Application of a neural net filter to improve the performances of an air pollution model

A. Pelliccioni[1], T. Tirabassi[2]
[1] *Ispesl-Dipia Via Urbana 167, 00184 Rome, Italy*
[2] *ISAO -CNR, Via Gobetti 101, 40129 Bologna, Italy*

Keywords: Neural net, neural filter, air pollution modeling, puff model, atmospheric dispersion.

1 Introduction

Air pollution models have not so far been able to reproduce satisfactorily ground level concentrations because the influence of important variables is not perfectly described. For example, it is recognised that deterministic models cannot provide an adequate correlation between hourly predictions and observed data paired in time and space. From another point of view, the concentration levels forecast by a supervised Neural Net model (Faussett, 1994) have to take account of the influence of the variables of the system, such as source emission factors, turbulence conditions, local topology, reactions rate, by using an appropriate training on the available experimental data.
The proposed approach relies on the development of an integrated model that optimises the performances of each methodology. In particular, we filter the concentrations evaluated by an air pollution model with a Neural Net so as to account for the systematic influence of important variables.
We have applied the Neural filter to a puff model (SPM: Skewed Puff Model) based on the Monin-Obukhov similarity theory, where continuously emitting sources are represented by the superposition of a series of puffs. The model utilises approximate solutions proposed by van Ulden (1992) for the dispersion of a cloud of passive contaminants released from an instantaneous source near the ground. Particular importance is attributed to describing the interaction between wind shear and vertical diffusion and to the process that transforms shear-produced skewness into diffusive variance in the wind direction. The equations describing the concentration field can be found in Tirabassi and Rizza (1995).

2 Neural net filter

In forecasting with neural networks, the most popular tool is provided by multilayer perceptrons with an error-backpropagation supervised learning rule (Rojas, 1996). This net architecture is able to reproduce non linear models, without any *a priori* assumptions, by means of an accurate choice of the variables of the system and of the meaningful patterns.
A learning algorithm is an adaptive method by which a network of computing units self-organises to reproduce the desired model. This is done in learning algorithms by presenting some examples of the desired input-output mapping to the network. A correction step (the error-backpropagation rule) is performed iteratively until the network learns to produce the desired response.
As architecture we used a 3-layer perceptron model. The first input layer contains the input variables of the net, in our case atmospheric turbulence through U* (friction velocity) and L (Monin-Obukhov length), the distance of the receptor from the source (X), and the concentration levels predicted by the model. The real novelty of the proposed methodology lies in the choice of the latter variable: in fact, the inclusion of the predicted model concentrations as input values of the network means that it must perform a twofold task. The first of these is to start from a situation close to reality (insofar as any dispersion model will provide a response to different conditions of turbulence and emission). The second, and conceptually more important task is linked to the fact

that models perform well under certain hypotheses, while tending systematically to differentiate in performance when reality falls short of the ideal situations. In this case, the NN functions as a filter of the model, correcting it so that it can give the best reproduction of the real situation.
The second layer consists of the neurons of the hidden layer. The number of neurons of the hidden layer is one of the parameters to be chosen in the perceptron model and was estimated from the number of the clusters present in the data set.
The third layer is the output layer, which consists of the target of the forecasting model.
As the activation function of the single neurons, we chose the sigmoid function:

$$F(P) = \frac{1}{1 + e^{-(p-s)}}$$

where P is the activation potential and S is the activation threshold.
The model optimisation mechanism takes place through the automatic update of the weights. The update of weights among neurons is guided by the following function:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}$$

where □ is the learning rate (a characteristic parameter for updating) and E is defined as:

$$E = \frac{1}{2} \sum_{j=1,npatt} (y_j - \overline{y_j})$$

and E provides a quantification of the overall difference over all the examples and for all the reproduced variables.
In the present work, we used the conjugate gradient method for weight correction, as it was judged to be best for identifying the absolute minima of the function E.

3 Validation against experimental data

We evaluated the performances of the SPM model in the cases of an emission source near the ground, using the well known data set of the Prairie Grass experiment (Barad, 1958), as well as an elevated source, using the Copenhagen data set (Gryning and Lyck ,1984). Generally speaking, the distributed data set contains hourly mean values of concentrations and meteorological data. However, for this model validation, we were able to use data with a greater time resolution, kindly made available to us by Gryning. In particular, we used 20-minute averaged measured concentrations and 10-minute averaged values for meteorological data.
The perceptron model used for the two simulations is made up of a 3-layer architecture with 4 neurons in the input layer (containing the variables U*, L, X and C_{SPM}), 9 neurons in the hidden layer and one output neuron (containing the concentration levels to be reproduced).
Given the low number of patterns available (60 for the Copenhagen data set and 167 for the Prairie Grass data set), all the data were used during the training phase for the two simulations. To optimise learning, the two nets were trained using respectively the following pairs of values for the learning parameter and for momentum for four different training phases: (0.4, 01); (0.3, 0.4); (0.2, 0.5);
(0.1, 0.6)
Each phase consisted of 500 patterns, so that the results obtained were eventually found after 2000 examples.

