PREDICTION OF SO_2 AND PM CONCENTRATION IN COASTAL MINING AREA (ZONGULDAK) WITH ARTIFICIAL NEURAL NETWORK

Lokman Hakan Tecer
Karaelmas University, Dept. of Environmental Eng., 67100, Zonguldak, Turkey. Email: ltecer@cumhuriyet.edu.tr

ABSTRACT

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

Keywords: Artificial neural network, SO_2 and PM pollution, Zonguldak.

1. INTRODUCTION

Many epidemiological studies have demonstrated the association of air pollution with a deterioration of human health. Experimental studies in humans have also shown that air pollutants including ozone, sulphur dioxide, inhalable particle and nitrogen oxides all can aggravate airway pathology by including or enhancing inflammation. The levels of air pollution have been associated with reduced pulmonary function, increased respiratory symptoms and even increased mortality [1-3].

Several methods have been used for modeling air quality. The literature survey was divided into various categories; mathematic models [4], meteorological models [5,6], dispersion models [7], statistical models [8-10] and artificial neural networks (ANN) ([11-13]. Computing with neural networks is one of the faster 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 particular, in analysing air pollution.

Prediction of SO_2 episodes [14], ozone levels [13] have been successfully used by many modellers. The applications of multilayer perceptions in prediction of SO_2 concentrations were used by Mlaker et al., (1997)[15]. In this context, the approach of ANNs have become a popular and useful tool for predicting air pollution problem. These models provide a better alternative to statistical models because of their computational efficiency and generalization ability. At first, a neural network model was used by Bonzer et al, (1993)[11]. Bonzer et al, (1993) applied it to predict SO_2 concentrations resulted by a power plant in complex terrains. In that study, meteorological parameters and historical data on observed SO_2 from surrounding power plant were used as inputs. The result of the study is that ANNs can be applied to predict air pollutants concentrations in complex terrains. In recent years, there has been a major focus in ANN research on the development of more efficient training algorithms. In general, two types of multilayer perceptrons are available; feedworward 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 [16].

In this study, artificial neural network models are used to predict SO_2 and PM concentrations observed two different stations in coastal mining area (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 the artificial neural network can be used efficiently for analyzing and predicting an efficient air quality.

2. THE STUDY AREA

Zonguldak is a coastal city located in the western Black Sea region at position $41^0 27'$ N, $31^0 46'$ E (Figure 1). It has a current population of about 108 000. The city is characterized by "black diamond", the name which signifies the importance attached to the coal produced in the area. Zonguldak is the main mining centre of Turkey with many underground coalmines, mainly run by the government. In fact, the local economy has heavily relied on coal mining and coal industry for decades [17]. The adverse consequence for the population and industrialization is the increase of environmental degradation, namely air quality in Zonguldak. In the Zonguldak, sulphur dioxide has been emitted into the atmosphere with no controls particularly in hard coal mining region. In addition to SO_2 emission, the hard coal mines emits particulate matter which contains hazard heavy metals, into the city's atmosphere. Measured monthly average concentrations of SO_2 and PM ($\mu g/m^3$) are shown in Figure 2. Yearly average concentrations of SO_2 and PM have not significant differences, 65,55$\mu g/m^3$, and 72,21$\mu g/m^3$, 70,65$\mu g/m^3$ for SO_2 and 84,35$\mu g/m^3$, 79,68$\mu g/m^3$, 73,09$\mu g/m^3$ for PM from 1999 to 2001. The Turkish Health Ministry, air pollution control regulations annual mean concentration criteria is 60$\mu g/m^3$ for SO_2 and PM and all this records are higher than 60$\mu g/m^3$.

In Zonguldak, chronic respiratory asthma, chronic bronchitis diseases are prevalent conditions. Epidemiological surveys have shown that children and young adults suffer from asthma [18]. Therefore, the prediction of the air quality is increasingly gaining importance.

