IL NUOVO CIMENTO Vol. 31 C, N. 3 Maggio-Giugno 2008
DOI 10.1393/ncc/i2008-10302-9

Air pollution model and neural network: An integrated modelling system

A. Pelliccioni([1])(*) and T. Tirabassi([2])(**)

([1]) *Ispesl-Dipia - Via Fontana Candida 1, 00040 Monteporzio Catone, Italy*
([2]) *ISAC-CNR - Via Gobetti 101, 40129 Bologna, Italy*

(ricevuto il 30 Luglio 2008; revisionato il 10 Dicembre 2008; approvato il 9 Gennaio 2009; pubblicato online il 2 Marzo 2009)

Summary. — It is well known that neural networks can work as universal approximators of non-linear functions and they have become a useful tool either where any precise phenomenological model is available or when uncertainty complicates the application of deterministic modelling as, for example, in environmental systems. Usually, NN models are using as regression tool. We have developed an integrated modelling system coupling an air dispersion model with a neural network method both to simulate the influence of important parameters on air pollution models and to minimize the input neural net variables. In our approach, an optimised 3-Layer Perception is used to filter the air pollution concentrations evaluated by means of the non-Gaussian analytical model ADMD. We applied this methodology to the well-known Indianapolis urban data set which deals with a release of pollutants from an elevated emission source.

PACS 92.60.Sz – Air quality and air pollution.
PACS 07.88.+y – Instruments for environmental pollution measurements.
PACS 87.85.dq – Neural networks.

1. – Introduction

The management and control of air quality implies a knowledge of the state of the environment. Such knowledge involves both cognitive and interpretative aspects. In general, monitoring networks, meteorological measurements and inventory of emission sources, are fundamentals for the construction of the cognitive picture, but not for the interpretative one. In fact, air quality control requires interpretative tools that are able to extrapolate in space and time the values measured by monitoring stations, while environmental improvement can only be obtained by means of a systematic planning of reduction of emissions, and, therefore, by employing tools (such as mathematical models

(*) E-mail: armando.pelliccioni@ispesl.it
(**) E-mail: t.tirabassi@isac.cnr.it

of atmospheric dispersion) able to linking the causes (sources terms) of pollution with the respective effects (pollutant concentrations measured at monitoring stations).

The processes involving transport and diffusion of pollutants are numerous; such complexity would be impossible to describe without the use of mathematical models. Air dispersion models therefore constitute a crucial technical instrument of air quality management.

There exist many and different mathematical models to simulate atmospheric pollutant dispersion, that may be utilized for the aforementioned target. In fact, the phenomenon of turbulent diffusion and transport in the atmosphere has no unique mathematical formulation, in the sense that none approach has yet been proposed that is able to explain all of the observed phenomena.

The air pollution models constitute a mathematical deterministic tool that reflects the knowledge on turbulent transport in the atmosphere and the results are affected by uncertainties, mainly linked with the description of the atmosphere and to local orography, by choice of model parameters and of model variables, and, moreover, by the position and intensity of air pollution sources.

It is well known that each dispersion model is designed to work in some specific conditions, and the results depend on the agreement between model assumptions and actual physical conditions.

A useful approach, alternative to the deterministic modelling, is based on statistical models.

In fact, often the statistical tools are used to forecast pollutant concentrations and the more used one is the neural network model.

Among the complex systems the neural networks can be used considered as universal approximators of non-linear functions and, consequently, can be used in assessing the dynamics of such systems. Usually, they have become a useful tool either where correct phenomenological models are not available or when uncertainty in input and output data complicates the application of deterministic modelling as may happen, for example, in environmental systems.

Pioneering works in developing Artificial Neural Net (ANN) applications for short-term forecasting in atmospheric systems have been conducted since the early 1990's [1] (*e.g.*, Boznar *et al.*, 1993, constructed a model to predict atmospheric sulphur dioxide in a polluted industrialized area of Slovenia).

ANN methods have been developed for forecasting daily maximum ozone levels in various urban areas, using average daily meteorological data as input parameters [2].

Gardner and Dorling [3] produced an overview of applications of ANN in the atmospheric sciences and in 1999 tested the benefits of using a MLP to model NO_2 concentrations in London relative to other statistical modelling approaches.

Kolehmainen *et al.* [4] evaluated various computational models using hourly concentration time series of NO_2. The methods have been combined in order to optimise their performance [5].

All above environmental applications use the NN model as regression tool.

In the field of post-processor systems of Global Climate Models (GLMs), an interesting work can be found in Pasini and Pelino [6]. They attempt to explain the predictability for the well-known Lorenz system by using a neural net architecture in recognising different zones of attractors. The use of NN seems to produce low values of errors in evaluation of the bred growth classes.

Our approach consists of a combination of statistical models (neural net) and deterministic models (air dispersion models).

We developed an integrated modelling system coupling an air dispersion model with a neural-network method in order to adjust the influence of remarkable variables on dispersion models (which may produce systematic under- or over-predictions) and, contemporaneously, to minimize the number of input variables to the neural net model. A review of this methodology on environmental applications can be found in [7]. In our approach, an optimised 3-Layer Perception with error-backpropagation learning rules is used to filter the air pollution concentrations estimated by using an operative analytical non-Gaussian model ADMD. We applied this methodology in the case of an emission from an elevated source using the Indianapolis urban data set.

In order to better understand the actual improvement of our approach, we will give a brief description of the main assumption and difficulty inside the air dispersion model.

2. – Models and methods

The relevance of the reproduction of pollutant levels by mathematical models is one of the most complex tasks. In fact, the pollutant depends on many factors, such as wind direction and velocity, intensity of turbulence, chemical reactions and the source apportionment. Further, the boundary conditions are very complex to be determined but they have high impact on the simulations. Ultimately, some new techniques of data assimilations [8] were applied to the boundary conditions to the actual meteorological data and to the concentrations measured at monitoring stations.

Given all above, it is possible to obtain reliability results on the pollutants dispersion in some simple situations. However, while simple situations to be modelled regard mainly low-density areas or relatively flat terrains, all people live mainly in high-density urban areas and the related prediction of the pollutant is very difficult to be fairly predicted.

In order to explain the novelty of our approach to the environmental data processing, we show below a description both of air pollution models and neural net model assumptions.

2˙1. *Brief description of air pollution models.* – Every air pollution model is characterised by some basic assumptions, that involve different mathematical approaches or parametrization (see, for example, [9]).

In the K approach, diffusion is considered, at a fixed point in space, proportional to the local gradient of the concentration of the diffused material. Consequently, pollutant diffusion is fundamentally of Eulerian type since it considers the motion of fluid within a spatially fixed system of reference. Such models are most suited for complex problems as, for example, the dispersion of pollutants over complex terrain or the diffusion of reactive pollutants. They are based on the numerical resolution, on a fixed spatial-temporal grid, of the equation of the mass conservation of the pollutant chemical species.

Among the Eulerian models, box models constitute the most simple mathematical approach since they neglect the spatial structure of phenomena. They assume that pollutants are uniformly distributed within a box. From the theoretical viewpoint, this is equivalent to assume infinite diffusion coefficients which induce an instantaneous propagation of the pollutant within the considered box. The pollutant present in the box originates from inside sources or from external contributions transported by the wind or flows across the upper layer of the box itself, which coincides with the height of the mixing layer.

Gaussian models are theoretically based upon an exact, but not realistic solution of the equation of transport and diffusion in the atmosphere, in those cases where both

wind and turbulent diffusion coefficients are constant with height. The solution is forced to represent real situations by means of empirical parameters, referred to as "sigmas" dispersion coefficients. They can be either stationary (plume models) or time-dependent (puff models). The name associated to these models is due to the vertical and crosswind pollution concentration, that is described by the famous distribution discovered by the physicist-mathematician Gauss. The various versions of Gaussian models essentially differ in the techniques utilized to calculate the sigmas as a function of atmospheric stability and the downwind distance from the emission source.

