursive algorithm to update the weights of Multilayer Perceptron (MLP) neural networks, based on the deepest-descent formula:

$$\Delta w_{ij}^{p} = -\epsilon \frac{\partial E_p}{\partial w_{ij}^{p}} = -\epsilon \delta_{i,p}^{(S)} O_{j,p}^{(S-1)} \quad (8)$$

where ϵ and w_{ij} are the learning velocity and the weight of the interconnections between the i-th neuron of the layer $S-1$ and the j-th neuron of the layer S. $\delta_{i,p}^{(S)}$ is the local gradient of the i-th neuron in the layer (S) and $O_{j,p}^{(S-1)}$ is the output of the j-th neuron in the layer $(S-1)$. The local gradient can be computed using different expressions depending on the different type of layer considered. In more detail for a neuron in the output layer $(S = n)$ when considering the traditional BP we have the expression (9)

$$\delta_{i,p}^{(n)} = (T_p - Y_p)\dot{f}(NET_{i,p}^{(n)}) \quad (9)$$

while for a neuron belonging to one hidden layer we have

$$\delta_{i,p}^{(S)} = \dot{f}(NET_{i,p}^{(S)}) \sum_{x=1}^{N_{s-1}} \delta_{r_p}^{(S-1)} w_{r_i}^{p} \quad (10)$$

In expressions (9) and (10) f is the activation function, $\dot{f}$ is the corrisponding first derivative and $NET_{i,p}^{(S)}$ and $O_{j,p}^{(S)}$ are defined as following:

$$NET_{i,p}^{(S)} = \sum_{j} w_{i,j}^{p} O_{j,p}^{S-1} \quad (11)$$

$$O_{j,p}^{(S)} = f(NET_{i,p}^{(S)}) \quad (12)$$

In view of implementing modified versions of the BP algorithm we observe that expression (10) is independent on the particular definition of the cost function. Thus adopting different definition for E_p will affect $\delta_{i,p}^{(n)}$ only, i.e. for the neuron of the output layer. We have computed $\delta_{i,p}^{(n)}$ for the five different cost functions given in (2) to (6) and the results are listed below.

Cost function J1:

$$\delta_{i,p}^{(n)} = (T_p - Y_p)(T_p - M)^2 \dot{f}(Net_i^p) \quad (13)$$

Cost function J2:

$$\delta_{i,p}^{(n)} = (T_p - Y_p)[(T_p - M)^2 + (Y_p - M)^2 - \\ -(T_p - Y_p)(Y_p - M)]\dot{f}(Net_i^p) \quad (14)$$

Cost function J3:

$$\delta_{i,p}^{(n)} = (T_p - Y_p)e^{-(\frac{Y_p}{T}-1)(\frac{T_p}{T}-1)}[2 - \\ -\frac{(T_p - Y_p)(T_p - T)}{T^2}]\dot{f}(Net_i^p) \quad (15)$$

Cost function J4:

$$\delta_{i,p}^{(n)} = e^{-(Y_p - T)(T_p - T)(T_p - Y_p)^2}(T_p - T) \\ (T_p - Y_p)[2(Y_p - T) - (T_p - Y_p)]\dot{f}(Net_i^p) \quad (16)$$

Cost function J5:

$$\delta_{i,p}^{(n)} = \begin{cases} (T_p - Y_p)\dot{f}(Net_i^p) & T_p < T \\ 2(T_p - Y_p)\dot{f}(Net_i^p) & T_p > T \end{cases} \quad (17)$$

4. MODELLING DAILY MAXIMUM OZONE CONCENTRATIONS AT MELILLI (SR)

The backpropagation algorithm proposed in this paper was considered to model 1 hour average daily maximum concentrations (DMAX) of Ozone (O3) recorded at various recording stations located in the industrial area of Siracusa (Italy). In particular we show some results concerning the station referred to as Melilli. The structure of the prediction model considered is given by the following expression:

$$O_{3_{MAX}}(t+1) = F(O_{3_{Mean[16-20]}}(t), \\ , Temp_{MAX[1-20]}(t), \\ , WS_{Mean[9-16]}(t)) \quad (18)$$

