A forecast model to predict the next day's maximum hourly SO_2 in the site of Priolo (Siracusa)

U. Brunelli, V. Piazza & L. Pignato
Dipartimento di Ricerche Energetiche ed Ambientali,
Università di Palermo, Palermo, Italy

Abstract

The purpose of this research is to develop a pure predictive model to forecast the next day's maximum hourly SO_2 in the site of Priolo in two representative monitoring stations. We globally propose a Recurrent Neural Network (RNN) with an inherent dynamic memory to forecast the fluctuations of ground concentration of SO_2 pollutant. The model uses an available time series which was recorded in the industrial site of Priolo, on the east coast of Sicily, Italy for the period between 1st April 1998 and 12th December 2001. The inputs used to train the neural network are three, where two are adimensional variable, which are obtained by the meteorological data of the site and the third is the stability classes of Thomas. This study has some important implications for health warning systems environmental management in places with high pollution concentration.
Keywords: air pollution, neural network, forecasting.

1 Introduction

Air pollution has a negative impact on the environmental and public health when it occurs in the lower atmosphere. In this paper the results obtained by an Artificial Neural Network to forecast the daily maximum SO_2 ground concentration are presented. A neural network is able to treat information that is uncertain and incomplete like the human brain, and it has been utilised to model complex non-linear functional relationships between predictor variables. The

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060571

pollutant concentration genesis and its dynamics are characterised by highly non-linear processes which are only partially known so then as artificial Neural Network (ANN) appears to be a good approach for pollutant concentration forecasting.

By using neural networks the authors have obtained very interesting results, as shown in Figures 4 and 5. Tasadduq et al. [2] use an MLP neural network trained with the BackPropagation algorithm and a batch learning scheme for the prediction of hourly mean values of ambient temperature 24 h in advance in Jeddah, Saudi Arabia. Full year hourly values of ambient temperature are used to train the network. The performance was evaluated by the mean percentage error between the predicted and measured values for three different years, considering a time window of 50 hours. Elman neural networks have been successfully used in other forecasting applications and time series prediction. Luk et al. [3] evaluated three alternative ANN models for rainfall forecasting: a Multi-Layer Feed-Forward Neural Network, an Elman Neural Network and a Time Delay Neural Network (TDNN). The study eventuates that the three approaches have comparable performance as long as the complexity of the network is variable. In details the Elman network had the simplest structure, but was complex at the same time. Kostela et al. [4] compared the performance of the Multi-Layer Perceptron, FIR and Elman Neural Networks in four different time series analyses: the performance of the Elman neural networks was better or similar to the other neural networks. In the paper, it is also argued that the efficiency of the learning algorithm is more important than the neural model used. In the last year neural networks have became an alternative to conventional methods and they are going to become an important instrument to model the distributions of air pollution (Nagendra and Khare [5]). Viotti et al. [6], use a multi-layer perceptron neural network to forecast short and middle long-term concentrations levels for O_3, NO_x, NO_2, CO. Hooyberghs et al. [7], describe the development of a Multi-Layer Perceptron neural network to forecast the daily average PM10 concentrations in Belgian urban areas one day ahead.

In this paper the authors use a Recurrent Neural Network for prediction of the next day's maximum hourly SO_2. The time series was recorded between April 1st, 1998 and December 31st 2001 and refers to two monitoring stations of air pollution parameters (Melilli, Farodromo) with a lead-time of one hour. These data are kindly provided by CIPA (Consorzio Industriale Protezione Ambiente). The paper is organised as follows. Section 2 gives a theoretical description of the area, section 3 a processing of the data, section 4 the background on the ANN model, and sections 5 and 6 the structures of the Neural Network model used and the obtained experimental results.

2 Site description

The region of Priolo is represented in Figure 1.

