Original Research

Prediction of SO_2 and PM Concentrations in a Coastal Mining Area (Zonguldak, Turkey) Using an Artificial Neural Network

L. H. Tecer*

Balıkesir University, Department of Environmental Engineering, 10100 Balıkesir, Turkey

Received: August 28, 2006
Accepted: January 24, 2007

Abstract

In this study, artificial neural networks are proposed to predict the concentrations of SO_2 and PM at two different stations in Zonguldak city, a major coastal mining area in Turkey. The established artificial neural network models involve meteorological parameters and historical data on observed SO_2, PM as input variables. The models are based on a three-layer neural network trained by a back-propagation algorithm. The models accurately measure the trend of SO_2 and PM concentrations. The results obtained through the proposed models show that artificial neural networks can efficiently be used in the analysis and prediction of air quality.

Keywords: artificial neural network, MLP, backpropagation, time series, prediction of SO_2 and PM pollution, Zonguldak.

Introduction

Many epidemiological studies have demonstrated the association of air pollution with deterioration in human health. Experimental studies in humans have also shown that air pollutants, including ozone, sulphur dioxide, inhalable particle and nitrogen oxides, can all aggravate airway pathology by causing or provoking inflammation. The levels of air pollution have been associated with reduced pulmonary function, increased respiratory symptoms and even increased mortality [1-4].

Public opinion is becoming more concerned relating to environmental issues and at the same time new regulations are being enforced to control pollution deterioration of ambient air quality. Many environmental modelers have already successfully accomplished studies conducted on atmospheric air quality modeling. The aim of these studies has been to predict ground-level concentrations at specified locations in order to define air pollution control strategies.

Several methods have been used for modeling air quality. In recent years, artificial neural networks (ANNs) have become a popular and useful tool for modeling environmental systems [5]. Computing with neural networks is one of the fastest growing fields in the history of artificial intelligence (AI), largely because NNs can be trained to identify nonlinear patterns between input and output values and can solve complex problems much faster than digital computers. Owing to their wide range of applicability and their ability to learn complex and non-linear relationships – including noisy or less precise information – NNs are very well suited to solving problems in environmental engineering and in particularly towards the analysis of air pollution [6, 7].

Many modellers have successfully used neural network models in prediction of SO_2 episodes [8], ozone and NO_2 [9-10], and particulate matter levels [11-13]. In this

*Corresponding author; e-mail: lhtecer@balikesir.edu.tr

context, the approach of ANNs has become a popular and useful tool for predicting air pollution problems. These models provide a better alternative to statistical models because of their computational efficiency and ability to generalize. Initially, a neural network model was used by Bonzer et al. [14]. and applied it to predict SO_2 concentrations resulting from a power plant in complex terrain. In that study, meteorological parameters and historical data on observed SO_2 concentrations from within the vicinity of where a power plant was located were used as inputs. The result of the study was that ANNs can be applied to predict air pollutant concentrations in complex terrain. In recent years, there has been a major focus on ANN research towards the development of more efficient training algorithms. In general, two types of multilayer perceptrons are available; feedforward networks and feedback networks. Feedforward networks can be used for function approximation and pattern recognition. Over the last few years, recurrent networks have been proposed as an alternative to feedforward networks, particularly in time series applications. Feedback networks can be useful in predicting temporal and spatial patterns from data series [5].

In this study, artificial neural network models with back propagation algorithms were used to predict SO_2 and PM concentrations observed at two different stations in the coastal mining area of Zonguldak using meteorological data. The models accurately match the trend of SO_2 and PM concentrations. The results obtained by the proposed models show that artificial neural networks can be used efficiently in the analysis and prediction of air quality.

