PERGAMON

Central and South America

Atmospheric Environment 35 (2001) 1783–1789

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

Prediction of NO and NO_2 concentrations near a street with heavy traffic in Santiago, Chile

Patricio Perez*, Alex Trier

Departamento de Fisica, Universidad de Santiago de Chile, Casilla 307, Correo 2, Santiago, Chile

Received 17 February 2000; received in revised form 18 May 2000; accepted 30 May 2000

Abstract

Based on NO concentrations and meteorological variables recorded hourly at a point close to an avenue with heavy traffic in the city of Santiago, we are able to build a simple model that allows prediction of NO concentrations several hours in advance. Predicted NO concentrations in conjunction with forecasted meteorological data may be used to predict NO_2 concentrations with reasonable accuracy. We compare predictions generated using persistence, linear regressions and multi layer neural networks. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Air pollution prediction; NO and NO_2; Neural networks; Meteorology forecast

1. Introduction

Nitric oxide (NO) is a gas which is a product of motor vehicle combustion. High concentrations of this gas in the atmosphere of a city are usually accompanied by high concentrations of particles in suspension and, in general, with conditions that may be threatening to human health. NO is known to react with hydrocarbons to generate ozone. NO plays also an important role in photochemical reactions which lead to the formation of nitrogen dioxide (NO_2). NO_2 has been shown to produce adverse effects in pulmonary function when inhaled at concentrations of the order of those found in urban air of polluted cities during certain episodes (Speitzer, 1980).

Given the topographical and meteorological conditions in Santiago, Chile, dispersion of pollutants is very slow, especially in winter. We analyze the period May through September which is characterized by fairly extended periods of atmospheric thermal inversion. Continued periods of daytime mixing layer heights as low as 400–600 m lead to severe pollution episodes. Unfortunately, no continued measurements of this parameter are available. Weather frontal passages are infrequent and total rainfall for the period is less than 300 mm. The city is surrounded by mountains, and during the time of the year under study minimum temperatures are between 0 and 5°C and maximum values are between 10 and 20°C. Wind speed in the area has an average value of $1.2\ m\ s^{-1}$. More details on topographical and meteorological conditions in Santiago can be found in Rutllant and Garreaud (1995).

It is very useful then to have reliable methods to predict pollutant concentrations, in particular NO and NO_2, several hours in advance, in order to give the opportunity that the public authorities take emergency actions to protect the population.

Classical statistical methods and neural network methods have been used by several authors for short-term prediction of gas and particulate matter pollution. A good review of applications of neural network techniques in atmospheric sciences with reference to forecasting is given by Gardner and Dorling (1998).

With respect to air pollution prediction in Santiago, Chile, we can mention the study by Rutllant and Garreaud (1995). They correlate a meteorological air pollution potential index (called MAPPI) with a normalized index (called IAC) that integrates in a single variable different air pollutants. MAPPI is modeled as a linear

* Corresponding address.
E-mail address: pperez@lauca.usach.cl (P. Perez).

PII: S1352-2310(00)00288-0

function of temperature and zone wind component, and a 12 h forecast on the high air pollution potential days is made with 73% accuracy. It is worth mentioning a study by Jorquera et al. (1998) where a linear model, a neural network and a fuzzy model are compared as predictors of daily maximum levels of ozone. Perez et al. (2000) have shown that a three-layer neural network may be a useful tool to predict PM2.5 concentrations in the atmosphere of downtown Santiago several hours in advance when hourly concentrations of the previous day are used as input.

The possibility to predict maximum NO_2 levels at the downtown area of Athens, Greece, has been studied by Ziomas et al. (1995). They forecast the possible increase or decrease of NO_2 levels based on discriminant analysis using the information of maximum levels on the previous day, forecasted temperature, wind velocity and wind direction, an index indicating short-term emission variations of the day of interest and an index characterizing the effect of precipitation during the day of interest. Average percentage of successful forecasts was about 80%. They also provide a quantitative estimation of the next day's maximum mean hourly NO_2 concentration performing a multiple linear regression analysis. In this case, percentage of error based on the root mean square error is of the order of 25%.

