

#### **4.18 COUPLING OF A LAGRANGIAN PARTICLE MODEL (SPRAY) AND NEURAL NETWORK TO IMPROVE THE ACCURACY ON IMPACT ASSESSMENT OF AN INDUSTRIAL FACILITY**

*A. Pelliccioni<sup>o</sup>, C. Gariazzo<sup>o</sup>, T. Tirabassi\**

<sup>o</sup> Ispesl-Dipia Via Urbana 167, 00184 Rome, Italy.

\* CNR (ISAC)- Via Gobetti 101, 40129 Bologna, Italy

##### **INTRODUCTION**

The accuracy of a modeling system in analyzing the consequences of either continuous or accidental releases in the atmosphere is important especially when adverse health effects are expected to be found. Stationary models are often used in consequences analysis studies. Unfortunately such models can only be applied in stationary and homogeneous conditions, which are often far to be true and consequently the results are poor. An improvements on model accuracy can be obtained using non conventional approaches where dispersion models are coupled with statistical ones. Among the statistical models the neural network (NN) have shown to better deal with non linear problems coming in different fields of study such as prediction of air pollution concentration levels (Gardner et al., 1999), atmospheric turbulence parametrization (Agnello et al., 2001) and so on. Recently Pelliccioni et al (Pelliccioni et al., 2003) applied this kind of statistical approach to the results obtained by a gaussian dispersion model to investigate the ability of a NN in improving the accuracy on reproduction of the observed ground concentrations. A net improvements on accuracy of the coupled dispersion-NN model system were observed. This kind of study was conducted using the Kincaid validation data set (Bowne and Londergan, 1981), collected in flat terrain conditions. When the pollutants are dispersed in complex areas the reconstruction of the actual ground concentrations becomes more complicated. In fact on such areas the land/sea breezes and topography effects give rise to complex circulation patterns which have a great influence on local meteorology and in turns on pollutants dispersion. On such conditions a non stationary modeling system can achieve better accuracy than that obtained using stationary models such as the gaussian ones. Among the non stationary model the Lagrangian particle models have demonstrated to better deals with non stationary non homogeneous conditions like those described above. Although better results are obtained with this kind of model, the accuracy is sometimes poor and it needs to be improved. In order to get this aim a Lagrangian particle model was coupled with a neural network. This model system was then applied to reconstruct the ground concentrations produced by a cement plant located in a complex area and to investigate the possibilities to improve the accuracy on reproduction of the observed concentrations.

##### **INVESTIGATION AREA**

The studied area is located in central Italy, approximately 30 km far from Rome. The south-west side of the domain is rather flat and includes a military airport, while in the north and east parts hills mountains up to 1200 m are present. Guidonia is the most important town in this area with a population of approximately 30000 inhabitants.

Different emission sources are present. Vehicular traffic is an important emission source in the town and its suburbs. A large cement plant is located out of the city of Guidonia, about 1 km north from the boundary of the urban area. The plant facilities cover an extension of 3x2 km<sup>2</sup> including extraction activities, components preprocessing and final burnings. The main emissions consists of NO<sub>x</sub> and SO<sub>2</sub> pollutants from a 54 m height stack which is connected with a large burning system.

## FIELD CAMPAIGN DESCRIPTION

A field campaign has been conducted to feed the models with real data and to validate the modeling system. Three monitoring stations were used for this study: the first is close to the cement plant (UNI hereafter) and located 800 m south of the stack. The second monitoring station is placed downtown the Guidonia city (GUID hereafter) handled by the local environmental protection agency. Both UNI and GUID stations collect hourly chemical and meteorological data. The last station is composed by a Mobile Meteorological Laboratory (MML) and by a Mobile Chemical Laboratory (MCL-ISP). The MML was located 300 m east of the stack. Its location was chosen in order to measure meteorological parameters close to the emission point. The MML calculates averaged values (10 minutes time period) of the main standard meteorological data and turbulence parameters. Wind and turbulence vertical profiles were also collected at this station. The MCL-ISP, equipped to measure air pollutants concentrations, was located 600 m north-east of the stack, in a rural area at the foot of the Montecelio hill.

