Use of Neural Net Models to Forecast Ozone pollutant in Rome Urban Area.

A. Pelliccioni (*), F. Pungì (*), A. Ianni (**), L. Bellina (**)
(*) ISPESL-DIPIA, Via Fontana Candida 1, 00040 Monteporzio Catone (RM), Italy.
(**) University of Rome "La Sapienza"- DIS Via Ariosto 25 00185 (RM), Italy.

Abstract

The aim of the present work is to provide a methodological procedure to forecast Ozone concentrations 24h in advance. The study area is the urban center of Rome, and the results for two measurements locations are presented here. Artificial neural network models have been developed using an ensemble of meteorological and pollutants variables as input.
The Ozone levels have been predicted using two neural network architectures: the first one to forecast the meteorological variables 24h in advance and the second one for the Ozone using as inputs the meteorological data coming from previous net and the pollutants, taken one and two days before.

1. INTRODUCTION

Air quality problems produced by high levels of ozone (O_3), that is a well known air pollutant, concern for their effects on human health related with respiratory problems.
The latter are critical especially in large metropolitan areas where transport emissions are relevant causing a greater exposure of population with consequent health problems.
Ozone is a very reactive gas and presents concentration levels which are strongly dependent both from the micro-meteorological conditions of the site and the seasonal effects. The prediction of Ozone levels is very complex to obtain as described in different studies [1][2][3][4][5].
For Ozone models one of the most difficult problems to deal with is the simulation of the chemical reactions that occur in atmosphere, linked to the long range transport, to the incoming solar radiation and to the atmospheric turbulence conditions [6][7].
NN methods have been developed to forecast daily maximum ozone levels in various urban areas, using average daily meteorological data as input parameters [2].
Among the complex systems, an important tool to forecast air pollution data, by advanced statistical methods, is the neural network (NN) that can work as universal approximator of non-linear functions and, consequently, can be used in assessing the dynamics of such systems [8].
In our work, NN methods have been developed to forecast daily maximum ozone levels, the 8h average and the hourly levels, using daily and hourly meteorological and concentration data as input parameters. The aim of our work is to provide a methodological

procedure in order to forecast Ozone 24 hours in advance.
During our simulations, we applied both a cascade of Neural net models and a process of data optimization, which aim is to select patterns and variables having an high meaning for explaining data variability related to NN model performance.
For the simulations, the optimization of the input patterns is a critical point. In fact, the presence of extreme events (Ozone levels higher than 80 $\mu g/m^3$) has interest for the related human health, but is less significant from a statistical point of view, constituting about 1% of the collected data.
The NN parameters have been obtained by a training procedure based on use of an efficient unconstrained minimization algorithm.
The training procedure target is not only to reproduce Ozone trend, but it is also to try to simulate the process of Ozone diffusion and the chemical reactions in atmosphere. The meteorological conditions play an important role during the summer season, when relevant photochemical peaks of pollutant can be observed and give rise to health effects on children and elderly.
The goodness of the Ozone forecast is strongly dependent by selection of the patterns during the training phase, and the results are related with the statistical distribution of the Input/Output data set.
Our research shows that a preliminary study of input data is always needed in order to remove not very meaningful data from the training sets, to choice a suitable normalization rule and to compute the usual statistical indexes of correlations among the variables. The skill of the NN to capture the environmental information inside the data is highly dependent by the preliminary study of patterns. The generalization capacity of the net to forecast Ozone peaks has to be connected with the essential information inside the data set and this information is not necessarily regularly distributed among all patterns.

2 DATASET DESCRIPTION

Data used in our simulations come from two monitoring stations of the ARPA LAZIO (Regional Agency for Environmental Protection in Lazio) network in the urban center of Rome (Magna Grecia and Corso Francia), and regard hourly data during all the year 2005.
The variables used for the simulations are:

1. observed pollutants variables:
   - Carbon monoxide ($\mu g/m^3$) – CO
   - Nitrogen oxide ($\mu g/m^3$) – NO
   - Nitrogen dioxide ($\mu g/m^3$) – NO_2
   - Ozone ($\mu g/m^3$) – O_3 – (Input/Output variable)

2. meteorological variables:
   - Temperature (C°) – T
   - Global Solar Radiation (W/m^2) – GSR
   - Relative Humidity (%) – RH
   - Pressure (mbar) – Press

Data coming from Magna Grecia, consisting of 7897 hourly patterns, are used to train the NNs whereas those coming from Corso Francia, consisting of 7202 hourly patterns, are used to test the performance of the NNs.