The Study Area

Zonguldak is a coastal city located in the western Black Sea region of Turkey at 41°27'N, 31°46'E (Fig. 1). It has a current population of about 108,000. The city is characterized by a "black diamond," signifying the importance attached to the coal produced in the area. Zonguldak is the primary mining centre of Turkey, with many underground coalmines run mainly by the government. In fact, the local economy has relied heavily on coal mining and the coal industry for decades. At present, the decline in the industry has already begun having adverse repercussions upon the local economy [15]. The development of the city and rapid rise in the population were associated with the growth of this coal industry after the 19th century. An adverse consequence of the rise in population and industrialization was the increase in environmental degradation, particularly in terms of air quality in Zonguldak. In Zonguldak, sulphur dioxide has been emitted into the atmosphere unchecked, especially in the hard coal mining region. In addition to SO_2 emission, the hard coal mines also emit particulate matter containing hazardous heavy metals into the city's atmosphere. It is well known that underground mining impacts directly the health of those working underground, but opencast mining creates wider air-quality deterioration due to dust and gaseous pollutants in and around the mining complexes [16]. The measured monthly average concentrations of SO_2 and PM (µg/m^3) are shown in Fig. 2. Annual average concentrations of SO_2 and PM display no significant variations, 65.55µg/m^3, and 72.21µg/m^3, 70.65µg/m^3 for SO_2 84.35µg/m^3, 79.68µg/m^3, 73.09µg/m^3 and for PM from 1999 to 2001. According to the Turkish Ministry of Health's air pollution control regulations, annual mean concentration limits are 60µg/m^3 for SO_2 and PM; however, all these records exceed 60µg/m^3.

In Zonguldak, chronic respiratory asthma and chronic bronchial diseases are found to be prevalent. Epidemiological surveys have shown that children and young adults suffer from asthma [17].

In 1997, an Environmental Management Committee was founded to investigate the environmental problems resulting from mining, industrial and domestic activities in Zonguldak and to coordinate subsequent environmental protection and improvement activities. The committee

Fig. 1. Map of Zonguldak City, Turkey, showing the locations of monitoring stations (M: Meteorological station, BE: Bahçelievler Station and CC: City Centre Station).

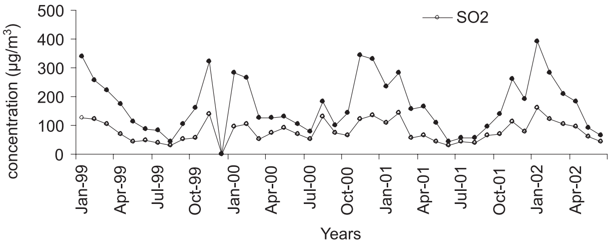

Fig. 2. Monthly average concentrations of SO_2 and PM during 1999-2002.

consists of representatives from the central government, the municipality of Zonguldak, industry and the coal company. Therefore, prediction of air quality is increasingly gaining importance.

The Method and Data

The Data

The data are provided by two air quality measurement stations established by the Turkish Ministry of Health. SO_2 and PM concentration data for the season through January-December 2002 are monitored for 24-hour periods at the two sites. These 24-hour SO_2 and PM concentration averages were used as pollutant parameters. One of the stations, Bahçelievler, has a hospital, residential area and some social centres in its locality. The other is the City Centre station, which was located directly on the city's main traffic artery close to schools and a business district in Zonguldak. The input parameters for the model include the meteorological variables provided by the Governmental Meteorology Office. The meteorological station is also located very close to the Bahçelievler station at a distance of approximately 100 meters. All parameters are shown in Table 1 with their descriptive statistics. The predominant wind direction is expressed showing 16 directions exhibited by the wind rose. The directions are numbered according to the reference numbers: N=0, E=0.5, S=1, W=-0.5.