2˙2. *K models*. – Eulerian approach for modelling the statistical properties of the concentrations of contaminants in a turbulent flow as the Planetary Boundary Layer (PBL) is widely used in the field of air pollution studies. Within this frame, the diffusion equation that describes the concentrations arising from a continuous point source can be written as

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} + v\frac{\partial c}{\partial y} + w\frac{\partial c}{\partial z} = -\frac{\partial \overline{u'c'}}{\partial x} - \frac{\partial \overline{v'c'}}{\partial y} - \frac{\partial \overline{w'c'}}{\partial z} + S, \tag{1}$$

where c denotes the average concentration; u, v and w are the Cartesian components of the wind and S is the source term. The terms, $\overline{u'c'}$, $\overline{v'c'}$ and $\overline{w'c'}$ represent, respectively, the turbulent fluxes of contaminants in the longitudinal, crosswind and vertical directions.

The most widely used closure for eq. (1) is based on the gradient transport hypothesis which, in analogy to molecular diffusion, assumes that turbulence causes a net flux of material down the gradient of material concentration at a rate which is proportional to the magnitude of the gradient [10]:

$$\overline{u'c'} = -K_x\frac{\partial c}{\partial x}, \qquad \overline{v'c'} = -K_y\frac{\partial c}{\partial y} \qquad \text{and} \qquad \overline{w'c'} = -K_z\frac{\partial c}{\partial z}\,, \tag{2}$$

where K_x, K_y and K_z are the eddy diffusivities.

The simplicity of the K-theory of turbulent diffusion has led to its widespread use as the mathematical basis for simulating urban, photochemical pollution. However, K-closure has its own limits.

The gradient-transfer theory works well when the dimension of dispersed material is much larger than the size of turbulent eddies involved in the diffusion process, *i.e.* for ground-level emissions and for large travel times [11]. A further problem is that the down-gradient transport hypothesis is inconsistent with observed features of turbulent diffusion in the upper portion of the mixed layer, where countergradient material fluxes are known to occur [12].

Despite these well-known limits, the K-closure is widely used in several atmospheric conditions, because it describes the diffusive transport in a Eulerian framework, where almost all measurements are of Eulerian type. It produces results that agree with experimental data as well as any more complex model, and it is not as computationally expensive as higher-order closures.

The reliability of the K-approach strongly depends on the eddy diffusivity parametrization. This parameterization can be determined on the basis of the turbulence structure of the Planetary Boundary layer (PBL), and on the model's ability to reproduce experimental diffusion data. A great variety of the K-turbulent profile formulations exist [13-16] and can be used for the parameterization of the K-turbulent profile.

2˙3. *Gaussian models.* – The Gaussian approach is widely used in operative air pollution studies to model the statistical properties of the concentration of contaminants emitted in the PBL [17].

The conditions under which the mean concentration of a pollutant species emitted from a point source can be assumed to have a Gaussian distribution are highly idealised, since they require stationary and homogeneous turbulence. In the PBL the flow may be assumed quasi-stationary for suitably short periods of time (ca. 10 min to 1 h). However, due to the presence of the surface, there are variations with height of both the mean wind and turbulence that cannot always be disregarded.

Much effort has been devoted to the development of non-Gaussian models, for handling the non-homogeneous structures of PBL turbulence. However, they still result in excessively large computer runs, either for emergency response applications or for calculating concentration time series over a long time (*e.g.*, a year). The latter are especially important in the evaluation of violations of air pollution standards, which are often expressed in high percentiles.

Conversely, Gaussian models are fast, simple, do not require complex meteorological input, and describe the diffusive transport in an Eulerian framework, making easy use of the Eulerian nature of measurements.

In practice, they can be relatively easy to use, and can be applied in numerous conditions (*e.g.*, isolated sources, cities, road traffic, complex terrain, etc.).

For these reasons they are still widely employed for regulatory applications by environmental agencies all over the world. Nonetheless, because of their well-known intrinsic limits, the reliability of a Gaussian model strongly depends on the way the dispersion parameters are determined on the basis of the turbulence structure of the PBL and the model's ability to reproduce experimental diffusion data. A great variety of formulations exist [18-23].

The Gaussian model can be modified so as to extend its applicability to non-stationary and non-homogeneous conditions, as well as to more complex orography. In particular, the breaking down of the plume into puffs has permitted the simulation of pollutant dispersion in pseudo-stationary conditions. Such models decompose the plume into puffs, whose characteristics evolve in time and space together with the changing meteorological and emission conditions [24].

Several schemas exist for the calculation of σ_y and σ_z as functions of stability classes and of the downwind distance from the source.

Stability classes can in fact be calculated with semi-empirical techniques using, for example, the method of Pasquill [25] based on simple meteorological observations such as wind velocity, solar radiation and, at night, cloud cover. Other techniques adopt measurements of the standard deviations of vertical wind velocity σ_w, horizontal wind direction σ_θ, the vertical gradient of temperature $\Delta T/\Delta z$, and the Richardson number.

Once the stability classes have been evaluated, the sigmas are expressed as a function of downwind distance x, using one of the many formulae available in the literature, retrieved from experimental campaigns.

The sigmas can be expressed as functions of continuous variables of atmospheric turbulence. Most of them are based on the approach proposed by Pasquill [11]. There exist schemes based on the similarity theory, their related time scale, and the micrometeorological variables. Up-to-date expressions of the sigmas, in terms of wind variance and the Lagrangian integral scale time, on the basis of an atmospheric turbulence spectra model, are presented in [26].

New operative models that describe pollutant diffusion using as input ground-based meteorological data (which can be acquired by automated stations), but which are able to evaluate directly atmospheric turbulence, through the value of the Monin-Obukhov length and the friction velocity, rather than empirical classes [27].

Among the new generation models, of particular note are: the Danish model OML [21] and British model ADMS [28]; the American model HPDM [29], and new model AERMOD, proposed by the US EPA. Among the puff models are CALPUFF [30], M4PUFF [31] and SPM [32].

2˙4. *Model performances and reliability.* – It must be kept in mind, when using air dispersion models, that, while they are rather sophisticated instruments that ultimately reflect the current state of knowledge on turbulent transport in the atmosphere, the results they provide are subject to a considerable margin of error. Models, in fact, provide values expressed as an average, *i.e.* a mean value obtained by the repeated performance of many experiments in the same meteorological scenario, while the measured concentrations are a single value of the sample to which the ensemble average provided by models refer. This is a general characteristic of the theory of atmospheric turbulence and is the ground for the statistical approach such as the Reynolds used in attempting to parameterize the chaotic character of the measured data. At the same time, the uncertainty linked to the stochastic character of the parameterization of the atmosphere depends on turbulence intensity and is a function of the mean sampling time. Atmospheric diffusion models present errors that can be considered as an uncertainty inherent in the phenomenon they describe. Moreover, errors originate from the use of an incorrect or insufficient set of input data and/or from the intrinsic inadequacies of the particular model. Studies of model performance validation indicate errors in input data (both of emission and meteorology) to be the factor responsible for the greater contribution of the total uncertainty of models. Irwin [33] using Monte Carlo techniques to simulate the propagation of errors from those of input data, showed that the interval of error of the concentration maximum and of its distance from the source may be double the interval of error of the input data. A model is generally deemed acceptable if the estimated values are within a factor of two of the observed data.

Further errors can be due to the different spatial reproduction of the model respect to the resolution of the monitoring stations. Often, where the model calculates pollutants on $1\,\mathrm{km}^2$, the observed concentrations can be significant on $200\,\mathrm{m}^2$, if data are referred to primary pollutants, or $10\,\mathrm{km}^2$ or more if secondary. At this time, this aspect is to be solved.