In this article we report a study on the possibility to predict hourly averages of NO and NO_2 concentrations based on data obtained in a station located at a fixed point in Santiago. This point is close to a thoroughfare which shows a heavy traffic, especially at rush hours, when high emissions of NO are expected. Encouraged by the results obtained with PM2.5, we compared the predictions produced with a feed forward neural network, which uses previous values of NO hourly concentrations as input, and a scheme based on persistence of average conditions (see Section 3 for details). However, in this case prediction by persistence seemed to be a better method. This may be an indication that memory effects due to accumulation of the pollutant is more important for particles than for gases. Then, we intended an improvement of the forecast using values of temperature, relative humidity and wind velocity on the previous day as the only inputs to linear regressions and neural networks. In this situation, predictions of NO concentrations with a linear regression are better than persistence for the first hours on the next day. When considering forecasted values of the meteorological variables for the day under consideration as input, an additional improvement is possible by using a three layer feed forward neural network. Since maximum of NO_2 concentration is most of the time around noon, and it depends on the available NO few hours earlier, it will be very useful to know NO concentrations at first hours of the day (which is precisely what we could predict) in order to use that information to predict NO_2 concentrations at noon and later. Assuming that we could know NO concentrations in the first 8 h on the next day, and including forecasted temperatures as input, we are able to predict NO_2 concentrations around noon with an accuracy of the order of 70%, which is considerably better than forecasts based on persistence.

2. The data

For our study we consider measurements of NO and NO_2 concentrations taken at station B, which is one of the eight stations installed and supervised by the government throughout the city of Santiago in order to monitor the air quality. We consider data corresponding to hourly averages for the period that goes from 05/01/94 to 09/30/94, which includes the winter season of year 1994. NO concentrations are in parts per billion (ppb) with an average value of 154 and a standard deviation of 138. For NO_2, the average is 61 ppb and the standard deviation is 39 ppb. We also have data corresponding to temperature, relative humidity and wind velocity measured at the same station with the same frequency.

Fig. 1 shows the average NO concentrations per hour of the day for each day of the week. The period considered permits the inclusion of 21 complete weeks in the averages. The sharp peaks at around 8:00 AM and 8:00 PM may be explained by the traffic congestion produced at rush hours, because measurements are taken at a point which is close to an avenue that is a enforced pass for cars traveling between downtown and a residential area. For this reason, it is reasonable to find that the peaks are less pronounced on Saturday and Sunday. For working days, there are no great differences on daily variations of average concentrations of NO. We have

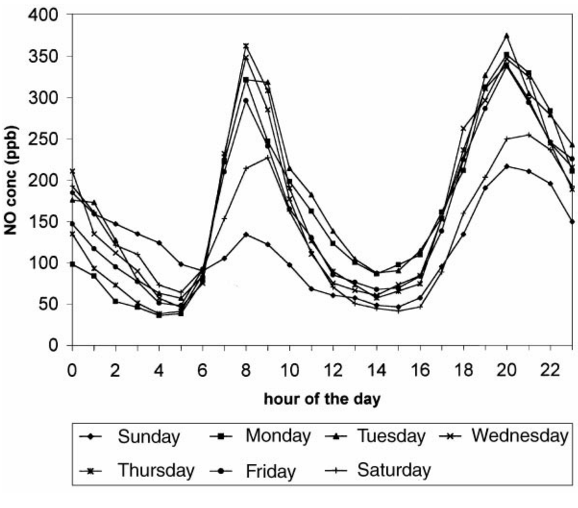

Fig. 1. Average daily variation of NO concentrations (p.p.b.) from 05/01 to 09/30, 1994.

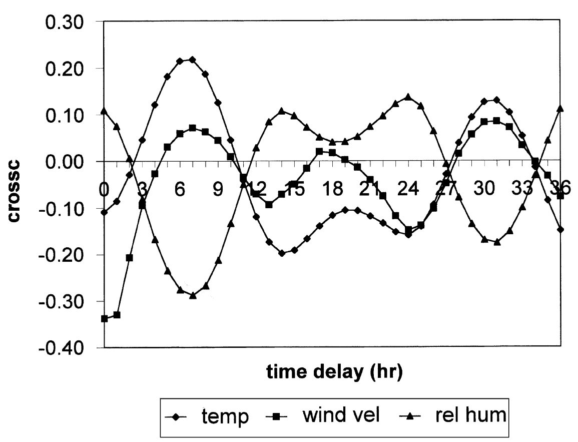

Fig. 2. Delayed cross-correlation function between NO concentrations and temperature, wind velocity and relative humidity.