The neural network model was developed for each station (Bahçelievler, City Centre) and for each pollutant (SO_2, PM), individually. Thus, the four prediction models were trained and tested. In these models, meteorological data were used as input parameters in the prediction of SO_2 and PM concentrations a day later. Furthermore, the SO_2 and PM concentrations pertaining to the day the meteorological data was taken (i.e. pollutant levels for the day prior to the prediction day) were also used as inputs. In this way, the potential effect of the perpetuity of pollution in the atmosphere was also included in the models. All daily data were divided into three parts. The odd days during the aforementioned 2002 period were used as the training set comprising 175 data, thereby increasing the ability of the model to generalize through understanding or inferring the state of pollution throughout the year using the meteorological data in the study period. The even days were used as the testing set comprising 150 data. Here it was desired that the model that had previously been trained using the 175 data set would use meteorological data to predict the SO_2 and PM concentrations for the next day, following each of the 150 test set days. Ad-

Table 1. Target and input variables considered in the neural network models.

Parameters	Unit	Minimum	Maximum	Mean	Std. Deviation
SO_2_BE* (target)	µg/m³	13.00	196.00	64.3529	35.9930
PM_BE (target)	µg/m³	2.00	335.00	70.2389	67.6290
SO_2_CC* (target)	µg/m³	13.00	299.00	91.1378	53.2464
PM_CC (target)	µg/m³	4.00	579.00	99.3314	97.8300
$SO_{2(t-1)}$_BE	µg/m³	13.00	196.00	64.5777	36.2837
$PM_{(t-1)}$_BE	µg/m³	2.00	335.00	70.0294	67.3942
$SO2_{(t-1)}$_CC	µg/m³	14.00	299.00	91.3167	53.0580
$PM_{(t-1)}$_CC	µg/m³	4.00	579.00	99.9971	98.7997
Pressure	mb	978.00	1018.00	1000.3021	6.1963
Cloudiness	x/10	.00	10.00	4.7713	3.2003
Relative Humidity	%	30.00	97.00	74.9589	13.3175
Wind speed_mean	m/sec	.00	8.00	2.2053	1.0949
Wind speed_max	m/sec	2.00	24.00	8.0381	3.6800
Wind direction		-	-	-	-
Temperature_max	°C	.10	35.20	11.3912	10.0242
Temperature_mean	°C	-.30	2.90	1.4120	.7202
Temperature_min	°C	-4.40	24.30	7.5842	7.5777

* BE: Bahçelievler station, CC: City Centre station

ditionally, 16 even days were used as a validation set for the model.

The Neural Network

Neural network models are biologically oriented structures and have been used for prediction of gas and particulate matter pollution. A brief description of the applications of neural network in atmospheric science is given by Gardner and Dorling (1998)[18]. A neural network consists of a number of processing elements, normally arranged in layers. Each processing element is linked to elements in the previous layer by connections that have an adaptable strength or weight. The adaptation of these weights is performed by a learning algorithm and this adaptation process improves the learning capability of the system, thereby enabling it to generalize for new situations.

The multilayer perceptron is the most popular type of neural network in such applications. It was developed during the 1980's and was the first recognized model that could process nonlinear data. Training of a multilayer perceptron proceeds in the following way: at the outset, the weights and biases in the network are initialized with small random initial values. Then a training pattern is applied to the input units, and the activations of neurons in the first hidden layer are calculated. The outputs produced by these neurons via the transfer function are fed to neurons in the next layer. This forward process is repeated at each layer until an output signal from the neurons in the output layer is obtained. In this learning process, one possible technique is to use a procedure known as gradient descent process that finds only the nearest local minimum in the mean square error from any given set of initial connection values. The backpropagation training algorithm uses this procedure to attempt to locate the global minimum of the error surface.

In this study, a three-layer recurrent network that consists of an input layer, output layer and one hidden layer was used.