2˙5. *Considerations on environmental data set.* – In order to optimise the use of NN to forecast pollutant in complex situations, some preliminary considerations on the meaning of environmental data set have to be done.

The values of the pollutants levels depend on various variables:

– The Monin-Obukhov length: L.

– The speeds measured and estimated to different z levels: $U(z)$.

– The turbulent eddy diffusion coefficient: K.

– The friction velocity: U^*.

– The PBL height: h.

– The height of the source emissions: Hs.

– The plume temperature: T_f.

– The source emissions rate: Q_E.

Therefore, the concentration can be computed by a mathematical formulation of the type

$$C(x, y, x; L; U(z), K, U^*, Hs, T_f, Q_E), \tag{3}$$

where x, y, and z are the coordinates of the receptor.

The main variables (such as L and U^*) depend on other meteorological variables and these last come from different mathematical formulations that depend on the atmospheric stability class and on the ground typology (*i.e.* roughness length Z_0) and emission sources factors.

The Gaussian plume models work well only in stationary or quasi-stationary conditions and they can be used only in simple homogenous situations.

The K models are more sophisticated models respect to the Gaussian ones, but their results are highly dependent on turbulence profile, that can be determined by theoretical parameterization.

The forecast coming from the urban dispersion models need additional consideration. In the case of the urban pollution the necessity exists to develop the calculation at different spatial scales in order to obtain satisfactory results. The models need to take into consideration the resolution starting from few meters to hundreds of kilometres.

2˙6. *Description and selection of the input variables to NN model.* – The above considerations on the environmental data set are fundamentals for the best selection of the input variables to be assumed when using NN.

Artificial Neural Networks (ANN) are a class of computational tools able to model non-linear dependences characterising the data and to learn complex relationships directly from the data during training phase.

One particular ANN architecture, especially adapted for forecasting tasks, is known as the Multi Layer Perceptron (MLP) with an error supervised learning rule [34]. This net architecture is able to reproduce non-linear models, by means of an accurate choice of the variables of the system and of the meaningful patterns.

In our work, as architecture we used a 3-layer perceptron model. 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 the hidden layer. The third layer is the output layer, which consists of the target variable (*e.g.*, pollutants concentration) to be reproduced.

Usually, NN is used as statistical model, in which the input variables regard strictly the system phenomenology.

The input variables have to satisfy the following task:

– They are related with the output.

– They play the role of independent variables respect to the output one.

– They need to train the NN; in particular, all the weights of the connections layers are coupled with input-output variables.

In the case of pollutants concentrations forecasting, conventional use of NN could involve, as input parameters, the variables linked to the dispersion (as L, U^*, h_{PBL}, Q_E, x, y, z, etc.). This means that the target concentration C_{NN} is a function of many parameters:

$$C_{\mathrm{NN}} = f(x, y, x; L; U(z), K, U^*, Hs, T_f, Q_E). \tag{4}$$

Our methodology consists of including the results coming from the dispersion models as input variables.

So doing, we facilitate the task of NN to understand the relations between input NN variables and the target one:

$$C_{\mathrm{NN}} = f(C_{\mathrm{ADM}}, \text{other parameters}). \tag{5}$$

In such a way, we combine two model approaches: statistical and deterministic models.

Mostly, deterministic and statistical methods are two well-separated techniques using to work for different tasks.

Deterministic models can be applied when the system is well known, that means all the variables and parameters are fixed and mutually correlated.

Statistical models are applied when mathematical formulations have been fixed and they utilize all the empirical variables of the system.

The inclusion of the dispersion model predicted concentrations (C_{ADMD}) as input values of the network means that it must perform a twofold task. The first is to train the NN starting from a situation close to the actual such as provided by a dispersion model, which already incorporates emission and turbulence factors. The second task is linked to the fact that dispersion models perform well under certain hypotheses, while tending systematically to decrease their performance when reality is far from the ideal situations.

In this case, the NN functions as a filter of the model allowing to improve the reproduction of the real situation.

In substitution of the main variables linked to the turbulence and to the dispersion, we consider a subset of seven variables, all connected to the spatial resolution, to the dispersion and advection of pollutants. The variables are more easy to be measured and they could synthesise the behaviour of the actual variables of the system (meteorological, turbulence, source emissions, etc.).

2˙7. *Data pre-elaboration.* – In our work, the choice of the best input data to be provided to the NN model is made empirically through various trials (30 tests). As evident in table I, at each trial the input net is modified choosing a maximum of 7 variables of interest, and the correlation coefficient between simulated and measured variable target and the presence of negative predicted concentrations has been calculated.

The 7 input variables concern the spatial resolution of the dispersion model (Distance), the main meteorological information related to the turbulence (Z_i, U^*, Z/L) and the conventional surface data (W, T).

At this step, we neglected other variables, such as humidity, air pressure and wind direction because, as input data to NN, we considered the important variables C_{ADMD}, for which the influence of all environmental factors on air dispersion could be included.

TABLE I. – *Choice of NN parameters and of the main variables.*

	Distance	CONC(ADMD)	Z_i	U^*	Z/L	Wspeed	T	C negative	R
Test 1	X							0	0.475
Test 2		X						0	0.565
Test 3			X					0	0.417
Test 4				X				0	0.597
Test 5					X			0	0.566
Test 6						X		0	0.615
Test 7							X	0	0.424
Test 8	X	X						0	0.753
Test 9	X		X					1	0.593
Test 10	X			X				0	0.819
Test 11	X		X	X	X			2	0.894
Test 12	X		X	X	X	X		4	0.891
Test 13	X		X	X	X	X	X	4	0.931
Test 14	X		X		X			1	0.844
Test 15	X			X		X		3	0.876
Test 16	X					X	X	0	0.875
Test 17	X			X	X	X		4	0.889
Test 18	X			X			X	0	0.868
Test 19	X				X	X		2	0.893
Test 20	X	X	X					0	0.752
Test 21	X	X		X				0	0.897
Test 22	X	X	X	X	X			3	0.914
Test 23	X	X	X	X	X	X		7	0.918
Test 24	X	X	X	X	X	X	X	1	0.941
Test 25	X	X	X		X			0	0.909
Test 26	X	X		X		X		2	0.897
Test 27	X	X				X	X	2	0.916
Test 28	X	X		X	X	X		1	0.902
Test 29	X	X		X			X	2	0.903
Test 30	X	X			X	X		1	0.893

The purpose of this pre-elaboration regards an optimization procedure of NN applied to the choice of input variables. This procedure implies the minimization of input data without loss of information concerning actual plume dispersion.

At this step, the role of the NN inputs is different respect to their conventional use.

Aim of our pre-elaborations is to find those important variables necessary to the adjustment of the variable C_{ADMD} calculated by a deterministic model to reproduce the pollutants levels.

We consider 100% of data during the training and correlations coefficients are related to all data set. At this stage, our aim is not to find the best performance of NN, but to give an indicator of input variables to NN.

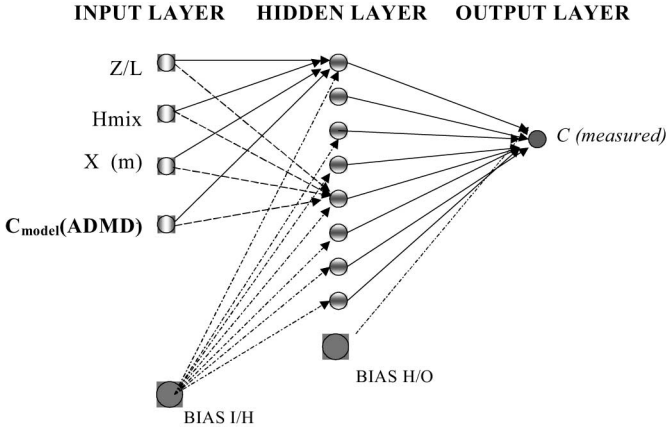

Fig. 1. – Architecture used in the simulation for the ADMD-NN model.