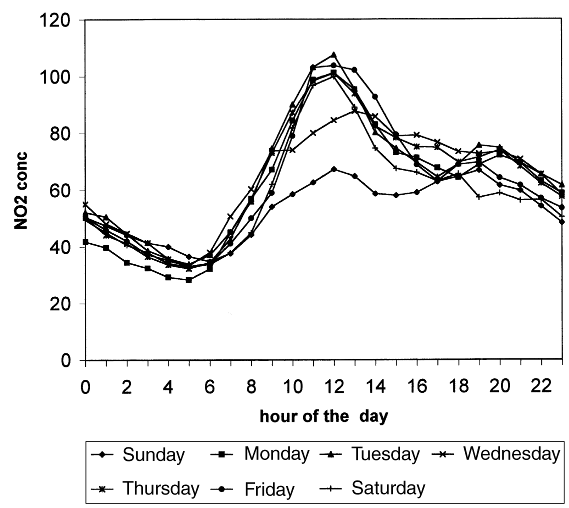

Fig. 3. Average daily variation of NO_2 concentrations (p.p.b.) from 05/01 to 09/30, 1994.

also looked for correlations between NO concentrations and some meteorological variables, namely temperature, wind velocity and relative humidity. We have calculated a delayed cross-correlation function given by

$$C_{xj}(T) = \frac{\langle x(t+T)y_j(t)\rangle - \langle x\rangle\langle y_j\rangle}{\sqrt{(\langle x^2\rangle - \langle x\rangle^2)(\langle y_j^2\rangle - \langle y_j\rangle^2)}}, \quad (1)$$

where $\langle\rangle$ means average over the whole series and T ($T = 0, 1, \ldots T_{\max}$) is the time delay of measurement of variable x with respect to the measurement of variable y_j. $x(t)$ is the NO series and $y_j(t)$ is the series for the meteorological variable ($j = 1$ is the temperature, $j = 2$ is the wind velocity and $j = 3$ is the relative humidity) and t in Eq. (1) runs through the whole series ($t = 1, \ldots, 3672$ for $T = 0, t = 1, \ldots, 3671$ for $T = 1$, etc.). C_{xj} varies between 1 (total correlation) and -1 (total anticorrelation). Fig. 2 shows this delayed cross-correlation function for $T_{\max} = 36$ h. The peak of anticorrelation for wind velocity at $T = 0$ is explained by the fact that wind favors dispersion of pollutants. The mirror symmetry between the curves for temperature and relative humidity is consistent with the known anticorrelation between these two variables, so the information provided by them may be somewhat redundant. We observe that actual NO concentration is weakly correlated with the temperature 7 h before. The coincidence of local maxima of correlation or anticorrelation with a 24 h delay for all three meteorological variables suggests that we could include values of this variables as inputs to an algorithm aimed to predict NO concentrations on the basis of information of the previous day.

Fig. 3 shows the average NO_2 concentrations per hour of the day for each day of the week. The period considered is the same as for NO. The peak around noon

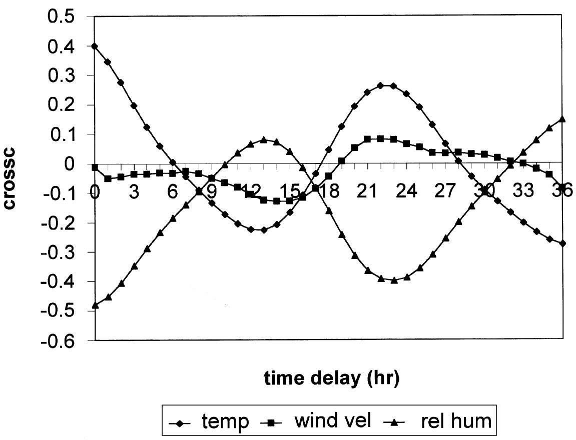

Fig. 4. Delayed cross-correlation function between NO_2 concentrations and temperature, wind velocity and relative humidity.