Here t represents the current day, $O_{3_{MAX}}(t+1)$ represents the daily maximum concentration of Ozone on day (t+1), $O_{3_{Mean[16-20]}}(t)$ represents the average concentration of Ozone between 4 p.m. and 8 p.m. on day t, $Temp_{MAX[1-20]}(t)$ represents the max temperature on day t computed between 1 a.m. and 8 p.m., $WS_{Mean[9-16]}(t)$ represents the mean wind speed computed between 9 a.m. and 4 p.m. at time t, and, finally, F is a non linear unknown function to be approximated. The prediction model structure shown in (18) was obtained by using a trial and error approach. Fig. 2 shows the one hour average daily maximum concentration of Ozone recorded at the Melilli station during the time period (1995-1998). Model identification was performed considering recorded data deprived of 365 consecutive values (a one year set of data) that were used as test set. To compare the different algorithms 10 trials have been carried out changing the test set. The process being modelled was assumed to be stationary during the time interval considered.

4.1 Performance Indices

In order to evaluate the capabilities of the training algorithms to predict exceedances of the attention level the indices defined in (19)-(25) were computed.

$$SP = \frac{N_p}{N_o} \quad (19)$$

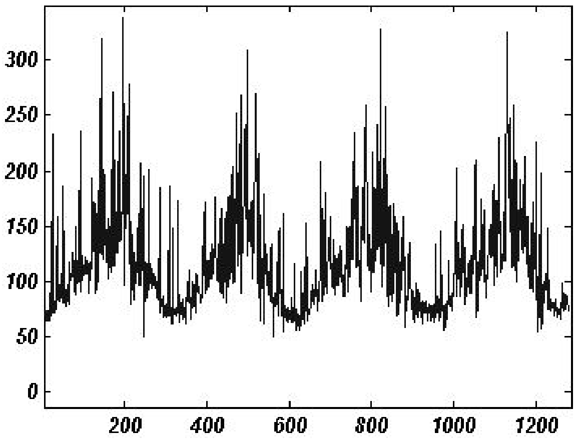

Fig. 2. 1-hour average daily maximum concentration of Ozone recorded during 1995-1998

$$SR = \frac{N_p}{N_F} \quad (20)$$

$$FA = 1 - SR \quad (21)$$

$$SI = (\frac{N_p}{N_o} + \frac{N + N_p - N_o - N_F}{N - N_o} - 1) \quad (22)$$

In expressions (19)-(25) N_o is the total number of observed exceedances of a given threshold, N_p is the number of correctly predicted exceedances, N_f is the total number of forecast exceedances and N the total number of data points. The meaning of the indices defined above is the following. SP indicates the percentage of exceedances correctly forecast, FA is the percentage of false alarms, SR gives the percentage of predicted exceedances which actually occurred and, finally, SI is the success index which gives a cumulative evaluation of how well the exceedances are predicted. Details about the measuring of the mentionated performance indices can be found in (Aalst and Leeuw, 1997).

$$PI = (1 - \frac{N_o + N_f - 2N_p}{N}) \quad (23)$$

$$PI = P(O, Y) + P(\overline{O}, \overline{Y}) \quad (24)$$

Unfortunately the SI index does not express a probability of success in strictly probabilistic sense. To overcome this drawback we propose here a new index referred to as PI (Probability Index) expressed by (23). It is easy to demonstrate that PI can also represented as indicated by (24). The right term of expression (24) represents the sum of two probabilities: $P(O, Y)$ which gives the probability that an observed exceedance will be correctly predicted by the model and $P(\overline{O}, \overline{Y})$ which represents the probability that non exceeding values will also correctly forecast. In (24) the argument O represents a boolean variable defined as following:

$$O = \begin{cases} True & \text{when the pollutant time series to be modelled exibits an exceedance (e.g. } O_{3_{MAX}} > T) \\ False & \text{otherwise} \end{cases}$$

$$\overline{O} = notO$$

The arguments Y and $\overline{Y}$ have the same meaning of O and $\overline{O}$ but refer to the estimated values (i.e. the output of the prediction model). Furthermore in this paper we introduce another new index, referred to as GI (Global Index) expressed by (25) which gives a measure of the success of the forecasting model independent on the number N of samples in the modelled time series.

$$GI = (\frac{N_p}{N_o + N_f - N_p}) \quad (25)$$

It is to be stressed that both PI and GI assume values in the $[0, 1]$ interval. For a good prediction model PI and GI should approach to 1.