## THE DISPERSION MODEL AND THE NEURAL NETWORK

The dispersion modeling system used in this study is composed by three models: the MINERVE meteorological model, the SURFPRO turbulence model and the SPRAY Lagrangian particles dispersion model. The Lagrangian particle model SPRAY 3.0 (Tinarelli et al., 1994) was used to reproduce the concentration fields produced by the buoyant plumes emitted by the cement plant stack. SPRAY is a 3-D model able to simulate air pollution dispersion in the atmosphere in non homogenous and non stationary conditions. The model deals with a number of computational (fictitious) particles which are emitted and dispersed taking into account the three basic dispersion components: the transport due to the mean fluid velocity; the random turbulent fluctuation of wind components; the molecular diffusion. Minerve is a mass-consistent meteorological model (Geai, 1987) which uses a diagnostic approach to reconstruct 3D wind and temperature fields using the provided meteorological data. A two steps procedure is conducted by the model: a first guess interpolation scheme and a divergence adjustments.

As Neural Network architecture, we have considered a 3 layers perceptron structure with a hidden layer, one of input and one of output layer.

## MODELS SETUP

A 10x10 km<sup>2</sup> domain, centered on the plant stack, was considered to cover all possible plume impacts on the surrounding areas and all relevant towns. The Minerve-Spray domain has been horizontally divided into 41 x 41 grid cells with 250 m resolution and vertically splitted from the ground level to the top, set to 1500 m, using layers of variable thickness.

The period of November 2nd-5th 2001 was used for test simulations. It can be considered as a typical local atmospheric circulation in autumn. Furthermore, significant NO<sub>x</sub> peaks were observed in this period which could be ascribed to the stack emissions. All available upper air and ground based meteorological stations have been provided as input to the meteorological model MINERVE to calculate 3D wind and temperature fields at 10 minutes time resolution. Emissions factors and its parameters were measured at the stack level. These parameters allowed to provide to the model the real stack emission factors and their time modulation in connection to the actual plant working conditions. NO<sub>x</sub>, SO<sub>2</sub> and CO were considered as simulated pollutants. Further details can be found in Gariazzo et al. (Gariazzo et al., 2004)

In order to find the best architecture NN which reproduce to the best the experimental data, we have tested three different neural nets, each associate to the number of neurons of the

hidden layer (corresponding to the choice of 5, 8 and 10 neurons). As any NN needs to be trained before to be applied for reproduction of the target parameter, we evaluated the performance of the neural net taking into account different percent of training input data from 40% up to 100% of the total data set.

The choice of the training parameters of the neural net is another key factor for the success of these techniques. Six variables were selected as input. Three meteorological parameters, the friction velocity, the vertical gradient of temperature and the mean temperature, were considered, all measured by the MML. Others two variables (the geometrical coordinates of the monitoring stations  $X(t)$  and  $Y(t)$ ) have been selected with the introduction of a new system of coordinates linked with the wind direction and the stack-monitoring stations distance to provide NN the information related with downwind-upwind conditions and impact distances. As these geometrical coordinates are calculated starting from the wind directions, they are time dependent. As a result the point of coordinates  $X(t), Y(t)$  moves on a circumference of radius equal to the stack-station distance (736.2m, 698.9 m and 1859.3 m for the stations of the ISP, of the UNI and of the GUID respectively) according to the wind direction. The different anemological conditions that are measured during the simulation correspond to different upwind distances ( $Y > 0$ ) and downwind ( $Y < 0$ ) for the three stations.

The last input training variable was the  $\text{NO}_x$  concentration values foreseen by the Lagrangian model (so called the Concentration Levels Predicted by Dispersion Model or CLPDM) at the three considered monitoring stations. This variable is very important because it contains the information related with the atmospheric dispersion of pollutants. To evaluate the importance of the choice of this variable, a further simulation has been performed without its inclusion in the NN input parameters (without CLPDM).

As NN target output parameter the  $\text{NO}_x$  ground concentrations at the three monitoring stations was chosen. This pollutants was selected due to its abundance on emission and for the presence of significant peaks on the observed ground concentrations.

## RESULTS AND DISCUSSION

The NN conducts its training phase starting from the input dataset patterns. A total number of 325 patterns coming from all monitoring stations were used for such a phase.

The NN produces some negative values of target concentrations, which don't evidently have any physical meaning. Their number depend on the NN architecture (number of hidden neurons). The best architecture was searched by minimizing the number of negative concentration produced. The relative percentage of their presence respect to the total dataset has been used as an indication of the simulation quality. The correlation factor R has also been considered as a quality index. The results are shown in the table 1.