We can see that the coastal strip occupied by the industrial settlements extends from S to N laying on a bay, and is limited by land not only on the Eastern side, but also on the Northern one (Capo Izzo with the small town of

Augusta). More importantly, there are relatively high elevations at W-SW (Climiti Mountains, about 400-500 m high), which probably interfere with winds and breezes to generate vertical patterns. As will be seen, they influence the classical partition of the wind directions near the ground, causing an apparent reduction of the frequency of the breezes coming from the sea when compared to those coming from the land. In the natural cavity surrounding the bay there are scattered villages, roads and single houses at different heights above the sea. The pollution sources on the coast can be roughly divided into extensive and fixed; the first ones cover, with a high degree of continuity, the whole costal strip, while the second ones consist mainly of stacks of relevant heights (frequently comprised between 100m and 200 m). This is usually considered a good design, because the maximum ground concentration on a flat plane is inversely proportional to the square of the source height; in this case however, the ground is all but plain and in stable conditions the plumes can impinge on the sides of the mountains at heights which are the same as those of the polluting sources. This results in very high pollutant concentrations in the local area.

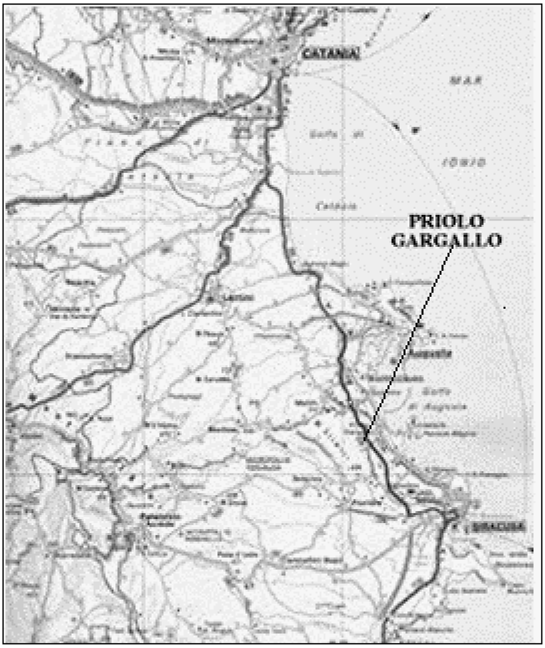

Figure 1: Description of Priolo site.

The facilities of interest are:

- A RASS (Radio Acoustic Sounding System), which is a remote sensing device that allows air temperature measurements of up to 1000 m of height with a spatial resolution of 20-30 m. Its working principle relies on the property of sound waves to propagate with different speeds depending on air temperature: a train of acoustic waves is then emitted along the vertical, letting the latter be reflected by the travelling wave fronts of the former. The Doppler effect between emitted and reflected radio waves then enables the computation of the sound speed and hence the ambient temperature at the height where the reflection takes place

- A pole of 10m supporting an anemometer located at 10m and two thermometers at 2 and 10m. Both thermometers are shielded to avoid ground radiation and provide an accurate estimation of the temperature gradient at the ground level.
- A SODAR (Sound Detection and Ranging) is able to determine the vertical distribution of wind velocity and direction and to give information on the vertical thermal structure and the turbulence in the first five hundred meters of the atmosphere.
- 11 fixed stations for measuring ground level concentrations of pollutants such as (NHC, CH_4, SO_2, O_3, H_2S, NO, NO_2).

On the Priolo site the monitoring stations are indicated with numbers 1-11; in particular the two monitoring stations of our interest are number 8 for Melilli and number 5 for Farodromo. The CIPA laboratory is indicated by number 12 and it is located on a slightly elevated spot (30m above sea) about 1 Km inland (Figure 2).

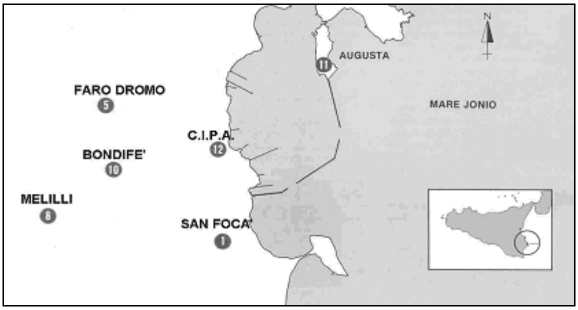

Figure 2: Map of monitoring stations.

3 Data processing

A neural network is a computational model based on data; it is clear that they are the principal component to obtain good results by the neural network.