The Magna Grecia monitoring stations represents typical urban situations (high traffic sources and domestic heating during winter season) and is located in the central area of Rome. The Corso Francia station is similar to Magna Grecia, but is located outside respect the Rome central area and traffic sources are representative of speedier traffic flows.

3 METHODOLOGY

We used in our simulations two NN architectures to forecast meteorology at first, and Ozone finally. For both architectures we used a 3-Layer Perceptron model, which is considered to be able to approximate every measurable function. The first input layer contains the input variables of the net linked with all relevant physical parameters.

The second layer consists of the neurons of hidden layer. The third layer is the output layer, which consists of the target variable to be reproduced (e.g. pollutants concentration or atmospheric variables).

We used a cascade of MLP models, consisting of two separated architectures in which the outputs of the first one (the meteorological NN architecture) are the input for the second one (which aim is to forecast the ozone one day before).

3.1 NN architecture to forecast Meteorology

As meteorological variables, in our analysis we considered the temperature (T), the relative humidity (RH), the global solar radiation (GSR) and the pressure (Press). We didn't take into account the wind direction and speed for two reasons. Usually, the monitoring stations give wind at 2 meters height and no correlations with pollutant transport can be associated to that measurement. Secondly, the more important consideration, is that ozone is typically a secondary pollutant so it is influenced more by air parcel conditions (linked to scalar meteorological variables) than by transport factors.

Our strategy to simulate the ozone 24h in advance, consists of two steps equivalent to a cascade of NN models.

We start from the consideration that actual ozone levels depends by urban

micrometeorological conditions and by the levels of primary pollutants. We separate the task of NNs in order to optimize the reproduction of the urban micrometeorological environment and the available chemical conditions.
We operate with two separated NN architectures.
Aim of first NN architecture is to simulate 48-24h before, those meteorological conditions related to urban micro climate.
The first NN, used to forecast meteorology take 8 input data (T, RH, GSR and Press) at different time lags: $T(L_{48-24})$, $RH(L_{48-24})$, $GSR(L_{48-24})$, $Press(L_{48-24})$.
The outputs of this NN are the same meteorological variables at time lag zero: $T(L_0)$, $RH(L_0)$, $GSR(L_0)$, $Press(L_0)$ (Fig. 1).

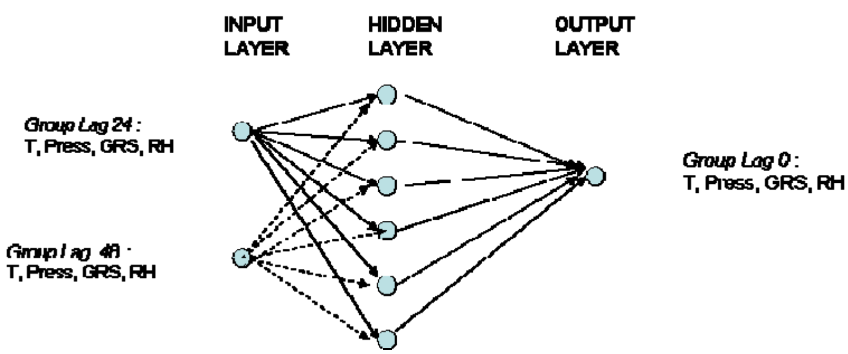

Fig. 1: MLP architecture for meteorological net

Behind the forecasting of urban micrometeorological variables, further interesting results can be investigated by this particular NN configuration. The input/output of meteorological NN are self-consistent variables. This means that we can evaluate the performance of our net in forecasting the urban micrometeorology at different days in advance. If we consider as NN meteorological input the results obtained by the same NN at different days, we can simulate the variables for 1-2-3-4 and 5 days in advance (denoted Lag 24, Lag 48, Lag 72, Lag 96 and Lag 120 respectively). It is obvious that 5 days in advance simulations have no means, but the results are interesting and a similar strategy could be used to improve meteorological forecast with the assistance of NN algorithm.
As last considerations, we organize our results in such a way that they are coherent and self- consistent (use all available data given by ground monitoring stations).
This step, linked to the forecast of urban micrometeorology one day before, can be substituted by the meteorological forecasting that meteorological office agency offers. In such a way, the training of next MLP would utilize as input data the forecasted meteorological parameters coming from the agency some day in advance.

3.2 NN architecture to forecast Ozone

The previous net has the task of predicting the micrometeorological variables one day before in which the ozone can react or disperse by the turbulence conditions.

The aim of this step, is to estimate the ozone levels one day before using both the estimate of micrometeorological conditions and the estimate of the pollutants (calculated by levels at one-two days before the target).