Results and Discussion

The aim of this study was to predict the values of SO_2 and PM concentration using the artificial neural network model. For the neural network models, the meteorological variables at one site were used as input variables for two stations. The neural networks were selected to train the data collected at the two stations. The accuracy of the neural network models are depicted graphically in a time series of SO_2 and PM concentrations for all months. Fig. 3 presents the observed SO_2 and PM concentrations versus the model-predicted concentrations for January-December 2002. As seen from Figs. 3 a–b, the predicted pollutant concentrations are concordant with the observed concentrations on a majority of the days. The best results were obtained in the training data set. It was seen that the maximum SO_2 and PM concentrations were observed approximately during the period of winter season as shown in Figs. 2 and 3. Such concentrations increased as a result of urban heating systems, local meteorology and topography. The averages of the observed SO_2 concentrations were 100, 129 and 133 µg/m^3 in November, December and January, respectively. During this period, the averages of the observed PM concentrations were 145, 172 and 194 µg/m^3. The observed PM concentrations were higher than those of SO_2 at both stations. This can be attributed to the coal mining, energy (power plant) and steel industry in the city and its vicinity.

For validation of the models, regression method was used with the predicted values as independent variables and observed values as dependent variables. At the Bahçelievler station, the determination coefficient (R^2) between observed and predicted values of SO_2 concentrations for the training data was 0.829 and for testing data was found to be 0.668. (Table 2 which shows the determination coefficients and correlations between observed and predicted values for the pollutants at the two stations.) For PM concentrations, R^2 was 0.820 for the training set, and 0.808 for the testing set, respectively. The determination coefficient indicates that the fitted model explains the percentage of the variability between observed values and the neural network model predictions. ANOVA analysis was performed to check this statistical relationship. The P-value in the ANOVA analysis was found to be less than 0.01, which indicated a statistically significant relationship between the variables at the 99% confidence level. Correlations between observed and predicted values were highly significant at the 0.01 level.

Table 2. Determination coefficients and correlations between observed and predicted values for neural network models.

Station	SO_2				PM			
	Training set		Testing set		Training set		Testing set	
	R^2	Correlation	R^2	Correlation	R^2	Correlation	R^2	Correlation
Bahçelievler	0.829	0.910	0.668	0.817	0.820	0.905	0.808	0.899
City Centre	0.776	0.881	0.776	0.881	0.768	0.876	0.789	0.888