Usually, people choose the NN input variables in order to connect the input/output and to try to simulate a mathematical relation between themself.

Our task consists of selecting lesser but significant input variables as possible (we apply the parsimony principle [35]).

Table I shows also some sensitivity studies (Test 1–Test 7) concerning the influence of every variable standing alone to forecast the observed pollutants levels.

The maximum correlation is verified for the wind speed variables ($R = 0.615$).

To keep away from our analysis the influence of important variables (that could create spurious correlation), we neglect as input data those with high and low correlation (Test 6, $R = 0.615$ and Test 7, $R = 0.424$). Such choice is coherent with the adjusting of C_{ADMD} as mentioned above.

In our simulation, by table I, the correlation result (Test 25, $R = 0.909$) seems to suggest the followings 4 variables as NN input: the mixing height (H_{mix}), the Monin-Obukhov Length (Z/L), the downwind distance Source-Receptor (X) and, the most important ones, the concentration levels predicted by ADMD dispersion model (C_{ADMD}).

To select the relevant input data for the NN model, a compromise must be obtained between the smallest number of input variables and the best simulation performance of the NN model.

The perceptron model used to forecast the pollutants levels is made up of 3-layer architecture with 4 neurons in the input layer, 8 in the hidden layer and one output neuron (containing the concentration levels to be reproduced).

The choice of the hidden layer was investigated by examining 6, 8 and 12 neurons. The best performances were obtained with 8 hidden neurons.

The our final architecture consists of a 4-8-1 type net (fig. 1). The choice of 8 hidden neurons is based on two opposite considerations: maximizing the hidden neurons to increment the NN parameters and minimizing this number in relation with the main situations linked to the input patterns.

The NN input variables (Z/L, H_{mix} and X) identify various turbulent regimes and different transport-diffusion scenarios where the air pollution model could present different behaviours.

Further, it is important to underline that our NN model uses only four variables in order to forecast the pollutant concentrations in complex situations such as a high-density urban area.

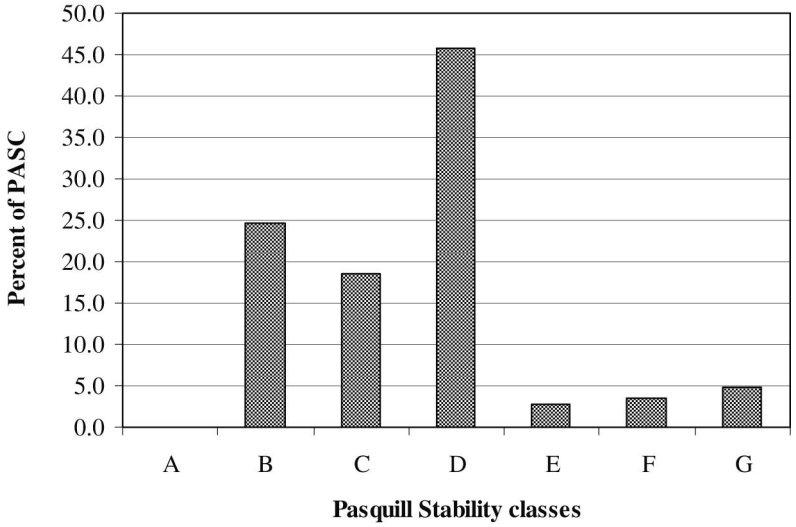

Fig. 2. – Distribution of Pasquill atmospheric stability classes.

As marked in the paragraph above, the conventional use of NN should involve the inclusion of many other input variables (such as all the meteorological factors, the source emission, the position of all the main sources and receptors, the vertical structure of turbulence parameters) to obtain the same information included in the C_{ADMD}. The pollutant concentration (calculated by dispersion model) takes into account all effects on the plume dispersion by other inputs variables (*i.e.* ground effects, meteorology and sources characteristic).

The variables C_{ADMD}, used as input to NN, optimize the environmental information to the net so that during the training only the differences between the deterministic model results and actual situations should be apprehended. In our simulations, we choose to use results coming from dispersion models. It could be considered as input variable the difference between the observed and the calculated pollutants. The NN do not make difference between the two choices. In our simulation we used 371 selected patterns, each determined by four input variables at some downwind distance and turbulence conditions.

2˙8. *Predicted concentrations of ADMD against the Indianapolis data set.* – The model ADMD has been validated through the results of the experiments conducted in an urban area of Indianapolis (Indiana, USA) in the months of September and October 1985 (EPRI [36]). During these experiments a SF6 tracer has been released at the height of 83.8 meters from a source, situated inside the urban area of Indianapolis. The concentrations have been monitored by ground stations, by a system of 160 receptors situated on arcs to 12 different distances from the source: 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, 8, 10 and 12 Km.

The Pasquill turbulence parameters observed during the field campaign at the site were given in fig. 2. The more representative class was the neutral one (D), followed by the instability classes (B-C). The stable conditions were measured very few times. The net prevalence of instability classes is the sign of the presence of heat island effect, typical of high-density urban area.

The roughness has been valued $Z_0 = 1\,\mathrm{m}$. The concentration measurements have been integrated in the cross wind direction. Table II shows the results coming from

TABLE II. – *Statistical indices related to the application of ADMD.*

	NMSE	FA2	R	FB
ADMD	1.26	0.52	0.33	0.52

a comparison of the concentrations observed and those calculated by ADMD considering all the stability classes and using the statistical indices described by Hanna [37].

By the examination of the statistic indexes, it is evident that the model ADMD presents poor performances (as expected) in simulating the releases of pollutant in the urban area of Indianapolis.

In fig. 3 we show the results obtained by the application of ADMD. The systematic under-prediction of the measured pollutant levels is evident. While the concentrations predicted by ADMD are within 1000 μg/mc, the measured concentrations range up to 6000 μg/mc. This means that ADMD model in the urban case does not succeed to calculate the right maximum plume impact in some turbulence conditions. When the observed concentration is in the range 1000–3000 μg/mc the model presents good performances. Taking into consideration the percent errors between predicted and observed ones, the error distribution presents a strong under-prediction (average of 27%—fig. 4). This behaviour is systematic (as evident by the skewness of the error distribution) and the work of the neural net could be concerned to adjust the performance of the air dispersion model. This is another crucial point to the elaboration.

The adjustment of our dispersion model by NN filter has meaning only if the model presents a systematic distortion (under or over prediction). In this case, our approach can be utilized to adjust those situations that deterministic models do not succeed to reproduce.

If ADMD did not present systematic errors, we should observe an error distribution similar to the Gaussian shape.

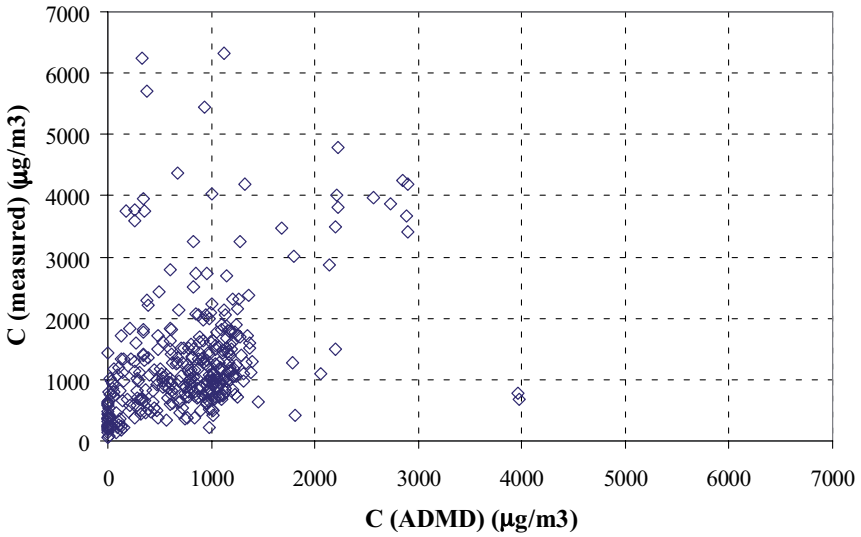

Fig. 3. – Comparison of CADMD model with measured pollutant levels.