This second NN used to forecast Ozone levels, takes as input data those coming from the previous NN as well as the observed pollutants at different time lags: $CO(L_{48\text{-}24})$, $NO(L_{48\text{-}24})$, $NO_2(L_{48\text{-}24})$, $O_3(L_{48\text{-}24})$. The scheme of architecture used is shown in Fig. 2.

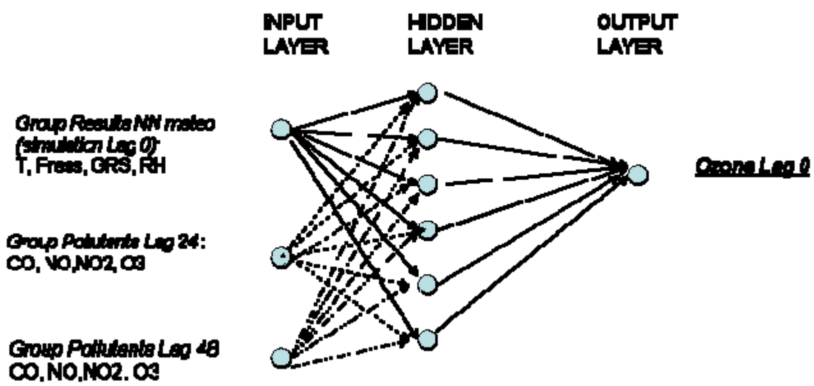

Fig. 2: ***MLP architecture for Ozone forecast***

We can observe that all the pollutants, considered as NN input, are important for the determination of the ozone levels. CO and NO coming mainly from traffic sources and are primaries fingerprint of combustion exhaust process. NO_2 can be considered as the main secondary tracer of combustion process and is linked to the NO levels. The direct use of Ozone helps to estimate the ozone one or two days before. We tested other NNs without using Ozone and we tried that results are not so good as with the presence of ozone itself.

Another consideration has to be done. We could use only data related to Lag 24 (one day before). We preferred to extend the simulations up to Lag 48 (two days before) to give the trend both of the pollutants and meteorological situations.

Data Lag 24 provide to simulate initial conditions, while data Lag 48 supply for the trend both of meteorological and sources contributions.

3.3 Main NN parameter

The MLP can present different performances in according to the choice of activation function and to the number of hidden neurons. As activation function, we use the standardized sigmoid one. A different choice of the activation function could improve the NN performance but, given the complexity of our task, we focus basically on patterns and on variables of the net rather than on the neural network algorithm optimization.
We trained the net with 5, 10, 15, 20 neurons for the hidden layer, and finally we chose 15 that gives the best performance in terms of minimizing the error function; so the final architecture is a 8x15x4 Multi Layer Perceptron for the meteorological NN and 12x15x1 for the hourly ozone simulation 24 hours in advance.
The results related to predictions 8h average and to the maximum daily are calculated directly by the hourly simulation.

4 RESULTS AND DISCUSSION

4.1 Validation and test results for hourly ozone

According to our planning, the results are referred to Magna Grecia station during the training phase and to Corso Francia during the test one.
The results show, using as input data 48h-24h past measurements of primary pollutants and meteorological variables predicted at time lag zero, a correlation coefficient for the Ozone of 0.84 during the training and 0.81 for the Corso Francia monitoring station. The generalization is excellent (R_{train}/R_{test}=0.97), taking into consideration that Corso Francia station has about the same amount of data respect to Magna Grecia and that no pattern is used during the training.
The trend of measured ozone and the ozone reproduced by the NN models cascade are shown in Fig. 3 and 4.
While the trend of the model (Fig 3) is satisfactory, the study of scatter plot (Fig 4) suggests that could be necessary some improvement to predict high and low Ozone levels. This behavior is mainly related to the saturation of the activation function and could be due to complexity of relation between the input variables and the Ozone (taking into consideration the NN cascade model).
When we consider trend and scatter plot of ozone for Corso Francia (test case), we obtain the results as in Fig. 5 and 6 respectively.