4.2 Experimental Framework

To evaluate the peculiarities of adopting the modified cost functions (2) to (6) in comparison with the traditional MSE given in (1), a software tool was coded which implements the modified back propagation algorithms as described in the previous section. All these algorithms were considered to train the NARX model given in (18). In order to obtain a measure of the generalization capabilities not affected by a particular training and testing set, the learning phase was organized as follows. The available data set spanning for 1995 to 1998 was divided in ten overlapping data sets, each containing one year data (i.e. 365 samples of daily maximum ozone concentration). For each back-propagation algorithm ten different trials were performed. During each trial 9 of the 10 data set were considered for the training and the remaining one for the test. This should guarantee a non biased evaluation of the performance (i.e. the set of indices is representative of the generalization capabilities of the neural model). During all the experiments the number of learning cycles, hidden neurons (in the unique hidden layer considered) and the learning velocity were considered constant in order to assure a more objective inter-comparison exercise. In particular the number of hidden neurons was set to 6, the learning velocity ϵ to 0.1 and the number of learning cycles was set to 10000. Results are reported for the six different algorithms in Table 2 to 7. Each table gives the performance indices for the ten trials carried out by using a given algorithm. Furthermore the averaged values of the performance indices are summarized in Fig. 3 and 4. In particular Fig. 3

gives the SP, FA and SI indices and Fig. 4 the introduced set of indices (PI and GI). From Fig. 3

Table 2. Performance of models obtained using the traditional BP (cost function J0)

$SP\%$	$SR\%$	$FA\%$	$SI\%$	$PI\%$	$GI\%$	N_o
47.8	63.5	36.5	40.2	82.8	37.5	69
67.5	78.3	21.7	61.2	87.1	56.8	80
70.8	78.0	22.0	65.7	90.0	59.0	65
67.2	77.6	22.4	62.0	89.0	56.3	67
70.8	67.1	32.9	60.7	85.6	52.6	2
52.8	63.3	36.7	43.9	82.4	40.4	72
55.1	64.4	35.6	46.7	83.7	42.2	69
54.7	66.0	34.0	47.6	85.3	42.7	64
57.1	69.0	31.0	49.9	85.0	45.5	70
62.3	67.6	32.4	52.8	83.7	48.0	77

Table 3. Performance of models obtained by using the BP1 (cost function J1)

$SP\%$	$SR\%$	$FA\%$	$SI\%$	$PI\%$	$GI\%$	N_o
85.5	38.8	61.2	48.3	67.7	36.4	69
95.0	46.3	53.7	58.2	71.2	45.2	80
90.8	42.4	57.6	59.3	73.0	40.7	65
88.1	43.7	56.3	57.9	73.7	41.3	67
88.9	44.1	55.9	56.1	72.1	41.8	72
86.1	40.8	59.2	49.7	68.7	38.3	72
89.9	40.0	60.0	52.7	68.7	38.3	69
96.9	40.3	59.7	60.8	70.5	39.7	64
92.9	43.3	56.7	58.7	71.8	41.9	70
92.2	44.4	55.6	55.4	70.2	42.8	77

Table 4. Performance of models obtained by using the BP2 (cost function J2)

$SP\%$	$SR\%$	$FA\%$	$SI\%$	$PI\%$	$GI\%$	N_o
68.1	53.4	46.6	51.7	80.3	42.7	69
82.5	59.5	40.5	63.7	81.5	52.8	80
83.1	54.0	46.0	65.0	82.1	48.6	65
79.1	55.2	44.8	62.0	82.1	48.2	67
84.7	55.5	44.5	64.9	81.2	50.4	72
75.0	52.9	47.1	55.6	79.3	45.0	72
75.4	52.5	47.5	56.6	79.9	44.8	69
79.7	56.0	44.0	64.0	83.4	49.0	64
77.1	60.0	40.0	62.7	83.7	50.9	70
77.9	57.7	42.3	59.7	80.9	49.6	77

it appears that all the modified back-propagation algorithms (except BP3) perform better than the traditional BP in terms of SP and SI. In particular SP is about 0.60 for BP, 0.90 for BP1, 0.78 for BP2, 0.84 for BP4 and 0.70 for BP5. However this result is accompanied by a larger number of false alarms. This agrees with the fact that the PI and the GI are almost constant for all the considered algorithms. In other words the proposed back-propagation algorithms do not perform globally better than the traditional BP but if the modeller