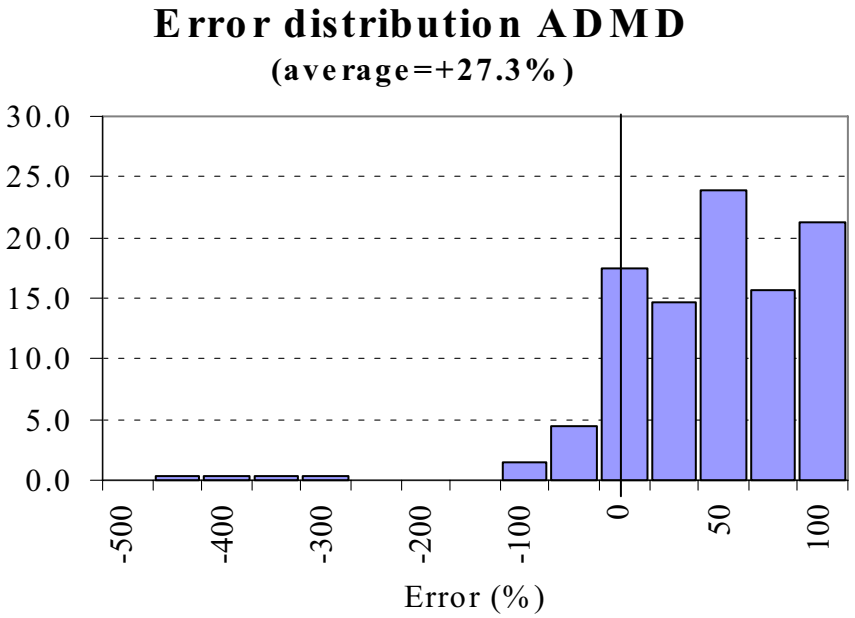

Fig. 4. – Error distribution using the ADMD model.

3. – Results and discussion

3˙1. *Validation results*. – To conduct the net training phase we have selected randomly the 50% of the data. The remaining 50% have been used for the generalization phase (where the data are independent and different from those used for the training). This choice should assure to avoid problems with the over-fitting (usually people use more than 50% during the training). The results always all referred to generalisation phase.

We use during the training phase the conventional activation functions

$$F(P) = \frac{1}{1 + e^{-(P-S)}}\,, \tag{6}$$

where P is the activation potential related to neurons of the different layers, S is the threshold related to the inner neurons layer and calculated by the bias neurons.

The NN weights correction was performed using two methods: the conjugate gradient and the popular back propagation algorithms.

The best NN results were obtained with the conjugate gradient method [34]. Output results were obtained using 30000 epochs during training phase and we operate with the *batch* corrections of weights.

In the environmental study, some typical statistical descriptors are used in order to evaluate models performances.

In our work we utilize the conventional descriptors suggest by Hanna [37]. Applying the NN model and calculating all the main statistical indexes, we obtain the results as in table III. To check the goodness of NN during training phase, we show also the indexes at this step. The table gives results for the pollutant calculated by ADMD and filtered by NN (ADMD + NN) in relation with the training and test phases.

TABLE III. – *Statistical indices related to application integrated model (ADMD+NN) for train and test phase.*

	NMSE	FA2	R	FB
ADMD+NN(Train)	0.47	0.61	0.88	−0.52
ADMD+NN(Test)	0.38	0.68	0.77	−0.41

The improvement obtained by our methodology could be deduced by the comparison of table II and table III.

During the training phase a good correlation coefficient was obtained ($R = 0.88$), while for the generalization a value of $R = 0.77$ was achieved. In fig. 5 the results coming from the generalisation are shown. It is important to underline that this NN model presents a good reproduction for all pollutants range. The FB indicates a low overestimation ($FB = -0.41$) of pollutants when NN is upstream to the dispersion model. Statistical indexes calculated by the air dispersion model ADMD alone (table II), show that the model does not succeed to reproduce adequately the pollutant levels and the inclusion of NN improves all these parameters.

When we calculate the error between the forecasting and measured levels, the error distribution (fig. 6) shows a net improvement respect to the error distribution obtained with ADMD model. A decrease of skweness can also be noticed respect to the ADMD alone.

The application of NN as a filter of deterministic models appears to have to be successful for these applications. In fact, the combined approach adequately reproduces the pollutant levels, despite the urban system can be considered as complex situations. Besides, to evaluate if the input variable C_{ADMD} was useful to improve the accuracy of results, we have run the same net without the concentration derived by the dispersion

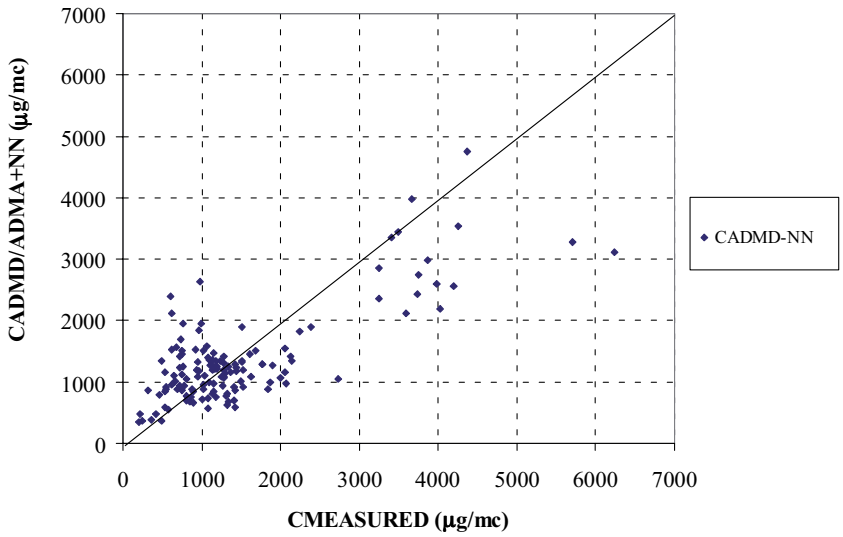

Fig. 5. – Comparison of CADMD-NN with measured pollutant levels.