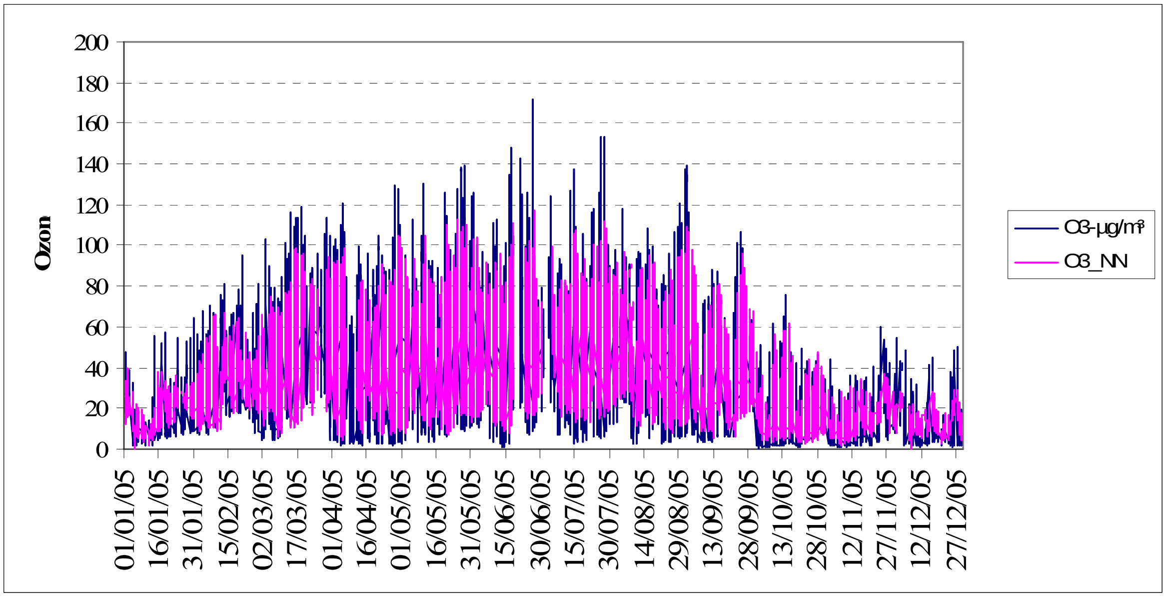

Fig. 3:Trend of Ozone as reproducing by Magna Grecia monitoring station (Train Case)

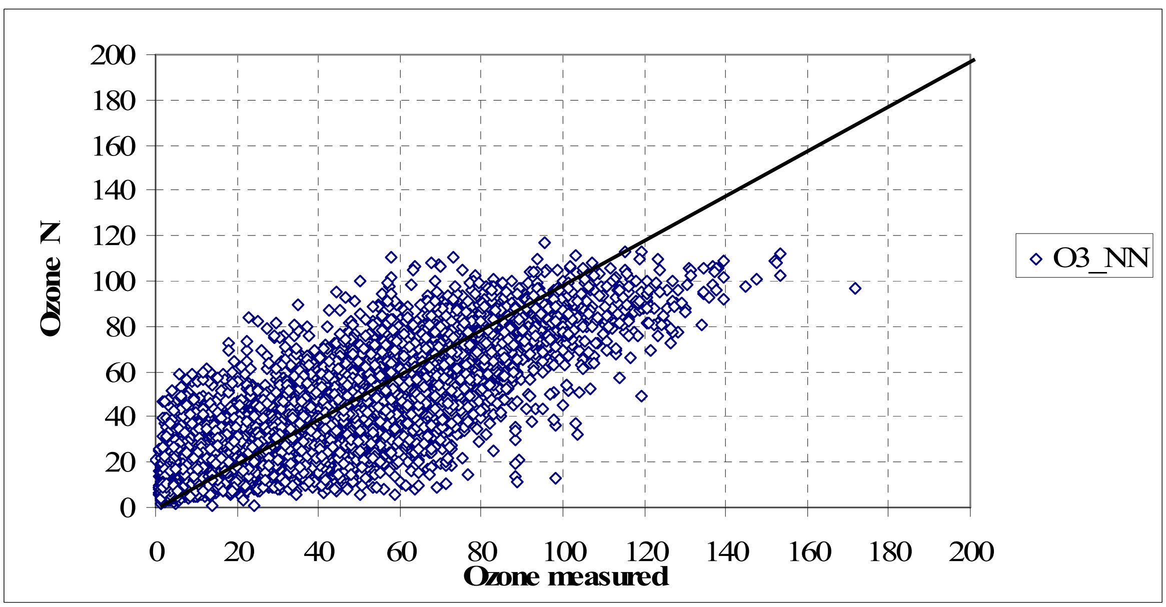

Fig. 4: Magna Grecia monitoring station: Ozone by NN vs Ozone measured

The Corso Francia station presents higher ozone levels respect to Magna Grecia one (in according to the suburban position) and the good reproducibility, as evident by the trend and by the scatter plot, means that our modeling strategy is winner. By Fig 6, we observe that peaks of ozone greater than 120µg/m^3 are difficult to simulate one day before at the suburban station of Corso Francia, starting by the training on urban station of Magna Grecia. This results is very interesting because the ozone background for the two station could be different (with high values for the suburban station respect to the urban one) and suggests that further information could be necessary to cover all Rome area.