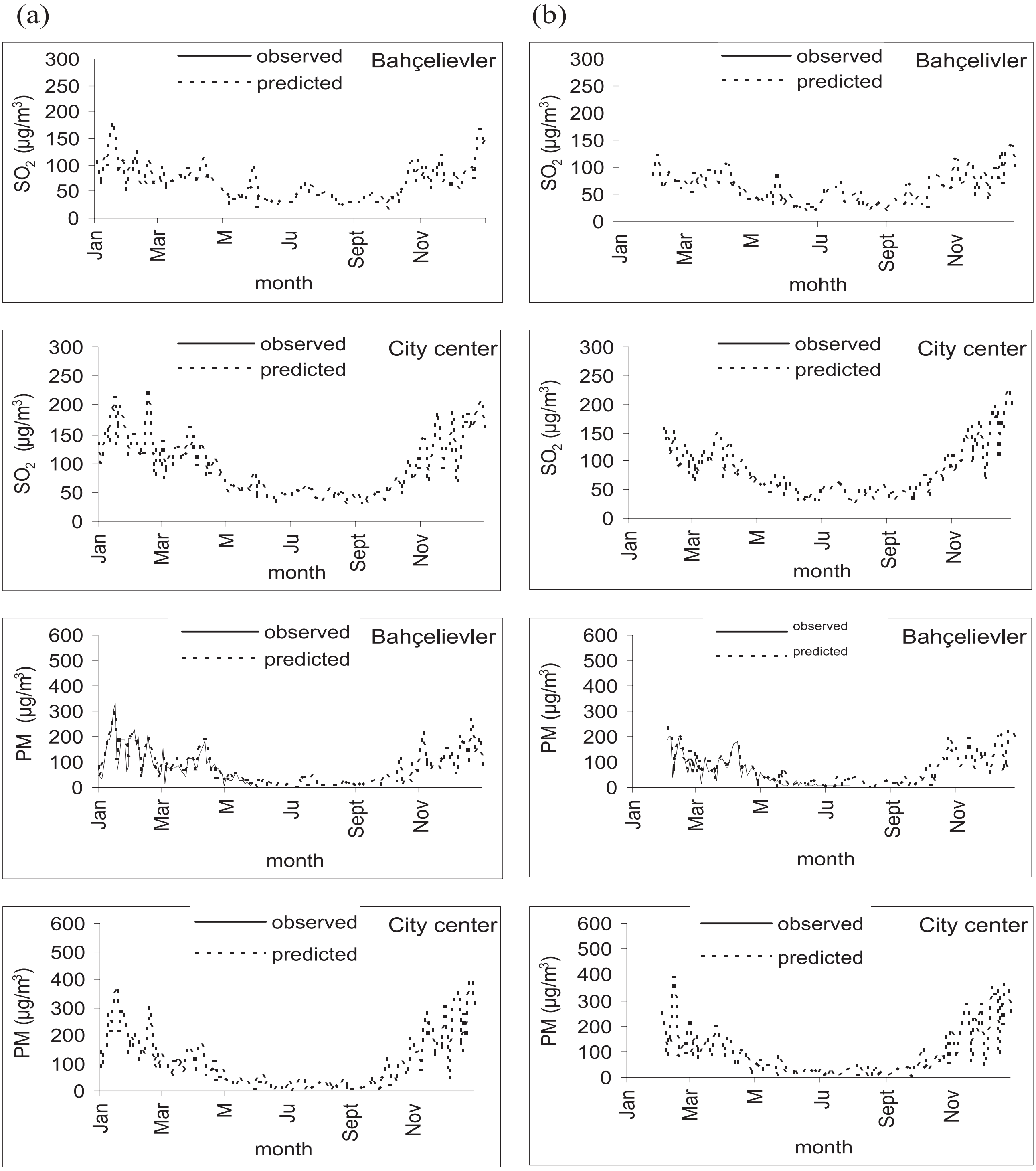

Fig. 3. Results of neural network models with (a) training set; (b) testing set for two stations (January-December 2002 period.)

Conclusion

In this study, the neural network model based on past (historical) values of SO_2, PM concentrations and meteorological variables were developed using artificial neural networks. The neural networks were trained in the hidden layer, which predicted the values by making use of the measured values.

The models explained the daily variation in pollutant concentrations and effectively predicted the observed values for the city of Zonguldak. We can conclude that knowledge of the sequence of past values of air pollutant concentrations is of considerable significance in order to predict its future values. Meteorological conditions at the intended time of prediction also have an important effect on air pollution modeling. It was also shown that the neural network model provides a good agreement with measured values of air pollutants.

Zonguldak's air quality has been poor due to the city's nearby coal mines. The main source of air pollution is the hard coal mines. To overcome air pollution in Zonguldak, the Environmental Committee decided to draw up and implement the Zonguldak Clean Air Project, reflecting the authorities' decision to take an active role in reducing air pollution. Air quality monitoring and modeling serves to assess the success of suitable policies, projects and programs. The results of modeling studies will show the health impacts of pollution and will promote local capacity to identify the goals and objectives of environmental policy of Zonguldak.