In order to assess the reproducibility of the concentration levels, we calculated the percentage of difference between the levels calculated by the two models and the experimental ones:

$$\Delta C(\%) = \frac{(C_{sperim} - C_{SPM+RN})}{C_{sperim}} \cdot 100$$

A comparison between the performances of the dispersion model and the new proposed methodology is shown in Table 1, which reports the squared correlation coefficients (R^2) between the measured and reproduced concentration levels. The values are relative to Figures 1 and 2 for the Copenhagen and Prairie data sets.
As can be seen from the correlation coefficient values, the levels of concentration reproduced for both data sets show a marked improvement when the neural network is added downstream of the dispersion model.

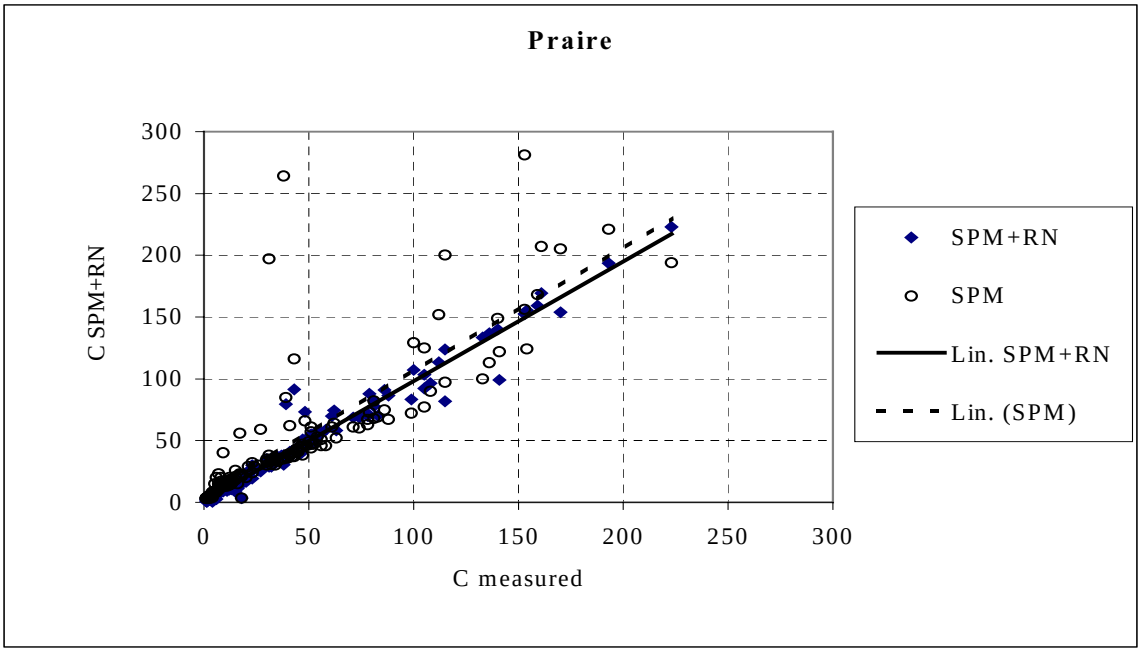

Figure 1 Praire dataset: Comparison between measured concentrations levels vs SPM model and SPM+RN model.

For both data sets, the SPM model produces underestimations. In fact, the mean value of reproducibility for the dispersion model is –45.77% for the Prairie data set and –16.01% for the Copenhagen one. Applying the neural net one sees a clear improvement of the concentration levels, with mean values of reproducibility of –6.95 and –0.80 for the Prairie and Copenhagen data sets respectively.

Table 1 Values of the correlation (R^2 and Bias) and mean value (m) and standard deviation (σ) of the distribution of reproducibility of concentration levels (ΔC(%)).

	Prairie			
	R^2	Bias	m	σ
SPM Model	**0.72**	**5.59**	**-45.77**	**85.42**
SPM+RN	**0.97**	**1.30**	**-6.95**	**31.98**
	Copenhagen			
	R^2	Bias	m	σ
SPM Model	**0.35**	**1.26**	**-16.01**	**43.73**
SPM+RN	**0.98**	**0.11**	**-0.80**	**9.89**

4 Conclusions

Using the results of a dispersion model, the reproduced concentration levels were evaluated for two typical calibration data sets. The model reproduces concentration levels with a precision that is not always optimal.