However, our simulations demonstrate the following:

- It is possible to simulate ozone levels one day before
- By appropriate strategies, the Neural Network can be used to reproduce data spatial correlation. This is a novelty on the pollutants prediction by NN, because, up to now, only the temporal trend was the target of the net. This work open the way to the spatial-temporal reproduction by neural networks.
- Training the net using all data picked up from a monitoring station, while take into account the seasonal effects as the emissions sources, can neglect some effects strictly associated to the local site. In this case, it could be necessary a deeper study for the choice of the reference station for the training. The two stations chosen are classified by same typologies by ARPA Lazio network. The results show that, for the ozone, the stations probably are not similar and further classification based on the background from secondary pollutants could be necessary .

If we use the conventional parameters [9] to compare the results between training and test phase, we obtain the following (Tab.1):

Tab. 1: Main performance index for training and test phase

	Train (*Magna Grecia*)	**Test (*Corso Francia*)**
NMSE	0,22	0,31
R	0,84	0,81
FA2	73,10	63,50
FB	0,00	-0,07
FD	0,18	0,12

All indexes, except the NMSE, are good both for training and test phase. The NMSE behavior can be affected by the underestimation of ozone peaks due to the improvement of the optimization of the net and to the local ozone background.

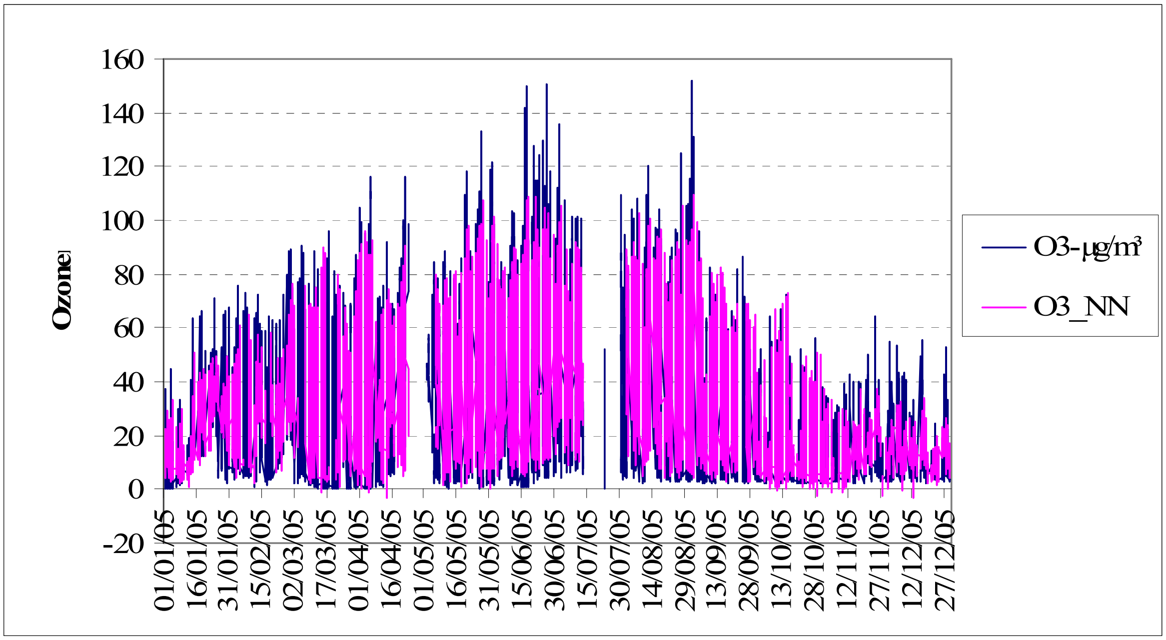

Fig. 5: Trend of Ozone as reproducing by Corso Francia monitoring station (Test Case)