seems to be mainly the result of the photochemical reactions that produce NO_2 from NO, given the high concentration of the latter around 8 AM. There is no peak following the 8 PM NO maximum due to the absence of solar radiation at that time. As in the NO case, NO_2 levels are much lower on Sundays. Fig. 4 shows the delayed cross-correlation between NO_2 and temperature, wind velocity and relative humidity. We can verify that there is a high correlation between present values of NO_2 and temperature, which signals that with NO present, high temperatures favor the creation of NO_2. Correlation between NO_2 and wind velocity is relatively small. The information provided by the NO_2-humidity curve seems again redundant when comparing it with the NO_2-temperature curve.

3. Prediction of NO concentrations

In the work of Perez et al. (2000), it is shown that a good prediction of PM2.5 particle concentrations in the atmosphere of Santiago is obtained up to 24 h in advance using a three layer feed forward neural network.

A feed forward neural network may be described as a regression using the following scheme:

$$y_t = f_t(x_1, x_2, \dots, x_p), \tag{2}$$

where $x_1, \dots, x_p$ represent the data used as input and y_t is the predicted value of the variable of interest at time t. For every $t (t = 1, \dots, p)$ there will be a different function f_t (usually called transfer function). The form of this function is obtained after adjusting a set of parameters using a training set of data. The quality of the prediction may be obtained from the performance on the test set of data. One way to estimate percent errors is

$$\mathrm{PE} = \frac{\langle |y_{tp} - y_{ta}| \rangle}{\langle y_{ta} \rangle}, \tag{3}$$

where y_{tp} is the predicted value, y_{ta} is the actual value, and $\langle \rangle$ means average over the test cases.

In the case of a linear perceptron we assume that prediction of the variable of interest at a given time is obtained from a linear combination of the values of the input variables:

$$y_t = w_{t1}x_1 + w_{t2}x_2 + \cdots + w_{tp}x_p + w_{t0}. \tag{4}$$

For every t, the coefficients w_{tj}, which allow global adjustment of the training cases, may be calculated using any linear regression technique. Since Eq. (4) represents a special case of a two-layer neural network, we have applied an iterative algorithm (the Delta rule) widely used in this area (Rumelhart et al., 1986).

If in Eq. (2) the function f_t were a nonlinear function of a linear combination of the inputs, we would have a nonlinear perceptron. Additional room for a good fitting of the data may be achieved by introducing a set of hidden nodes z_{tk}, $(k = 1, \dots, n)$, in such a way that

$$z_{tk} = f(w_{tk1}x_1 + \cdots + w_{tkp}x_p + w_{tk0}), \tag{5}$$

and

$$y_t = f(v_{t1}z_{t1} + \cdots + v_{tn}z_{tn} + v_{t0}). \tag{6}$$

The function f we have used is

$$f(X) = \frac{1}{1 + \mathrm{e}^{-X}}. \tag{7}$$

In order to find the w's and v's we have used a generalized Delta rule (Rumelhart et al., 1986), which is a back propagation of errors, starting from the difference between calculated and actual outputs. The number of hidden units n_h is determined by trial and error. Since the number of adjustable parameters in a three-layer feed forward neural network with n_p input units, n_o output units and n_h hidden units is $n_o + n_h(n_p + n_o + 1)$, for n_p, n_o fixed and with a given training set, we cannot choose an arbitrary big n_h.