*Table 1. Main results of Spray-NN model and NN alone (without  $\text{Nox}(\text{Spray})$ ).*

8 Hidden Neurons	Percent of data during the training phase								
	40%	50%	60%	70%	80%	85%	90%	95%	100%
<b>Spray-NN</b>									
Correlation ( R )	0.44	0.57	0.61	0.72	0.77	0.77	0.80	0.81	0.86
Percent of Negative Concentrations (%)	11.4	10.2	8.0	3.7	6.2	4.0	2.8	2.8	0.9
<b>NN alone</b>									
Correlation ( R )	0.55	0.59	0.68	0.70	0.65	0.75	0.76	0.76	0.76
Percent of Negative Concentrations (%)	8.0	7.1	4.3	3.4	1.5	2.2	2.8	3.1	3.1

In general, the best result was obtained with 8 neurons of the hidden layer and using 100% of the data ( $R=0.86$  and 0.9% of negative values). Nevertheless, some good results was also given beginning from 70% of the data ( $R=0.72$  and 3.7% of negative values).

The simulation, besides, demonstrate that the use of the  $\text{NO}_x$  Spray derived as input variable is very important for the performance of the net. In fact, as shown in table 1, the R values are

lower (0.76 versus 0.85 using 100% of training data and 8 hidden neurons) in comparison with the results obtained using the NO<sub>x</sub> Spray derived data. Short-term results produced by Spray were deeper analyzed by Gariazzo et al. (Gariazzo, 2004). They showed Spray sometimes missed the observed peaks due to other emission sources (eg. traffic) not included in that study or to an incorrect reproduction of the actual wind field which in turn affects the simulated peak shape that exhibits a temporal shift when compared with the observed one. Episodic incorrect peak reproduction are also outlined in this study although daily concentration are in a good agreement with the observed one

Figure 1 shows the Spray predicted Nox concentrations, the 8 hidden neurons 100% training data NN derived NO<sub>x</sub> concentrations and the related observed NO<sub>x</sub> concentrations at the three selected monitoring stations. In the figure 2 the Spray-NN Vs observed values scatter plot results are given. The net succeeds in adjusting the Spray results operating on two main factors. The first factor attempt to adjust the peaks of the maximum plume impact (to certainly be ascribed to the cement factory) and to fix the temporal shift produced by Spray.

The second factor operates on situation where observed values are mainly produced by other emissions sources different from the stack, which was the only one considered in the Spray simulations. This has particular relevance in the Guidonia monitoring station, where traffic emissions are, at rush hours, the main contributors to the measured pollutants concentrations.

It is important to be noticed that the introduction of the new spatial coordinates in the NN input variables, allows to extend the spatial estimation of ground concentrations, taking into account the location of the selected station and its distance from the impacting sources as well as the dispersion conditions as provided by the dispersion model.

In conclusion, the comparison of simulation results with the observations collected at selected monitoring stations have shown good agreement for NO<sub>x</sub>. To net improvement in the overall models accuracy is observed when the Neural network was applied downstream to the particle model.

## REFERENCES

- Agnello P., Gariazzo C., Pelliccioni A., 2001*: “Modeling PBL turbulence parameters by neural network to improve meteo station”; 7<sup>th</sup> International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Belgirate 28-31 May 2001.
- Bowne N.E. and Lodergan R.J., 1981*: Overview, result and conclusions for the EPRI plume model validation and development project: plane site. EPRI report EA-3074.
- Gardner M.W, Dorling S.R., 1999*: Neural network modelling and prediction of hourly Nox and NO concentrations in an urban area in London, *Atmos. Environ.*, 33(5), 709-719.
- Gariazzo C., Pelliccioni A., Bogliolo M. P., Scalisi G., 2004*: ‘Evaluation of a Lagrangian Particle Model (SPRAY) to assess environmental impact of an industrial facility in complex terrain.’, *Water, Air and Soil Pollution*, in press.
- Tinarelli G., Anfossi D., Brusasca G., Ferrero E., Giostra U., Morselli M.G., Moussafir J., Tampieri F., and Trombetti F., 1994*: ‘Lagrangian particle simulation of tracer dispersion in the Lee of a schematic two-dimensional hill, *Journal of Applied Meteorology*, 33, pp. 744-756.
- Geai, P., 1987*: ‘Methode d’interpolation et de reconstitution tridimensionnelle d’un champ de vent: Le code d’analyse objective MINERVE’, Rep. ARD-AID: E34-E11, EDF, Chatou, France.
- Pelliccioni A., Tirabassi T., Gariazzo C., 2003*: Coupling of Neural Network and Dispersion Models: a novel methodology for air pollution models. *Int. J. Environment and Pollution*. Vol. 20, Nos 1-6, 2003. pp. 136-146