References

1. WONG, G.W.K., KO, F.W.S., LAU, T.S., LI, S.T., HUI, D., PANGI, S.W., LEUNG, R., FOK, T.F., LAI, C.K.W. Temporal relationship between air pollution and hospital admissions for asthmatic children in Hong Kong, Clinical and Experimental Allergy, **31**, 565, **2001.**
2. BALLESTER, F., SAEZ, M., HOYOS, S.P., INIQUEZ, C., GANDARILLAS, A., TOBIAS, A., BELLIDO, J., TARACIDO, M., ARRIBAS, F., DAPONTE, A., ALANSO, E., CANADA, A., GRIMA, F.G., CIRERA, L., BOILLOS, M.J.P., SAURINA, C., GOMEZ, F., TENIAS, J.M. The EMECAM project: a multicentre study on air pollution and mortality in Spain: combined results for particulates and for sulfur dioxide, Occup. Environ. Med., **59**, 300, **2002.**
3. TIMONEN, K.L., PEKKANEN, J., TIITTANEN, P., SALONEN, R.O., Effects of air pollution on changes in lung function induced by exercise in children with chronic respiratory symptoms, Occup. Environ. Med., **59**, 129, **2002**
4. LIFPERT, F.W., MORRIS, S.C. Temporal and spatial relations between age specific mortality and ambient air quality in the United States: regression results for counties, 1960-97, Occup. Environ. Med., **59**, 156, **2002**
5. MAIER, H.R., DANDY, G.C. Neural network based modelling of environmental variables: A systematic approach, Math. and Computer Modelling **33**, 669, **2001**
6. NISKA, H., HILTUNEN, T., KARPPINEN, A., RUUSKANEN, J., KOLEHMAINEN, T. Evolving the neurah network model for forecasting air pollution time series, Engineering Applications of Artificial Intelligence, **17**, 159, **2004.**
7. NISKA, H., RANTAMAKI, M., HILTUNEN, T., KARPPINEN, A., HUKKONEN, J., RUUSKANEN, J., KOLEHMAINEN, M. Evaluation of an integrated modelling system containing a multi-layer perceptron model and the numerical weather prediction model HURLEM for the forecasting of urban airborne pollutant concentrations. Atmospheric Environment **39**, 6524, **2005.**
8. CHELANI, H.R., RAO, C.V.C., PHADKE, K.M., HASAN, M.Z. Prediction of sulphur dioxide concentration using artificial neural networks. Environmental Modelling and Software, **17**, 161, **2002.**
9. PEREZ, P., REYES, T. Integrated neural network model for PM10 forecasting. Atmospheric Environment, **40**, 2845, **2006.**
10. GRIVAS, G., CHALAULAKOU, A. Artificial neural network models for prediction of PM10 hourly concentrations, in the Greater Area of Athens, Greece. Atmospheric Environment, **40**, 1216, **2006**
11. WANG, W., LU, W., WANG, X., LEUNG A.Y.T. Prediction of maximum daily ozone level using combined neural network and statistical characteristics, Environmental International, **29**, 555, **2003.**
12. BASURKO, E.A., BERASTEGI, G.I., MADARIAGA, I. Regression and multilayer perceptron-based models to forecast hourly O_3 and NO_2 levels in the Bilbao area, Environmental Modelling and Software, **21**, 430, **2006**.
13. LU, H.C., HSIEH, J.C., CHANG, T.S. Prediction of daily maximum ozone concentrations form meteorological conditions using a two-stage neural network. Atmospheric Research, **81**, 124-139, **2006**.
14. BOZNAR M., LESJAK M., AND MLAKAR P. A Neural Network Based Method for Short Term Predictions of Ambient SO_2 Concentrations in Highly Polluted Industrial Areas of Complex Terrain, Atmospheric Environment, **27B**, 2,.221-230, **1993.**
15. Zonguldak Local Agenda 21, http://www.iula-emme.org/la21/cities/Zonguldak
16. GHOSE, M.K., MAJEE, S.R. Air pollution caused by opencast mining and its abetement measures in India. J. Environmental Management, **63**, 193, **2001.**
17. TOMAÇ N., ACUN C., DEMIREL F., ERMIŞ B., AYOĞLU F.N. Zonguldak ilinde astim ve diger allerjik hastaliklarin prevalansi ve bazi risk faktörlerinin araştirilmasi. *X. Ulusal Allerji ve Klinik Immünoloji Kongresi*,, Adana, Turkey, 24-27 Eylül **2002**
18. GARDNER, M.W., DORLING, S.R. Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences. Atm. Environment, **32**(14/15), 2627, **1998.**