The network Perez et al. (2000) used was a three-layer feed forward that had 24 inputs, corresponding to the hourly concentrations of the pollutant on a given day and it had one output, which corresponded to the predicted value at a given time of the next day. Different networks were implemented for each hour on the next day. By dividing the 5 month series of data in non-overlapping groups of 24, 114 training patterns and 38 test patterns were generated. In their study they did not make a difference between working days and weekends. When values of meteorological variables were added as extra inputs, only a slight improvement in prediction was observed. Initially, we planned a similar strategy in order to generate a model to predict NO concentrations several hours in advance. Preliminary results however, showed very large prediction errors. One of the reasons for this may be the marked difference between working days and weekend days, especially Sundays, as shown in Fig. 1 (the difference was not so dramatic with PM2.5 data). We decided then to build a model leaving out Saturdays and Sundays. We kept the idea of using 24 inputs in a multilayer feed forward neural network, corresponding them to the data on a given day. The single output corresponding to the predicted NO concentration at a given time on the next day. We were left with the data from Mondays through Thursdays, in order to be able to predict concentrations from Tuesdays to Fridays. For this reason we were able to use just 84 sample cases (21 for each day) for each intended prediction on a following day. Since our plan was to predict hourly concentrations from 1 to 24 h in advance, we implemented the same number of neural networks. Sixty-three cases for training and 21 cases for testing were selected at random. Due to the small amount of training cases, we found it convenient to implement a neural scheme without hidden layers. This means a 24–1 perceptron trained using the delta rule for training. With the weights fixed, for each net we averaged the prediction error given by Eq. (3) over the 21 test cases. Results using a linear transfer function are shown by the curve 24–1 tu–fr in Fig. 5. Errors are rather large, especially at times when absolute concentrations are expected to be low. The smaller prediction errors during early hours may be explained by assuming that the perceptron is able to extrapolate the slope of the concentration curve coming from late hours of the previous day. This is consistent with the fact that errors increase dramatically when the curve for average concentrations has a change of sign in the slope. The fact that a prediction scheme that uses the pollutant concentrations at every hour of the previous day as input works better for PM2.5 than for NO has to do with the relative importance between the short-term behavior of the sources (mainly motor vehicles) and memory effects due to accumulation of the pollutant. The first being more important for NO and the second being more important for PM2.5. This is confirmed by the observation that on weekends, when emissions are expected to be lower, the decrease in PM2.5 is not as important as that in NO. After noticing the regularity of the daily variation of NO concentrations and the similarities upon comparing different working days (Fig. 1), we intended a prediction scheme based on persistence of the average conditions, as an alternative to the perceptron. We chose the concentration at a given time of the day averaged over the 63 training cases as a predictor for values at the same time of the day on test cases. However, since we are evaluating overall performance using Eq. (3), results should not differ significantly from the more traditional persistence scheme that uses the measured value at the same hour of the previous day as predictor. After comparing with the actual values we obtained the curve pers tu–fr shown in Fig. 5. We can verify that errors are, in general, significantly smaller than with the perceptron, except for the first 4 h. With the idea to improve these predictions we incorporated values of meteorological variables in a new scheme. We implemented a 3–1 perceptron, in which the inputs were the values of temperature, wind velocity and relative humidity as measured 24 h before the time of the intended prediction. The reason for this choice of delay was that we wanted to predict hourly pollutant concentrations on a given day based on information of the previous day and that we observe in Fig. 2 a peak at $T = 24$ for the three meteorological variables. Using again a linear transfer function, results are summarized by curve 3–1 met in Fig. 5. We observe that the new results represent an improvement in prediction with respect to simple persistence just from one to four hours in advance. Adding a hidden layer to this net does not lead to further improvement. Since this result implies that our ability to predict NO depends on our knowledge of meteorological conditions, we have also implemented a net with nine inputs, where these inputs are temperature, wind velocity and relative humidity at the time of the intended prediction, one hour before and one hour after. In this case, the best results are obtained with a feed forward neural network which has 9 units as input, a hidden layer of 5 units, and one output, which is represented by curve 9–5–1 met in Fig. 5. Here, prediction errors are smaller than persistence not only during the first hours, but also slightly smaller at most of the later times. We must notice that for this calculation we have used meteorological information corresponding to the actual values occurring at the day and time of the day we want to predict NO concentrations, which implies that, in practice, we should have at hand an independent source of meteorological forecasts.