The phases of their elaboration can be shared in:

1. collection and analysis of the data,
2. pre-processing following by the elaboration of the neural network,
3. post-processing if it is necessary to convert the output values of the neural network in the required format.

For our cases the hourly values used are:

- horizontal wind velocity (m/s)
- vertical wind velocity (m/s)

- base level of the layer of the atmospheric stability (m)
- gradient of the potential temperature (°C/m)
- the difference of the potential temperature of reference (°C)
- wind direction

The data set is recorded from April 1st 1998 to December 31st 2001 and is referred to sulphur dioxide monitoring stations with a lead-time of an hour. These data are submitted to the method of adimensional analysis to obtain same values which are independent of the characteristic of the industrial site which has been considered. By the application of this method the authors have obtained the following two adimensional numbers:

$$Pg = \frac{\gamma_{\vartheta}}{\Delta_{\vartheta}} H^3 \tag{1}$$

$$Pz = \frac{V_v}{V_o} \tag{2}$$

where V_v is the vertical wind velocity, V_o is the horizontal wind velocity, γ_θ is the gradient of the potential temperature, Δ_θ is the difference of the potential temperature of reference and H is the base level of the layer of the atmospheric stability. The third input used to train the neural network is the stability classes of Thomas [8]; they are adopted for a simple characterization of the stability of the superficial atmospheric layer. Thomas compared the values of the standard deviation of the vertical wind direction obtained with two different methods:

- measurements of the wind field, at a fixed height, with the SODAR;
- direct measurements of the standard deviation σ_φ, at the same height.

This last parameter was evaluated through the approximate relation:

$$\sigma_\varphi = arctg\left(\frac{\sigma_{vv}}{u}\right) \tag{3}$$

where σ_{vv} is the standard deviation of the vertical wind velocity and u is the wind in the prevailing direction.

In addition, each value was normalized in the range [-1, 1] using the following linear transformation:

$$X' = (X - V_m)/(V_{max} - V_{min}) \tag{4}$$

where X' is the new normalized value, X is the old value, V_{max} is the maximum of the considered data set, V_{min} is the minimum of the considered data set and V_m is the average value of the considered data set.

Finally the input pattern of the neural network is composed by three parameters: the adimensional numbers Pg and Pz and the stability classes of Thomas.

4 Artificial neural network

The Elman neural network (Elman, [1]) is also known as the partial recurrent network or simple recurrent network. In this network, the outputs of the hidden and output layer are allowed to feedback onto itself through a buffer layer, called the context layer. This feedback allows Elman networks to learn, recognize and generate temporal patterns, as well as spatial patterns. Every hidden neuron is connected to only one neuron of the context layer through a constant weight of value one. Hence, the context layer constitutes a kind of copy of the state of the hidden layer, one instant before. The number of context neurons is consequently the same as the number of hidden neurons. Optionally, every neuron of the output layer can be connected to only one neuron of a second context layer through a constant weight of value one (figure 3).

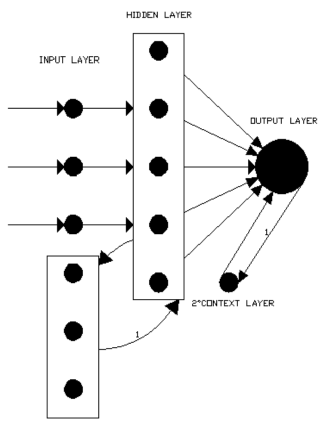

Figure 3: The Elman neural network.

Usually input, output, and context neurons have linear activation functions, while hidden neurons have the sigmoid activation function, but in this case the activation's function of output is the function RM.

$$a_j(t)=\begin{cases}0.85*a_j(t-1)+0.15*net_j(t)*(1-a_j(t-1))\xrightarrow{if} net_j(t)>0\\0.85*a_j(t-1)+0.15*net_j(t)*(1+a_j(t-1))\xrightarrow{if} net_j(t)\leq 0\end{cases} \tag{5}$$

and that for the hidden units is the function TanH:

$$a_j(t) = tanh(net_j(t) + \theta_j) \quad (6)$$

where $a_j(t)$ is the activation of unit j in step t, net_j is the input in unit j in step t and θ_j is the threshold (bias) of unit j. The training algorithm is the Resilient Back Propagation (RProp); it is a local adaptive scheme, performing fast and robust supervised batch learning in neural networks.

