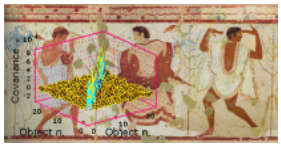

PM10 Forecasting by a Feed Forward Neural Network and Comparison with Multivariate Linear Regression Model

M. Caselli, L. Trizio, G. de Gennaro, P. Ielpo

Department of Chemistry, University of Bari, Via E. Orabona 4, 70126 Bari, Italy

Abstract

The problem of air pollution is a frequently recurring situation and its management has social and economic considerable effects.

Given the interaction of the numerous factors involved in the raising of the atmospheric pollution rates, it should be considered that the relation between the intensity of emission produced by the polluting source and the resulting pollution isn't immediate.

The development of a back propagation neural network is presented to predict the daily PM10 concentration one, two and three days early. The results obtained with the neural network were compared with those obtained by a multivariate linear regression model.

Introduction

The air we breathe every day can be contaminated by polluting substances emitted by industries, vehicles or other sources. These polluting substances can have bad effects both on human health and on the environment.

One of the most dangerous pollutant is PM10, that is particulate matter having an average aerodynamic diameter lower than 10 micron.

Health effects range from minor effects, such as nose and throat irritation, to more serious effects such as aggravation of existing respiratory and cardiovascular disease [1-2-3].

Short-term forecasting of air quality is needed in order to take preventive and evasive action during episodes of airborne pollution.

A classical forecasting method is based on multivariate statistical analysis, but now the Artificial Neural Network (ANN) is becoming an effective and popular mean alternatively to conventional methods [4-5-6].

Materials & Methods

The network type used in this paper is the feed forward back propagation. In this kind of network all the connections between neurons of the same level and the signal backward propagation are not allowed; the connections are fundamentally of the forward kind from which the name Feedforward Networks comes.

A monitoring network of the Municipality of Bari, located close the San Nicola stadium, constitute the primary data source of PM10 input data; the measuring is based upon the beta ray attenuation method on standard 47 mm membrane filters.

The forecasting of the major weather parameters (temperature, wind speed, pressure, relative humidity, rain) available on the web site (www.wunderground.com) and the forecasting of the Saharan dust, obtained by the "Centro Nacional de Supercomputaciòn" web site, satellite images and back trajectories analysis are used for the weather input data.

The optimization of the neural network was performed by means of a Simplex method [20] using the learning rate, the maximum weight increasing, the number of neurons and epochs as parameters. The network was trained with validated data of PM10 coming from San Nicola monitoring station from January 2005 to March 2006 and with meteorological data.

Results

After the training, the network was ready to predict the PM10 concentrations in the considered area.

The network was able to predict the PM10 concentration 12 h and 36 h before with an RRMSE of 26 and 30 percent respectively; the correlation coefficient between the foreseen and the target values was 0.7, showing a good accuracy in the prediction.

The results obtained by the neural network were compared with those from the multivariate regression model for S. Nicola monitoring station for one and two days forecasting.

In Fig. 1 the results for one and two days forecasting respectively are shown. In both graphs the grey line represent the real values, the triangles the regression values and the squares the neural networks values.

As one can see, the forecasting by the using of the neural networks gives more satisfactory results than that by the regression model in terms of errors (RRMSE). For both models, the higher errors verify for low PM10 concentrations. Moreover, particularly for two days forecasting, the neural network is able to fit better the measured PM10 trend, as it is shown in figure.

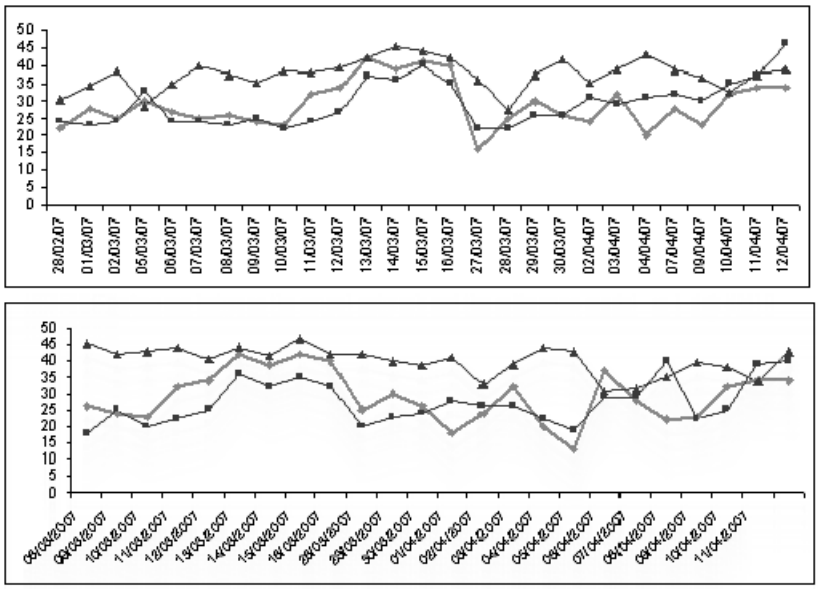

Fig. 1: Comparison between neural network and the regression model in 1-day (above) and 2-day (below) PM10 forecasting

Conclusions

Artificial neural networks provide a useful and effective tool for modelling the complex and poorly understood processes that occur in nature, as they are able to extract functional relationships between model inputs and outputs from data without requiring explicit consideration of the actual data generating process.

ANN can accurately model the relationships between PM10 concentrations and meteorological parameters and, increasing the number of input variables, improves the prediction performance of the model in terms of RRMSE.

The regression model fails to fit unexpected spikes, while it is quite able to fit spikes that are part of a trend, while the neural network proves more accurate in the forecasting of the overcoming of the alarm threshold.

References

1) D. W. Dockery, C. A. Pope III, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Jr Ferris, F. E. Speizer, An association between air pollution and mortality in six U.S. cities, N. Engl. J. Med., 329(24), (1993), 1753-9.
2) D. W. Dockery, F. E. Speizer, D. O. Stram, J. H. Ware, J. D. Spengler, B. G. Jr Ferris, Effects of inhalable particles on respiratory health of children, Am. Rev. Respir. Dis., 139(3), (1989), 587-594.
3) J. F. Gamble, PM2.5 and Mortality in Long-term Prospective Cohort Studies: Cause-Effect or Statistical Associations?, Environ. Health Perspect., 106(9), (1998), 535-549.
4) J. Hertz, A. Krogh, R. G. Palmer, Introduction to the Theory of Neural Computation (1991), Westview Press, ISBN: 0201515601.
5) R. Hecht-Nielsen, Proceedings of the International Joint Conference on Neural Networks, (1989), 543-611, IEEE Press, New York.
6) R. Hecht-Nielsen, Neurocomputing, Addison Wesley, (1990), ISBN:0-201-09355-3.