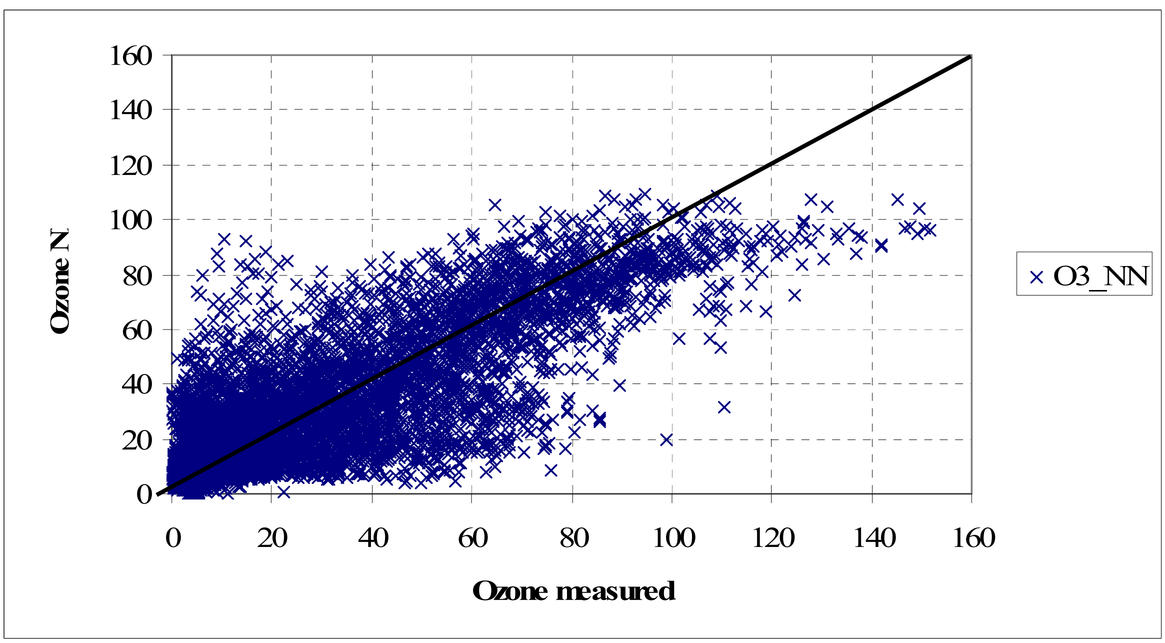

Fig. 6:Corso Francia monitoring station -Ozone by NN vs Ozone measured

4.2 Results for legislative Ozone levels

The normative levels for ozone health effects (Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe) refer (*120* **µg/m³** maximum daily eight-hour mean within a calendar year.

The Italian reference levels are:

	Hour (information)	**8h average**	**Maximum Daily**
O3 (µg/m³)	*180* **µg/m³**	*120* **µg/m³**	*200* **µg/m³**

The 8h average is referred to working hours and is associated with exposure levels. This average gives an idea of the low dose effects on human health.
The maximum daily ozone is related to the acute effects on health and cannot exceed *200* **µg/m³** time for year.
Trend of 8h ozone average is shown in Fig 7 for the test case. As evident, the prediction of the 8h average one day before is excellent, with a correlation index of 0.77 without bias, and 0.83 with a bias of 12.8 µg/m³.
It has to be underlined that results are representative of all the year and that Corso Francia station is suburban area distant from Magna Grecia station. Obviously, results for Magna Grecia present good correlation, higher respect to test case, of 0.84 without bias and 0.87 with a bias of 8.45µg/m³.

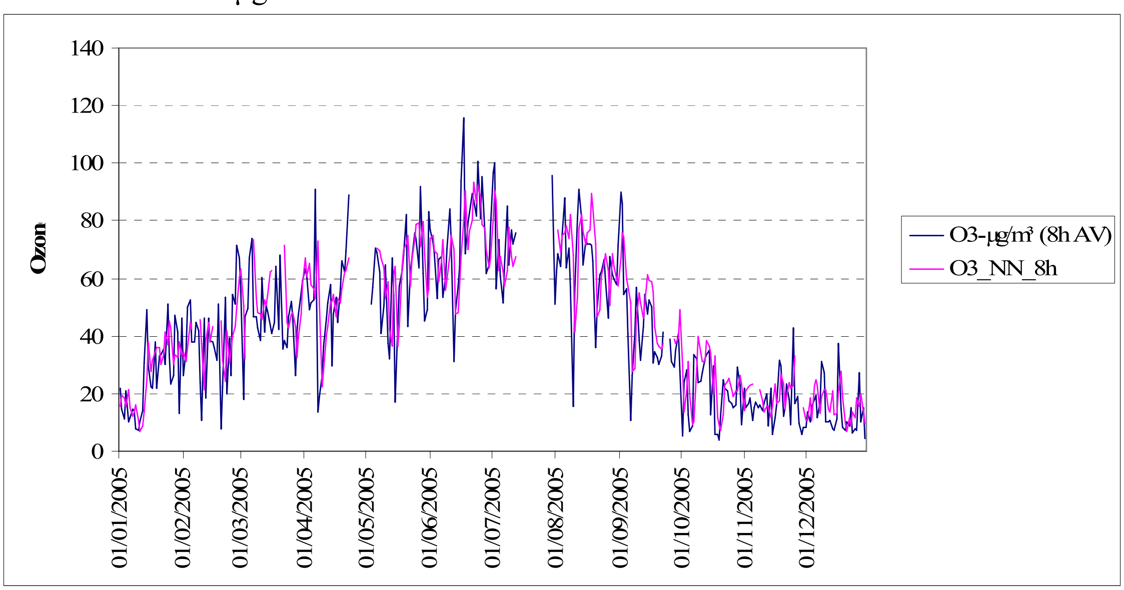

Fig. 7: Trend of 8h Ozone as reproducing by Corso Francia monitoring station (Test Case)