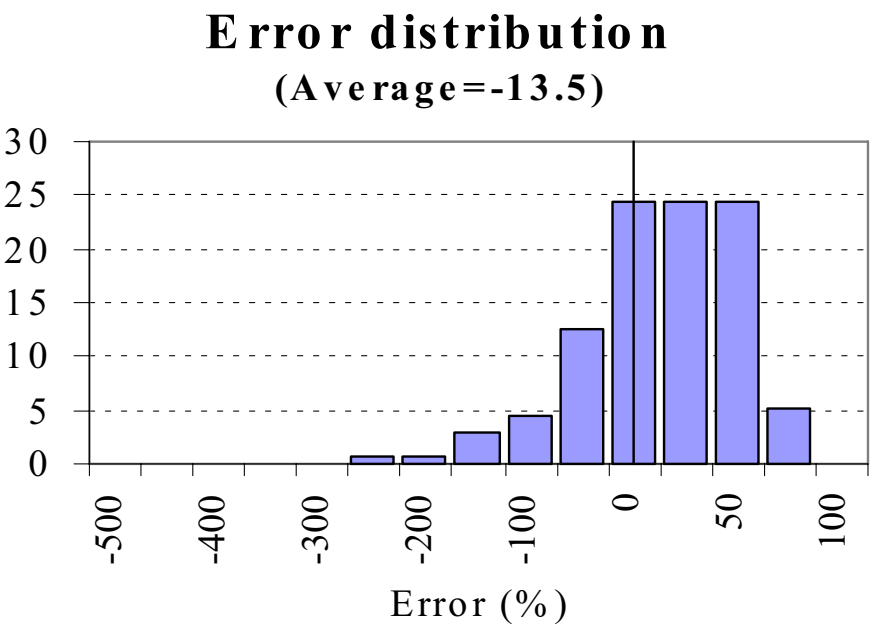

Fig. 6. – Error distribution coming from ADMD-NN model.

model (C_{ADMD}). In the generalization phase a correlation coefficient of $R = 0.68$ was found (to be compared with $R = 0.77$ obtained with the combined approach). Therefore the choice of using the concentrations predicted by the deterministic model (C_{ADMD}) is discriminating for the quality of results.

3˙2. *Plume impact evaluation study.* – While the results coming from deterministic models are easily to link with the effects on the plume by atmospheric dispersion, in general the NN are not necessarily connected with the physics of the system.

The question concerns if the NN results have some explanation by means of typical turbulence parameters.

As far as the physics related to the atmospheric dispersion is concerned, in the case of tall sources, it is recognized that: the concentration has to decrease with the distance; models have to reproduce for the pollution levels the same trend with the atmospheric stability classes (the pollution decreases from unstable to stable turbulence conditions); the distance of the maximum concentrations increases with the stability classes (from A to F) when we applied the deterministic models.

In order to verify the above features we run again new simulations fixing the internal neural net parameters as derived by the ADMD + NN model. We have considered in the simulations, as NN input, the Monin-Obukhov length, linked to the Pasquill Stability Classes (PSC) and three values of the mixing height ($Z_i = 400\,\mathrm{m}$, $1000\,\mathrm{m}$, $2000\,\mathrm{m}$).

In this case as air pollution model, in order to estimate the ground maximum concentration ($C_{\max}$), we used the well-known EPA ISCST model [38, 39]. The ground maximum concentrations are more easy to calculate from ISCST model respect to ADMD. ISCST is a reference model used worldwide for impact assessment applications. Conversely, ADMD is an advanced research model which uses turbulence parameters (such as the K-profiles) which are often not available for conventional applications. At this

TABLE IV. – *Value of the maximum pollutant levels ($C_{\max}$) and the maximum distance ($D_{\max}$) as calculated by a short-term Gaussian model (ISCST).*

$1/L$ (m^{-1})	PSC	$D_{\max}$ (ISCST) (m)	R_D (ISCST)	$C_{\max}$ (ISCST) ($\mu g/m^3$)	R_C (ISCST)
−0.0875	A	700	1.00	2.64	1.00
−0.0389	B	1500	2.14	1.64	0.62
−0.0081	C	2900	4.14	1.29	0.49
0.0000	D	11300	16.14	0.58	0.22
0.0081	E	33000	47.14	0.21	0.08
0.0389	F	183000	261.43	0.06	0.02

step, we need to calculate the behaviour of the plume impact at the ground as obtained by an air dispersion model easy to run. For the above reasons, ISCST was selected as a comparative model to evaluate the maximum plume impact.

The ratio distance (R_D) and concentration (R_C) are adimensional fractions, defined by us, between the maximum distance ($D_{\max}$) or concentration ($C_{\max}$) corresponding to A stability classes and all the others (from B to F). As is recognized, for the Gaussian model the levels $C_{\max}$ decrease with the stability classes (from A to F) and the impact distance $D_{\max}$ quickly grows.

For all simulations (ADMD-NN and ISCST), we have calculated the maximum pollutant levels ($C_{\max}$) and the related maximum distance from the emission ($D_{\max}$). The results for the typical Gaussian models (ISCST) are shown in table IV.

We can observe the rapid increase of the maximum impact distance ($D_{\max}$) with stability classes from A to F and the corresponding decrease of the ground pollutants. Taking as reference the values 700 m and 2.64 $\mu g/m^3$ and normalizing respect to these values (corresponding to A stability classes), we calculated the adimensional R_D and R_C. These values assumed for all the stability classes are typical of the dispersion models and reflect that plume impact increases with stability turbulence.

Once fixed the neural-network parameters, a NN simulation has been made to estimate the position of the maximum concentration $D_{\max}$ using three values of the mixing height (400–1000 m and 2000 m) in the input net variables.

In order to calculate the Monin-Obukhov length (as input for NN) we used the values from Golder diagram [40] that relates the Pasquill stability classes to the Monin-Obukhov length (MO). In this way, the typical MO length linked to the main Pasquill stability classes (from A to F) was obtained.

Table V shows relations between Pasquill-Gilford and the MO length used by us in the simulations as NN input (take the roughness length of 1 meter).

We simulate for all the Pasquill stability classes and for all upwind distances the ground impact as coming from the ADMD-NN model.

Further it is important to underline that these simulations concern data never sighted by NN (e.q. we investigate the NN performance in connection with the extrapolated data).

TABLE V. – *Value of the maximum pollutant levels ($C_{\max}$) and the maximum distance ($D_{\max}$) as calculated by ADMD+NN model.*

PSC	H_{mix} (ADMD) (m)	$D_{\max}$ (ADMD) (km)	R_{D} (ADMD)	$C_{\max}$ (ADMD) (μg/m^3)	R_{C} (ADMD)
A	400	7.031	1.00	7038.1	1.00
B	400	6.24	0.89	7323.7	1.04
C	400	6.635	0.94	4244	0.60
D	400	7.031	1.00	3876.6	0.55
E	400	7.031	1.00	3601.5	0.51
F	400	5.052	0.72	3360.5	0.48
A	1000	12.302	1.00	6637.4	1.00
B	1000	6.635	0.54	5999.1	0.90
C	1000	7.427	0.60	1839.3	0.28
D	1000	3.073	0.25	1587.9	0.24
E	1000	3.073	0.25	1639.2	0.25
F	1000	3.073	0.25	2062	0.31
A	2000	4.656	1.00	5713.4	1.00
B	2000	5.052	1.09	5166.1	0.90
C	2000	2.281	0.49	1098.9	0.19
D	2000	1.517	0.33	1009.4	0.18
E	2000	1.517	0.33	1058.4	0.19
F	2000	2.001	0.43	1468	0.26

The results for 1000 m mixing height are given in fig. 7 and for the Paquill-Gillford stability classes.

The ground concentrations levels predicted by NN resulted similar to the classic air dispersion model (with a well-defined maximum ground impact for different atmospheric stability).

The reproducibility of the shape ground distribution similar to the classic one is an important result and absolutely not an obvious consequence. In fact, the NN are non-linear mathematical models and when we change the internal NN parameters, the shape obtained in fig. 7 is sharply modified.

Behaviour with a well-defined maximum plume impact was generally not observed with training phase different from the chosen one. Mostly, the trend of NN increases with the upwind distance, coherently with the non-linear mathematical model inside the NN.

The $D_{\max}$ results obtained with the ADMD + NN are shown in table V.

In comparison to the Gaussian dispersion model (ISCST), the model ADMD + NN produces the same decrement of the pollutant concentrations with the stability classes (compare R_{C}(ISCST) and R_{C}(ADMD+NN)). Furthermore, when the mixing layer is low (400 m), the concentration levels are higher, while the calculated levels when the mixing layer is higher (1000–2000 m) are close to those derived from the Gaussian dispersion model (as evident by the RC trend).