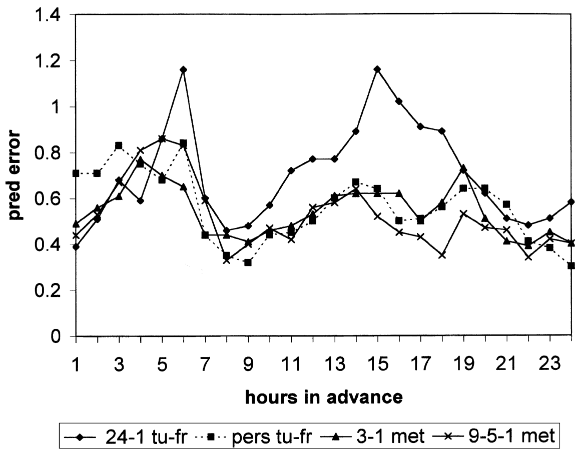

Fig. 5. Prediction errors for NO concentrations using 4 different methods. A linear perceptron with NO concentrations at every hour of the previous day as input (24–1 tu–fr), persistence (pers tu–fr) (where tu–fr means that we predict concentrations from Tuesdays to Fridays based on information from Mondays to Thursdays), a linear perceptron with values of temperature, wind velocity and relative humidity 24 h before as input (3–1 met) and a three layer neural network with temperature, wind velocity and relative humidity at the time of intended prediction, one hour before and one hour later as input (9–5–1 met).

It is important that we are able to predict NO concentrations with more accuracy than persistence, because by definition, pollution forecasts using the later method, are not able to detect critical episodes which are excursions from the average atmospheric conditions.

4. Prediction of NO_2 concentrations

It is likely that in the area where measurements are taken, NO_2 concentrations around noon depend strongly on NO concentrations at previous hours (see Figs. 1 and 3). Then, if we can predict NO concentrations at early hours with a better accuracy than persistence, we can use that information as input to predict NO_2 concentrations at noon or later. It is precisely at early hours that with a linear perceptron and a three-layer net we are able to forecast NO concentrations with relatively low average errors (see Fig. 5).

From the daily variation of NO_2 concentrations shown in Fig. 3, we found it convenient to leave Sundays out of modeling. For the period under study, we were left then with 126 days, from which we separated one of every four (31) for testing and the remaining (95) were saved for parameter adjustment. Persistence in this case means that we have averaged the NO_2 concentrations at noon for these 95 samples and used it as a predictor for the test cases. Calculating the percentage error as described above we found that its average value is 0.47. By using the actual values of NO concentrations at 0, 2, 4, 6 and 8 AM as inputs to a three layer network with 5 hidden units, we were able to decrease the average prediction error of NO_2 concentrations at noon to 0.43. If we added the actual value of temperature at noon as a sixth input, and keeping the same amount of 5 hidden units we were able to decrease the error to 0.33. After this result we decided to do a more complete calculation using a 6–5–1 neural network in the following way:

- Actual values of NO concentrations at 0 h, 2, 4, 6 and 8 AM as input.
- Actual value of temperature at the time of intended prediction as input.
- Prediction of NO_2 concentrations between 11 AM and 11 PM of the same day as output.

This means that we implemented 12 different networks where the training sets differed only in the sixth input and in the output.

Results are shown in Fig. 6, where a comparison with persistence is displayed. We observe that the neural network gives a better prediction than persistence between 11 AM and 1 PM, which are the times when maximum NO_2 concentration occurs most of the days. We notice, however, that this result depends on our capacity to forecast hourly NO concentrations and temperature.

Fig. 7 shows a case by case comparison between predictions of NO_2 concentrations using the neural network and the actual values at 1 PM. We observe reasonable good fit for most days in the test set. Two cases of high pollution are not very closely matched, which is not unexpected when using a regression model, because in general we have few of these cases that can be used for parameter adjustment during training.

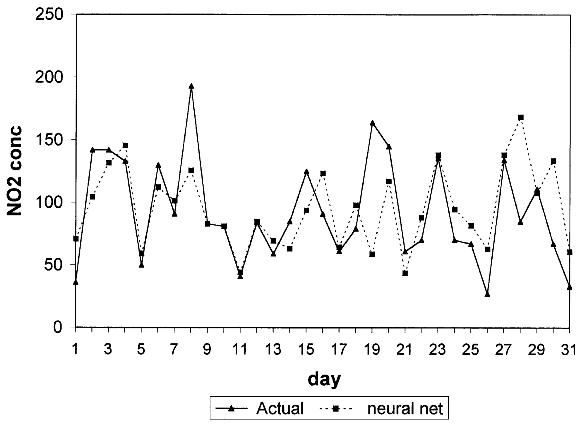

Fig. 7. Prediction of NO_2 concentrations at 1 PM for every day on the test set using a neural net which has NO concentrations on previous hours and temperature at 1 PM as input. Actual values are shown for comparison.