The similar performances were obtained for the maximum ozone levels.
In Fig. 8 are given the comparison between maximum ozone predicted by our NN cascade models and the observed ones for test dataset. Also in this case, the results are excellent, showing a correlation of 0.79 with a bias of 12.47 μg/m^3.

All results regarding forecasting of ozone one day before at different resolutions (hourly, 8h average and maximum daily), show very good performances, proving that our NN strategy is efficient to simulate the dispersion as the chemical behavior of this complex pollutant. The hourly simulation presents worse results. It is worth noting that is very difficult, in general, predict ozone one day before at same time (given that its levels is determined also by instantaneous chemical reactions). Easier is the forecast of maximum daily ozone, given that this could be happened during the day (no time information is request).

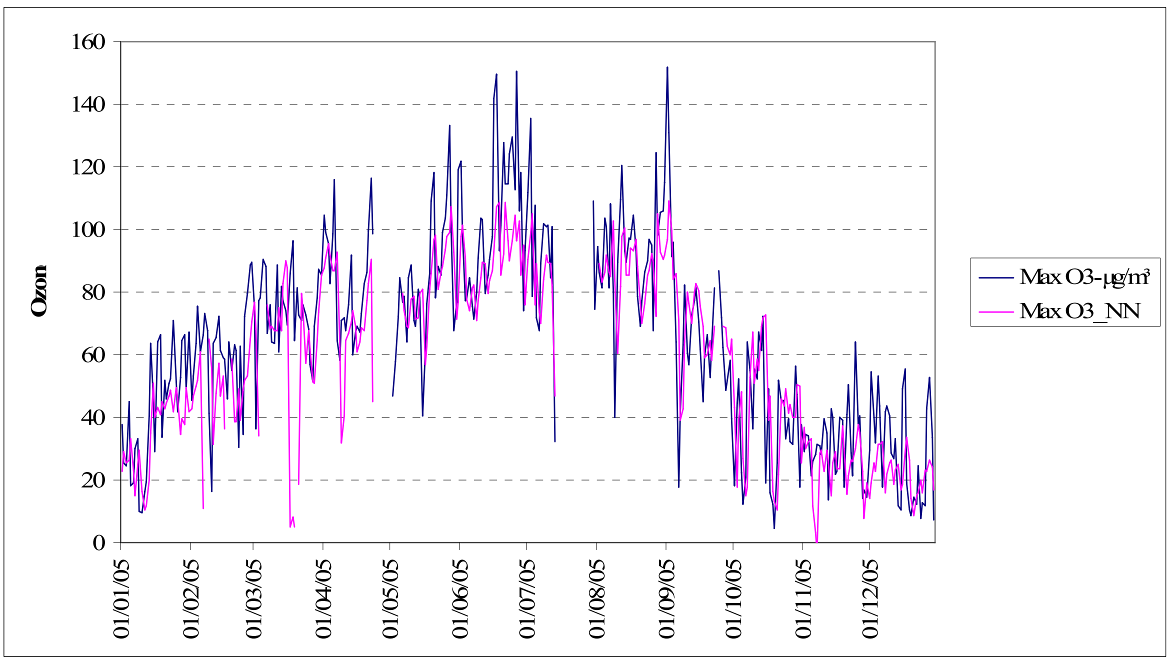

Fig. 8*: Trend of Maximum daily Ozone as reproducing by Corso Francia monitoring station (Test Case)*

4.3 Results concerning meteorological variables

A very attractive result derives from meteorological neural net. The variables choice is made so that is possible to provide as input for the net, the results coming from a previous net. So doing we can test the capacity of our net to forecast meteorological variables using

data one day before (Lag 24), two-three-four-five days before (respectively Lag 48-72-96-120).
Taking into consideration the meteorological variables forecasted in Magna Grecia and Corso Francia stations, we calculate the correlation decay trend at different temporal lags (Tab. 2).
The more difficult variable to predicted one day before is the relative humidity (R=0.73 and 0.78), while the easier is the temperature (R=0.96 and 0.96). Also the pressure and the solar radiation are easy to simulate one day in advance.
If we consider the correlation at different temporal Lags, we find the very interesting results that our net forecasts the temperature up to five days in advance in a very good way, while the more critical parameters to simulate in advance is the pressure, which presents a rapid decay only after three days. The solar radiation is the second best parameter that preserves information for different Lags, maintaining average correlations of 0.72 up to the third day.
This approach to simulate meteorological variables by NN could be very promising, because could be utilized in order to simulate micrometeorological variables in advance in complex situations as in the urban case. All meteorological models that attempt to forecast micro environmental parameters fall after few days. The Neural network can be considered a valid alternative or an help for physics models to improve their performance in forecasting.