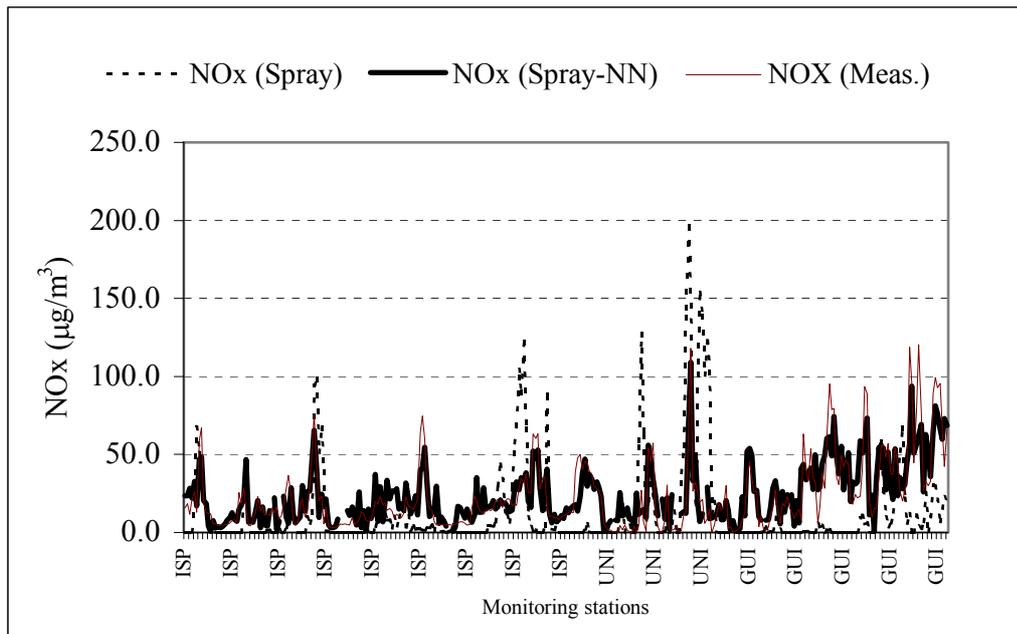


Figure 1. Time series of NO<sub>x</sub> concentrations predicted by Spray and Spray-NN superimposed with NO<sub>x</sub> observed values at the three selected stations.

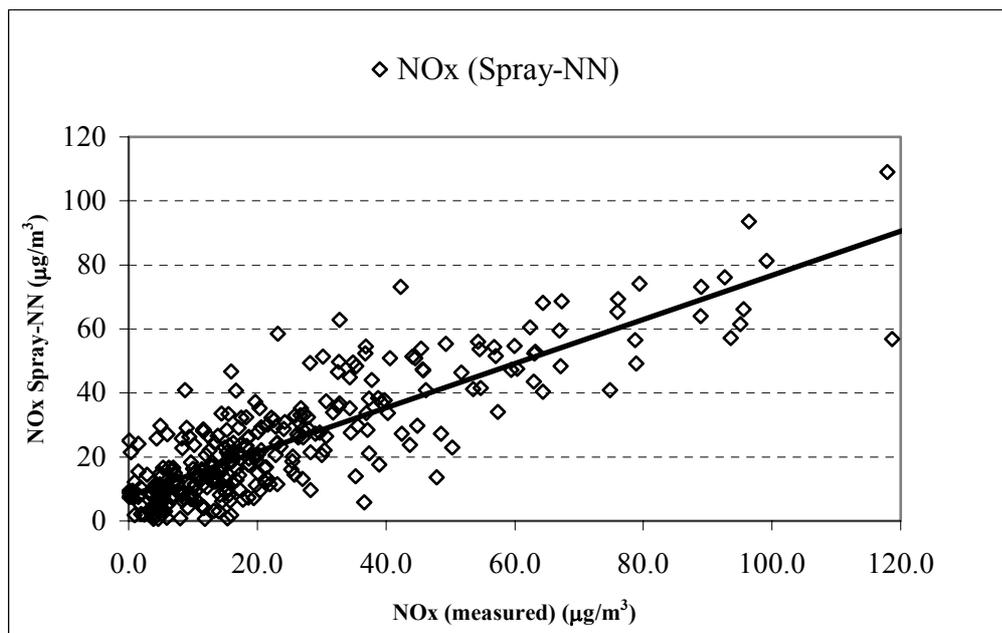


Figure 2. Scatter plot of Spray-NN derived Vs observed NO<sub>x</sub> concentrations.