Table 5. Performance of models obtained by using the BP3 (cost function J3)

$SP\%$	$SR\%$	$FA\%$	$SI\%$	$PI\%$	$GI\%$	N_o
44.9	62.8	38.0	37.3	82.1	35.2	69
67.5	79.4	20.6	61.6	87.5	57.4	80
66.2	84.3	15.7	63.0	90.6	58.9	65
62.7	85.7	14.3	59.9	90.0	56.8	67
66.7	73.8	26.2	59.8	87.1	53.9	72
50.0	67.9	32.1	43.1	83.4	40.4	72
52.2	67.9	32.1	45.4	84.3	41.9	69
48.4	73.8	26.2	44.1	86.2	41.3	64
57.1	78.4	21.6	52.7	87.1	49.4	70
54.5	75.0	25.0	48.8	84.6	46.2	77

Table 6. Performance of models obtained by using the BP4 (cost function J4)

$SP\%$	$SR\%$	$FA\%$	$SI\%$	$PI\%$	$GI\%$	N_o
71.0	50.5	49.5	51.8	78.7	41.9	69
87.5	58.3	41.7	66.6	81.2	53.8	80
84.6	52.9	47.1	65.3	81.5	48.2	65
80.6	53.4	47.6	61.2	80.6	46.6	67
84.7	55.0	45.0	64.5	80.9	50.0	72
86.1	53.7	56.3	53.7	71.8	40.8	72
82.6	50.0	50.0	59.8	78.4	45.2	69
95.3	43.3	56.7	63.9	74.0	42.4	64
84.3	58.4	41.6	67.4	83.4	52.7	70
85.7	55.5	44.5	63.8	79.9	50.8	77

Table 7. Performance of models obtained by using the BP5 (cost function J5)

$SP\%$	$SR\%$	$FA\%$	$SI\%$	$PI\%$	$GI\%$	N_o
59.4	57.7	42.3	47.4	81.8	41.4	69
71.3	67.1	32.9	59.5	84.0	52.8	80
73.8	61.5	38.5	62.0	85.3	50.5	65
70.1	61.8	38.2	58.6	84.6	49.0	67
76.4	58.5	41.5	60.6	82.4	49.5	72
65.3	52.8	47.2	48.3	79.0	41.2	72
69.6	53.3	46.7	52.8	80.3	43.2	69
67.2	58.1	41.9	55.0	83.7	45.3	64
72.9	60.7	39.3	59.6	83.7	49.5	70
75.3	61.7	38.3	60.4	82.8	51.3	77

is interested in maximizing the performance in terms of percentage of exceedances correctly forecasted it is quite evident that a benefit can be obtained adopting one of the introduced algorithms. The price to pay is an increasing level of false alarms which is usually acceptable provided that it is lower than a prefixed threshold (say 0.40). Fig. 3 shows that BP5, among the inter-compared algorithms, is the best compromise between a high level of SP (0.70) and an acceptable level of FA (0.40) whilst the traditional BP exhibits SP = 0.60 and FA = 0.30. It is interesting to stress here that the choice of $BP5$ is also confirmed by the following reasoning carried out in terms of PI

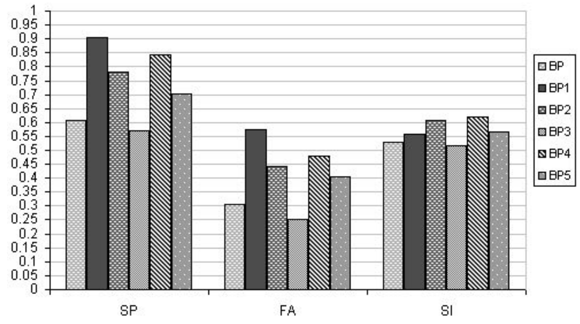

Fig. 3. Comparison among all proposed algorithms in terms of SP, FA, SI.