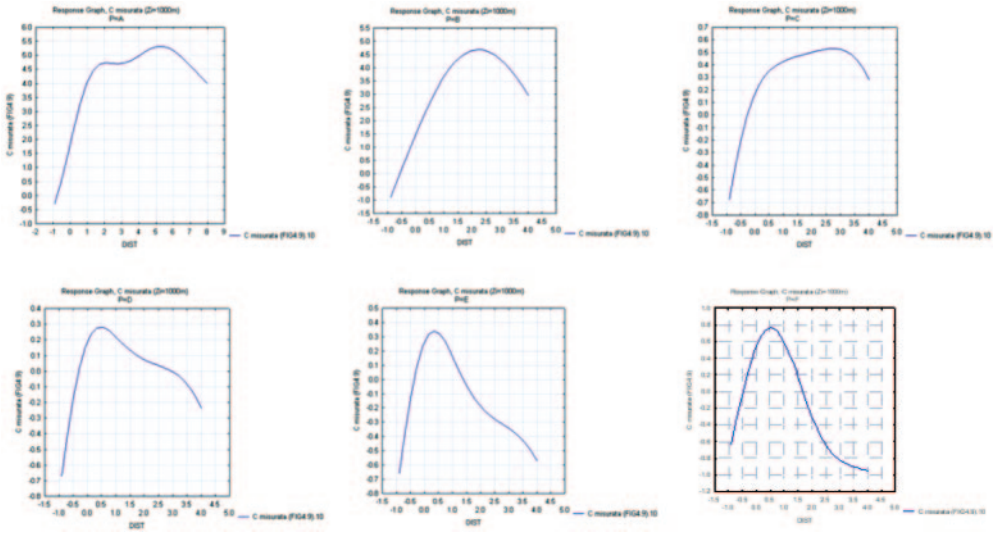

Fig. 7. – Concentrations calculated by ADMD-NN model (fixing $Z_i = 1000\,\mathrm{m}$) at different stability classes.

The results obtained for the $D_{\max}$ are very interesting (as evident by the ratio $R_{\mathrm{D}}(\mathrm{ADMD} + \mathrm{NN})$). We found that the simulated maximum distance decreases with the atmospheric stability (from A to F), in contrast with the behaviour of Gaussian model (see $R_{\mathrm{D}}(\mathrm{ISCST})$). This result is attractive, because it could be the main reason of unsatisfactory bad performances of Gaussian models in urban case.

Figures 8 and 9 synthesize the results as coming from tables IV and V. The figures show the ratio R_{C} among the maximum concentrations and the ratio R_{D} among the maximum distances of impact for different atmospheric stability.

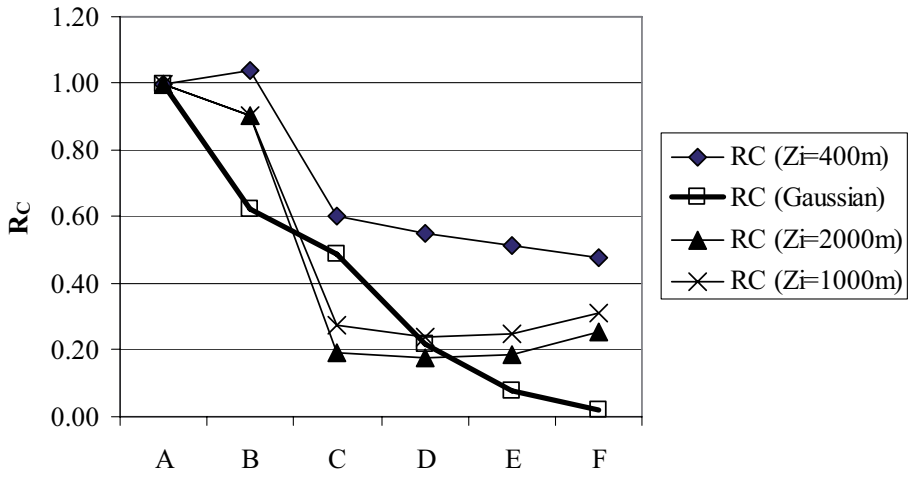

Fig. 8. – Trend corresponding to the stability classes of the adimensional fraction between the maximum concentration (R_{C}).

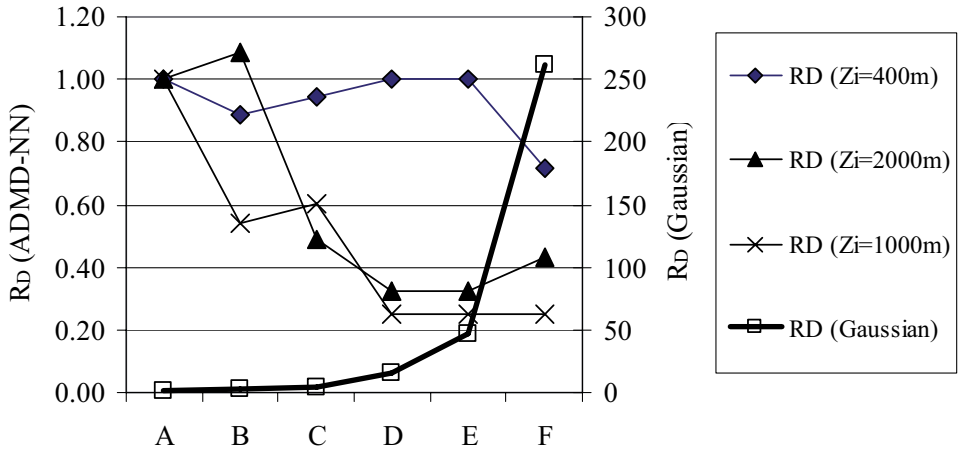

Fig. 9. – Trend corresponding to the stability classes of the adimensional fraction between the maximum distance (R_{D}).

Respect to the Gaussian dispersion model, the model ADMD-NN produces the same decrement of the pollutant relationship with the downwind distance (fig. 8). When the mixing layer is low (400 m), the ratios R_{C} are higher with respect to those calculated for 1000 m and 2000 m. When the mixing layer is higher (1000–2000 m) they are close to those derived from the Gaussian dispersion model. However, the trend for R_{C}, derived by ADMD-NN model, is similar to the Gaussian models (they decrease with the atmospheric stability).

As evident in fig. 9, the maximum concentration distance R_{D} simulated by the NN decreases with the atmospheric stability (from A to F), in contrast with the behaviour of Gaussian model. This fact means that ADMD-NN model succeeds to correct the performance of the deterministic model and suggests which could be the physical correction to be apportioned to the classical dispersion model.

The urban situation would involve an increasing of the concentration levels during stable conditions in comparison to the simple Gaussian model. At the moment, it is very difficult for a deterministic model to reproduce a similar trend for RD. The observed effects could be explained with the decrease of mean wind speed and the increase of surface drag due to the buildings height and the following wake turbulence.

4. – Conclusions

We have applied for a complex urban situation a mixed model composed by a deterministic model and a NN network. The mixed model, validated using the urban Indianapolis dataset, shows good performances. However, in the literature there are NN applications in forecasting residual values (between forecast and measured values) instead of real values [41].

A net improving of all main statistical indexes is observed, decreasing the error between the calculated values and the measured ones and we reproduce the same decrement, predicted by Gaussian models, of the pollutant ratio R_{C} with the stability classes. By using the ADMD + NN model, we obtained for the ratio R_{D} a trend in contrast with the Gaussian model.

The urban profile seems to involve a decrease of the plume impact during stable conditions in comparison to the simple Gaussian model.

Our study seems to suggest that the neural network can be considered as an alternative model approach to calculate the pollutant levels, most especially in complex situations.

The more evident advantage in the use of NN resides in building a model starting from direct variables related to the phenomenology and to the non-linear contributions to air pollution dispersion.