Given this fact, a neural net was trained on the basis of the meteorological situation and the values for the pollutant levels reproduced by the model.

Applying this methodology, the ground level concentrations indicate an improvement in the predictive capacity of dispersion models, when appropriately filtered by a neural net. Notwithstanding the limited nature of the example dealt with here, the proposed methodology can be generalised, thus opening new perspectives for integrated models in the simulation of complex situations.

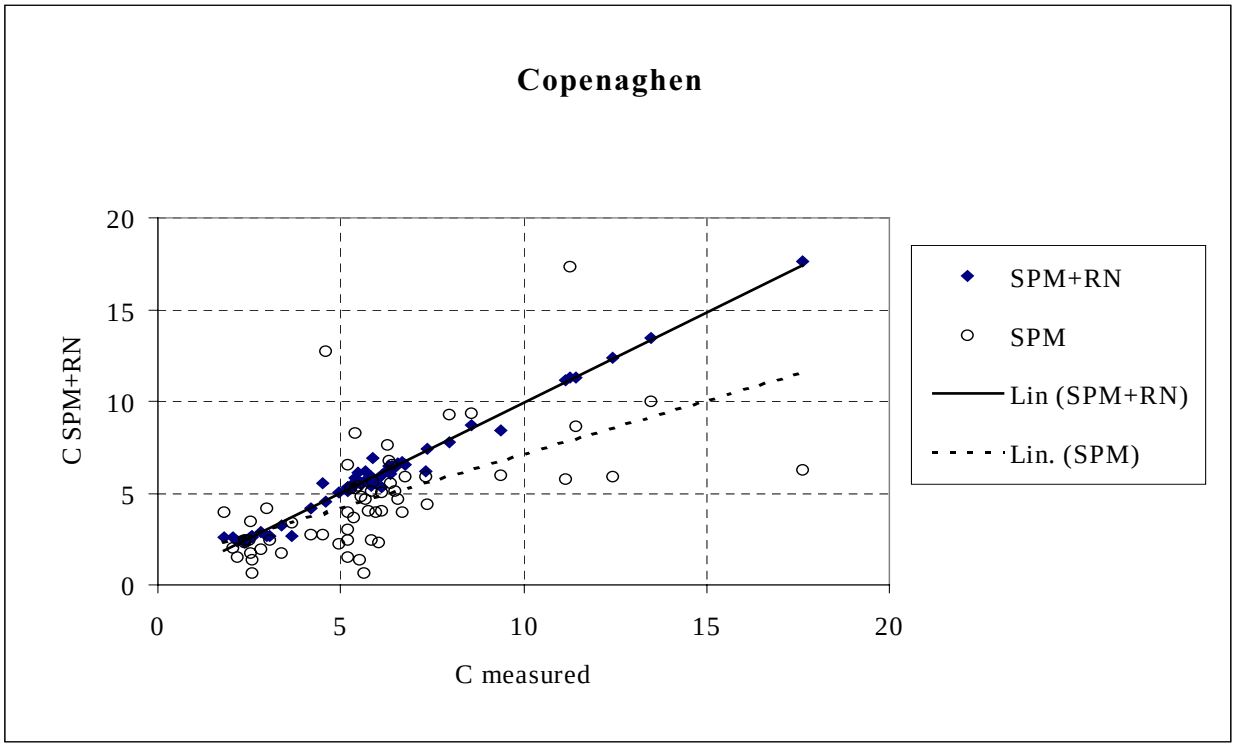

Figure 2 Copenaghen dataset: Comparison between measured concentrations levels vs SPM model and SPM+RN model.

5 References

1. Barad M.L., 1958. Project Prairie Grass, a field program in diffusion. Report AFCRL-TR-235.
2. Faussett, L (1994), Fundamentals of Neural Networks. Architectures, Algorithms, and Applications, Prentice Hall International Editions, 1994..
3. Gryning S.E. and E. Lyck, (1984). Atmospheric dispersion from elevated sources in an urban area: comparison between tracer experiments and model calculations. J. Clim. Appl. Meteor., Vol. 23, 651-660.
4. Rojas R., Neural Networks: a systematic introduction, Springer-Verlag, Berlin Heidelberg, 1996.
5. Tirabassi T. and Rizza U. (1995) A practical model for the dispersion of skewed puffs. J. Appl. Meteor., Vol. 34. pp. 989-993.
6. Van Ulden, 1992. A surface-layer similarity model for the dispersion of a skewed passive puff near the ground. Atmos. Environ., Vol. 26 (A), 681-692.