Tab. 2: Correlation at different temporal Lags for the test and the train cases.

Magna Grecia	**T**	**RH**	**Press**	**GSR**
LAG(24)	0,97	0,73	0,85	0,93
LAG(48)	0,94	0,60	0,64	0,83
LAG(72)	0,92	0,53	0,48	0,77
LAG(96)	0,82	0,13	0,36	0,14
LAG(120)	0,74	0,07	0,29	0,25
Corso Francia	**T**	**RH**	**Press**	**GSR**
LAG(24)	0,96	0,78	0,84	0,92
LAG(48)	0,93	0,68	0,62	0,78
LAG(72)	0,91	0,62	0,46	0,68
LAG(96)	0,89	0,60	0,36	0,60
LAG(120)	0,87	0,57	0,30	0,55

CONCLUSIONS

Usually a preliminary study of input data is always needed in order to remove not very meaningful data from the training sets, to choice a suitable normalization rule and to compute the usual statistical indexes of correlations among the variables.

Our research demonstrates that further considerations have to be applied in order to optimize the neural net performance.
In fact, the NN skill to capture the environmental information inside the data is highly dependent by the preliminary study of patterns. The generalization capacity of the net to forecast ozone peaks has to be connected with the essential information inside the data set and this information is not necessarily regularly distribute among all patterns. This problem is fundamental when the forecast of ozone is requested at different temporal lags.
To forecast ozone one day before, we take into consideration two MLP architectures. The first consists of meteorological variables, and the second of pollutants variables and the results obtained from the first net.
We test our strategy using two different datasets for the city of Rome during the 2005 year. Datasets refer to urban case and suburban case. We utilized the urban case for the training phase and the suburban as test case.
Results are in according to the measured ozone and substantially NNs cascade models seem a promising strategy to simulate the pollutants in very complex urban situations as verified for the Rome site, representing a typical Mediterranean area with the typical ozone by photochemical smog and heavy contributions of motor vehicles sources.

References

[1] *Carter, W.P.L, 1990. A detailed mechanism for the gas-phase atmospheric reactions of organic compounds. Atmospheric Environment, 24A, 481-518.*

[2] *Comrie R.S, (1997). Comparing neural network and regression models for ozone forecasting. Journal of the Air and Waste Management Association 47, 653-663.*

[3] *M.W. Gardner, S.R. Dorling" Statistical surface ozone models: an improved methodology to account for non-linear behaviour"Atmospheric Environment 34,2000, 21-34*

[4] *Gardner M.W, Dorling S.R., (1998). Artificial Neural Networks (the Multilayer Perceptron)- E Review of applications in the atmospheric sciences, Atmos. Environ., 32(14/15), 2627-2636.*

[5] *A.L. Dutot, J.Rynkiewicz, F.E. Steiner, J.Rude, "A 24-h forecast of ozone peaks and exceedance levels using neural classifiers and weather predictions", Environmental Modelling & Software, 22 , 9 September 2007, 1261-1269*

[6] *S.A. Penkett, M. J. Evans, C. E. Reeves, K. S. Law, P. S. Monks, S. J. B. Bauguitte, J. A. Pyle, T. J. Green, B. J. Bandy, G. Mills, L. M. Cardenas, H. Barjat, D. Kley, S. Schmitgen, J. M. Kent, K. Dewey, J. Methven" Long-range transport of ozone and related pollutants over the North Atlantic in spring and summer", Atmos. Chem. Phys. Discuss., 4, 2004, 4407-4454*

[7] *B. Finlayson-Pitt, J. Pitts, Atmospheric chemistry: Fundamentals and experimental techniques, A Wiley-interscience publication, John Wiley & Sons, 1986.*

[8] *Rojas R., (1996): Neural Networks: a systematic introduction, Springer-Verlag, Berlin Heidelberg.*

[9]*Hanna, S.R.," Confidence limits for air quality models, as estimated by bootstrap and jackknife resampling methods. Atmospheric Environment, 23, 1989, 1385-1395.*