In the classic use of NN, the target of the model concerns to link input variables with the output one. The input variables are determined by measurements and they can be completely different by the ones used in the deterministic models.

In our approach, we tested as input variables of NN some typical results coming from the application of the deterministic model or air dispersion model.

The use of an integrated model seems to suggest the direction to follow to improve the performances of deterministic models in complex areas as urban ones. The results of our methodology seem to suggest that (in urban cases) it could be necessary to find a model in which the ground plume impact falls more close to the source emission during the stable turbulence scenario respect to the instable one. The concentrations levels predicted by the dispersion model could carry on the same trend with the atmospheric stability.

Further applications on NN to real urban situations need be done to verify the goodness of our approach and validate the physical inference derived by our simulation.

REFERENCES

[1] Boznar M., Lesjak M. and Mlakar P., *Atmos. Environ. B*, **27** (1993) 221.
[2] Comrie R. S., *J. Air Waste Manag. Assoc.*, **47** (1997) 653.
[3] Gardner M. W. and Dorling S. R., *Atmos. Environ.*, **32** (1998) 2627.
[4] Kolehmainen M., Martikainen H. and Ruuskanen J., *Atmos. Environ.*, **35** (2001) 815.
[5] Kolehmainen M., Martikainen H., Hiltunen T. and Ruuskanen J, *Environ. Monit. Assess.*, **65** (2000) 277.
[6] Pasini A. and Pelino V., *Can we estimate atmospheric predictability by performance of neural network forecasting? The toy case studies of unforced and forced lorenz models*, in *Computational Intelligence for Measurement Systems and Applications, 2005. CIMSA. IEEE International Conference on. (July 2005)* (IEEE) 2005, pp. 69-74, doi:10.1109/CIMSA.2005.1522829.
[7] Pelliccioni A. and Tirabassi T., *Environ. Model. Softw.*, **21** (2006) 539.
[8] Amicarelli A., Gariazzo C., Finardi S., Pelliccioni A. and Silibello C., *Evaluation of data assimilation techniques for a mesoscale meteorological model and their effects on air quality model results*, in *14th International Symposium for the Advancement of Boundary Layer Remote Sensing, Riso National Laboratory, DTU, Denmark, 23-25 June 2008*, *IOP Conf. Series: Earth and Environmental Science*, **1** (2008) 012033 (IOP Publishing), doi:10.1088/1755-1307/1/1/012034.
[9] Finlayson-Pitts B. J. and Pitts J. N. jr., *Atmospheric Chemistry: Fundamentals and Experimental Techniques* (John Wiley & Sons) 1986 (ISBN 0-471-88227-5).
[10] Seinfeld J. H. and Pandis S. N., *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (J. Wiley, New York) 1998.
[11] Pasquill F. and Smith F. B., *Atmospheric Diffusion*, third edition (Halsted Press, John Wiley and Sons, New York) 1983.
[12] Deardoff J. W. and Willis G. E., *J. Appl. Meteorol.*, **14** (1975) 1451.
[13] Ulke A. G., *Atmos. Environ.*, **34** (2000) 1029.

[14] Degrazia G. A. and Moraes O. L. L., *Boundary-Layer Meteorol.*, **58** (1992) 205.
[15] Degrazia G. A., Rizza U., Mangia C. and Tirabassi T., *Boundary-Layer Meteorol.*, **85** (1997) 243.
[16] Degrazia G. A., Anfossi D., Carvalho J. C., Mangia C., Tirabassi T. and Campos Velho H. F., *Atmos. Environ.*, **34** (2000) 3575.
[17] Zannetti P., *Air Pollution Modelling* (Computational Mechanics Publications, Southampton (UK) and Van Nostrand Reinhold, New York, USA) 1990.
[18] Hanna S. R., Briggs G. A., Deardoff J., Egan B. A., Gifford F. A. and Pasquill F., *Bull. Am. Meteorol. Soc.*, **58** (1977) 1305.
[19] Briggs G. A., *J. Clim. Appl. Meteorol.*, **24** (1985) 1167.
[20] Hanna S. R., *J. Clim. Appl. Meteorol.*, **25** (1986) 1426.
[21] Berkowicz R. R., Olesen H. R. and Torp U., *The Danish Gaussian air pollution model (OML): description, test and sensivity analysis in view of regulatory applications*, *Proceedings of NATO-CCMS 16th International Meeting on Air Pollution Modelling and its Applications*, edited by De Wispelaere C., Schiermeier F. A. and Gillani N. V. (Plenum Press, New York, N.Y., USA) 1986, pp. 453-481.
[22] Erbrink J. J., *Boundary-Layer Meteorol.*, **74** (1995) 211.
[23] Mohan M. and Siddiqui T. A., *Boundary-Layer Meteorol.*, **84** (1977) 177.
[24] Arya S., *Air Pollution Meteorology and Dispersion* (Oxford University Press, New York) 1999.
[25] Pasquill F., *Q. J. R. Meteorol. Soc.*, **97** (1971) 369.
[26] Mangia C., Rizza U., Degrazia G. A. and Tirabassi T., *A new formulation of σ_y and σ_z for air quality dispersion models*, in *Air Pollution VI*, edited by Brebbia C. A., Ratto C. F. and Power H. (Computational Mechanics Publications, Southampton) 1998, pp. 839-848.
[27] Gifford F. A., *Nucl. Safety*, **2** (1961) 47.
[28] Carruthers D. J., Holroyd R. J., Hunt J. C. R., Weng W. S., Robins A. G., Apsley D. D., Smith F. B., Thomson D. J. and Hudson B., *UK atmospheric dispersion modelling system*, in *Air Pollution Modeling and its Application IX*, edited by Dop H. and Kallos G., *19th NATO/CCMS International Technical Meeting on Air Pollution Modeling and its Application, Creta, Greece, Sept. 29-Oct. 4, 1991* (Plenum Press, New York) 1992, pp. 15-28.
[29] Hanna S. R. and Paine R. J., *J. Appl. Meteorol.*, **28** (1989) 206.
[30] Scire J. S., Robe F. R., Strimaitis M. E. and Yamartino R. J., A Users's Guide for the CALPUFF Dispersion Model. Earth Tech, Inc., Concord, MA, USA (1999).
[31] Tirabassi T. and Rizza U., *J. Appl. Meteorol.*, **36** (1997) 1031.
[32] Tirabassi T. and Rizza U., *J. Appl. Meteorol.*, **34** (1995) 989.
[33] Irwin J. S., Rao S. T., Petersen W. B. and Turner D. B., *Atmos. Environ.*, **21** (1987) 1927.
[34] Rojas R., *Neural Networks: a Systematic Introduction* (Springer-Verlag, Berlin Heidelberg) 1996.
[35] Hansen M. H. and Yu B., *J. Am. Stat. Assoc.*, **96** (2001) 454.
[36] EPRI, Urban Power Plant Plume Studies, Rapport EPRI, EA-5468, Electric Power Research Institute 3412 Hillview Avenue Palo Alto California 94304 USA (1988).
[37] Hanna S. R., *Atmos. Environ.*, **23** (1989) 1385.
[38] Bowers J. F. and Anderson A. J., *An Evaluation Study for the Industrial Source Complex (ISC) Dispersion Model*, EPA-450/4-81-002, US Environmental Protection Agency, Research Triangle Park, NC (January 1981).
[39] Londerga R., Minot D., Wackter D. and Fizz R., *Evaluation of Urban Air Quality Simulation Models*, EPA-450/4-83-020, US Environmental Protection Agency, Research Triangle Park, NC (July 1983).
[40] Golder D., *Boundary-Layer Meteorol.*, **3** (1972) 47.
[41] Pasini A. and Ameli F., *Geophys. Res. Lett.*, **30** (2003) 1386, doi:10.1029/2002GL016726.