5 Network structure

The topology of neural network is a problem that depends on various factors. Whichever the type of ANN model employed, it is important to determine the appropriate network architecture in order to obtain the best results. Several artificial neural network topologies were implemented by changing the number of layers, and the number of the hidden and context units. Neural model simulations were performed using the Stuttgart Neural Network Simulator (SNNS) v. 4.1. The connection weights were initialized to zero-mean random values with adequate upper and lower bounds of (-1, 1). For the ANN the authors use the following values as input: two adimensional variable (Pg, Pz), which are obtained by the meteorological data of the site and the third is the stability classes of Thomas between April 1^{st}, 1998 and December 31^{st}, 2001, to define the optimum ANN structure using the common trial and error method. Different structures of each ANN have been tested with various different hidden nodes. It is found that fifteen hidden nodes are the optimum for both ANN, in this experiment. The primary aim of developing an ANN is to generalise the features of the processed time series. A popular technique to achieve generalisation, avoiding over fitting, is the early stopping method presented by Sarle [10]. According to this method, the generated data set was divided into two subsets: a training set and a test set. The whole training phase was stopped when the lowest error on the training set was reached. With more details, the training set is composed of 1310 values and the test set by 61 values (months of November and December 2001). In the conducted experimental trials, training epochs were set to 150 for each neural network model. To evaluate the model performance, the authors selected two parameters:

- MAE is defined by the following expressions:

$$MAE = \sum_{i=1}^{N} \left| O_i - P_i \right| / N \quad (7)$$

where O_i is the observed value at time i, P_i is the predicted value at time i and N is the total number of observations. For a perfect fit, $O_i = P_i$ and MAE = 0. So, the MAE index ranges from 0 to infinity, with 0 corresponding to the ideal condition, and in particular it permits one to compare the appropriateness of the networks used.

Linear Correlation Coefficient (r):

$$r = 1 - \left(\sum (O_i - P_i)^2 / \sum (P_i - P_m)^2 \right) \tag{8}$$

where P_m is the average value of the observed values. The linear correlation indicates the degree to which two variables are linearly related.

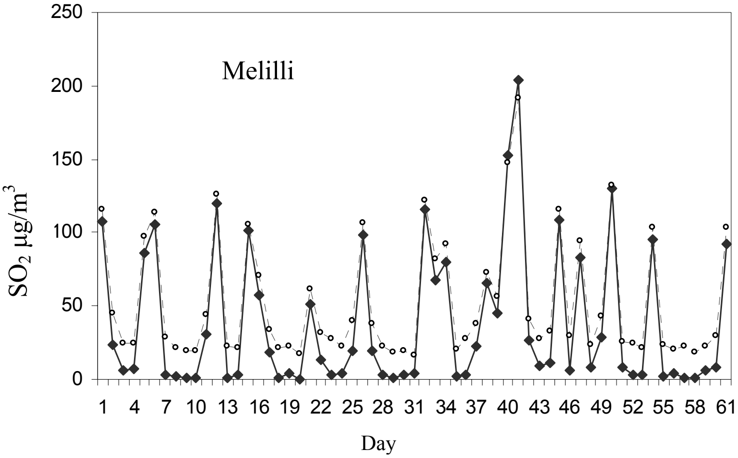

Figure 4: Results with Elman Neural Network, forecasted (white point) and measured SO_2 concentration (black point).