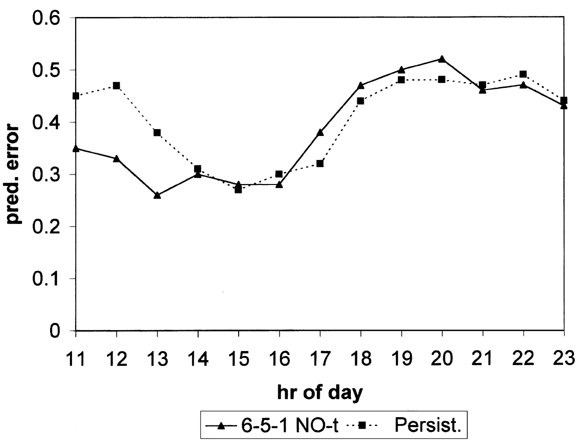

Fig. 6. Prediction errors for NO_2 concentrations using persistence (Persist.) and a three layer feed forward neural network which has NO concentrations at previous hours and temperature as input (6–5–1 NO–t).

5. Discussion

From the results obtained we can say that our capacity to predict NO and NO_2 concentrations in an area with heavy traffic depends in first place of our knowledge of the behavior of the sources (in this case persistence or variation of vehicle traffic). It is also important to have information of meteorological forecasts for the time of the day of the intended prediction. Unfortunately, we do not have detailed information about frontal passages in Santiago during the analyzed period. Although they are not very frequent, but considering that they alter the environment significantly, in an operational model, days when frontal passages occur could be extracted from the training set and treated with another model. Memory effects due to pollutant concentrations on previous days are not very important, at least much less important than the case of particulate matter like PM2.5 (Perez et al., 2000). We have verified that the main features of our results do not depend on the specific year nor the station chosen for analysis. Further improvement of the accuracy of the predictions seems possible using different variables as input. A feed forward neural network appears to be a convenient method for the problem under study because it is a nonlinear regression technique for which we have a reasonable control over the parameters to adjust. It allows a prediction of NO_2 concentrations with significantly lower error than persistence for the hours which would be of greatest interest to public authorities.

The results of this work could be very useful for researchers and technicians involved in the development of a more comprehensive operational model to be used for pollution control. The advantage of the neural network techniques and the effect of the variables we have taken into account for prediction should be considered in this task.

Acknowledgements

We would like to thank the support from the Research Department, Universidad de Santiago de Chile (DICYT), through project 049631PJ.

References

Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron) – a review of applications in atmospheric sciences. Atmospheric Environment 32, 2627–2636.

Jorquera, H., Perez, R., Cipriano, A., Espejo, A., Letelier, M.V., Acuña, G., 1998. Forecasting ozone daily maximum levels at Santiago, Chile. Atmospheric Environment 32, 3415–3424.

Perez, P., Trier, A., Reyes, J., 2000. Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile. Atmospheric Environment 34, 1189–1196.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning Internal Representations by Error Propagation. Parallel Distributed Processing. MIT Press, Cambridge, London, pp. 318–364.

Rutllant, J., Garreaud, R., 1995. Meteorological air pollution potential for Santiago, Chile: towards an objective episode forecasting. Environmental Monitoring and Assessment 34, 223–244.

Speitzer, F.E., 1980. Respiratory disease nrates and pulmonary function in children associated with NO_2 exposure. American Review of Respiratory Diseases 121, 3–10.

Ziomas, I.C., Melas, D., Zerefos, C.S., Bais, A.F., Paliatsos, A.G., 1995. Forecasting peak pollutant levels from meteorological variables. Atmospheric Environment 24, 3703–3711.