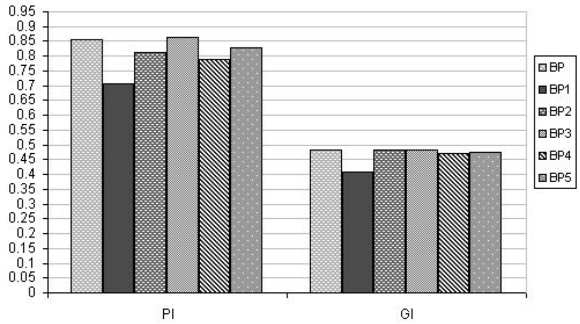

Fig. 4. Comparison among all proposed algorithms in terms of PI and GI.

and SP. Indeed Fig. 4 shows that the best three models in terms of PI are BP ($PI = 0.854$), $BP3$ ($PI = 0.863$) and $BP5$ ($PI = 0.827$) since they exhibit almost the same value. However $BP5$ is the best with respect BP and $BP3$ in terms of SP . Hence we may suggest this criterion to make the choice among various air quality prediction models.

5. CONCLUSIONS

In this paper a novel backpropagation algorithm to improve the capabilities of the traditional backpropagation algorithm to predict episodes of poor air quality has been proposed. The rigorous intercomparison, performed in the framework of the described case study show that despite the traditional and the proposed algorithms perform quite similarly in terms of success index and global index, this latter algorithms performs better in terms of the percentage of exceedences correctly forecast. The price to pay for this is a limited increasing in the percentage of false alarms.

REFERENCES

Aalst, R. M. Van and F. A. A. M. De Leeuw (1997). National ozone forecasting system and international data exchange in northwest europe. *European Topic Centre on Air Quality.*

Arena, P., S. Baglio, C. Castorina, L. Fortuna and G. Nunnari (1996). A neural architecture to predict pollution in industrial areas. In: *International Conference on Neural Network - ICNN*. Vol. 4. Washington, USA. pp. 2107–2112.

Chelani, A.B., D.G. Gajghate and M.Z. Hasan (2002*a*). Prediction of ambient pm10 and toxic metals using artificial neural networks. *Journal of Air and Waste Management Association* **52**, 805–810.

Chelani, A.B., R.C.V. Chalapati, K.M. Phadke and M.Z. Hasan (2002*b*). Prediction of sulphure dioxide concentration using artificial neural network. *Environmental Modelling and Software* **17**, 161–168.

Dorling, S.R., R.J. Foxall, P.D. Mandic and G.C. Cawley (2003). Maximum likelihood cost functions for neural networks models of air quality data. *Atmospheric Environment* **37**, 3435–3443.

Finzi, G., M. Volta, A. Nucifora and G. Nunnari (1998). Real-time ozone episode forecast: a comparison between neural network and grey-box models. In: *International ICSC/IFAC Symposium on Neural Computation*. Vienna.

Gardner, M.W. and S.R. Dorling (1998). Artificial neural networks (the multilayer perceptron), a review of applications in the atmospheric sciences. *Atmospheric Environment* **32**(14), 2627–2636.

Kolhmainen, M., H. Martikainen and J. Ruuskanen (2001). Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment* **35**, 815–825.

Nunnari, G., A. Nucifora and C. Randieri (1998). The application of neural techniques to the modelling of time-series of atmospheric pollution data. *Ecological Modelling* **111**, 187–205.

Rumelhart, D.E., G.E. Hinton and L. McLelland (1986). *Parallel Distributed Processing Exploration in the Microstructure of Cognition.* Vol. 1. MIT Press. Cambridge.

Schlink, U., S. Dorling, E. Pelikan, G. Nunnari, G. Cawley, H. Junninen, A. Greig, R. Foxall, K. Eben, T. Chatterton, J. Vondrcek, M. Richter, M. Dostal, L. Bertucco, M. Kolehmainen and M. Doyle (2003). A rigorous inter-comparison of ground-level ozone predictions. *Atmospheric Environment* **37**, 3237–3253.

Volta, A., A. Nucifora and G. Nunnari (1998). Real time ozone episode forecast: a comparison between neural network and grey box models. In: *International ICSC/IFAC Symposium on NEURAL COMPUTATION-NC'98, ICSC*. pp. 854–860.