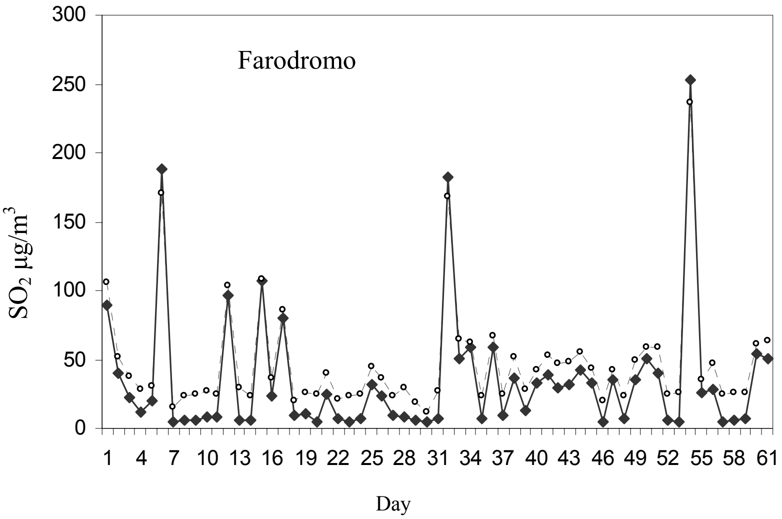

Figure 5: Results with Elman Neural Network, forecasted (white point) and measured SO_2 concentration (black point).

6 Experimental results and conclusions

The forecast of microclimatic phenomena represents a very complex activity with remarkable economic and sociable implications overall about the climatic changes which are happening. These processes, which are not linear at all and for whose cause the genesis and the dynamic are partially known, demand particular models of analysis and simulations to be forecasted with precision; in this situation the model based on artificial neural network represents a good solution for this problem.

The authors have used a Recurrent Neural Network (RNN) to forecast the values of hourly maximum daily sulphur dioxide, and they were tested in two stations of the Priolo region (Melilli and Farodromo). The set of test is constituted by the values of months of November and December 2001 and it is the same for all stations. The peculiarity of the model presented here is that it uses as input patterns two adimensional numbers (Pg, Pz) which are obtained from the meteorological characteristic of the studied place and the stability classes of Thomas. The quality of the developed model was assessed and compared on actual observed SO_2 concentration measurements of the test set and the results are showed in Figures 4 and 5. A statistic index has been used to provide a general indication of the relationship between observed and forecasted values. The selected performance measures included the mean absolute error (MAE) and the linear correlation coefficient (r). The general performance static for the two sites is summarized in Table 1. As we can see, the values of MAE, although being different numerically, are practically identical in terms of relative model performance; in particular, the values of r close to one show that the model gives good results.

Table 1: Statistic indexes.

Station	**r**	**MAE**
Melilli	0,85	16,25
Farodromo	0,84	7,58

Acknowledgements

The authors wish to thank the CIPA laboratory and the director Professor Giuseppe Zerbo for having supplied the experimental data used for this work.

References

[1] Elman, J. L. (1990), "Finding structure in time". In Cognitive Science, n.14 pp. 179-211.

[2] Imran Tasadduq, Shafiqur Rehman, Khaled Bubshait (2002) "Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia" in Renewable Energy n.25 pp.545–554.

[3] K.C. Luk, J.E. Ball, A. Sharma (2000) "A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting" Journal of Hydrology n.227 (2000) pp.56–65.

[4] Koskela T, Lehtokangas M, Saarinen J, Kaski K (1996), Time series prediction with Multilayer Perceptron, FIR and Elman neural networks Proceedings of the World Congress on Neural Networks. INNS Press, pp 491—496.

[5] S.M. Shiva Nagendra and Mukesh Khare. Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. Ecological Modelling, Corrected Proof, Available online 11 July 2005.

[6] P. Viotti, G. Liuti and P. Di Genova. Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia. Ecological Modelling, Volume 148, Issue 1, 1 February 2002, Pages 27-46.

[7] Jef Hooyberghs, Clemens Mensink, Gerwin Dumont, Frans Fierens and Olivier Brasseur. A neural network forecast for daily average PM_{10} concentrations in Belgium. Atmospheric Environment, Volume 39, Issue 18, June 2005, Pages 3279-3289.

[8] P. Thomas, "Stability classification by acoustic remote sensing", Atmospheric Research, 20, pp. 165-172, 1986.

[9] SNNS - Stuttgart Neural Network Simulator, url: http://www-ra.informatik.uni-tuebingen.de/SNNS/.

[10] W.S. Sarle, Stopped training and other remedies for over fitting, Proc. of the 27th Symp. On the Interface of Computing Science and Statistic, 352-360, (1995).