Figure 1. Map of Zonguldak City showing the locations of monitoring

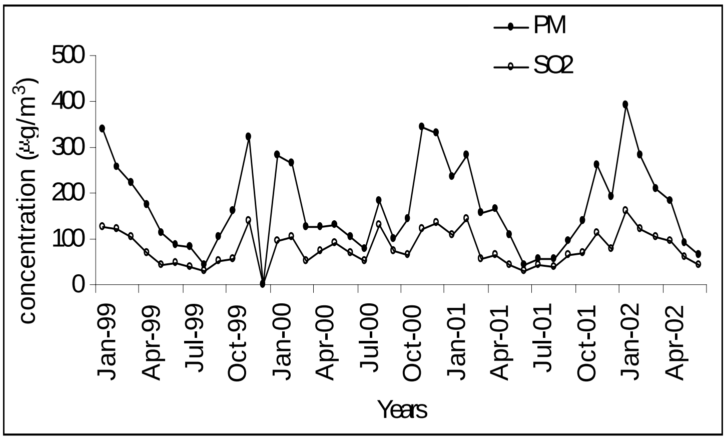

Figure 2. Monthly average concentrations of SO_2 and PM during 1999-2002

stations

3. THE METHOD AND DATA

The data

The data are provided from the two air quality measurement stations established by Ministry of Health. SO_2 and PM concentrations data for January-December 2002 season are monitored for 24 hour period at two sites. One of the station, Bahçelievler station (BE), surrounded by the hospital, houses and some social clubs. The other is City Centre station (CC) which was placed directly on the city's main traffic road close to schools and other offices in Zonguldak. The input parameters for the model include the meteorological variables which were provided by the Governmental Meteorology Office. The meteorological station (M) is also very close to Bahçelievler station with a distance of about 100 meters. All parameters are shown Table 1 with their descriptive statistics. The predominant wind direction is expressed with 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 is developed for each station (Bahçelievler, City Center) and for each pollutant (SO_2, PM), individually. Thus, the four prediction model have been trained and tested. For period of 2002, SO_2 and PM concentrations at two stations were used as target variable in the models. In the models, the same meteorological parameters were used as input variables. Also, the previous day concentrations of SO_2, PM concentrations were added into SO_2 and PM prediction models. The all daily data were divided into three parts. The odd days during the 2002 period were used as the training set consist of 175 data and the even days were used as the testing set consist of 150 data. Also, 16 even days were used as 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 applications of neural network in atmospheric science is given by Gardner and Dorling (1998) [19]. 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. This adaptation improves learning capability of the system and enables to generalize it for new situations.

The multilayer perceptron is the most popular type of neural network in the 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. First, the weights and biases in the network are initialized with small random initial values. A training pattern is then 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 that an output signal from the neurons in the output layer is obtained. At 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 form 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.

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

Parameters	Unit	Minimum	Maximum	Mean	Std. Deviation
SO_2_BE (target)	µg/m^3	13,00	196,00	64,3529	35,9930
PM_BE (target)	µg/m^3	2,00	335,00	70,2389	67,6290
SO_2_CC (target)	µg/m^3	13,00	299,00	91,1378	53,2464
PM_CC (target)	µg/m^3	4,00	579,00	99,3314	97,8300
$SO_{2(t-1)}$_BE	µg/m^3	13,00	196,00	64,5777	36,2837
$PM_{(t-1)}$_BE	µg/m^3	2,00	335,00	70,0294	67,3942
$SO2_{(t-1)}$_CC	µg/m^3	14,00	299,00	91,3167	53,0580
$PM_{(t-1)}$_CC	µg/m^3	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 Humudity	%	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	^{0}C	,10	35,20	11,3912	10,0242
Temperature_mean	^{0}C	-,30	2,90	1,4120	,7202
Temperature_min	^{0}C	-4,40	24,30	7,5842	7,5777

4. RESULTS AND DISCUSSION

The aim of this study is 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 are used as input variables for two stations. The neural networks are 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. Figure 3 presents the observed SO_2 and PM concentrations versus the model-predicted concentrations for January-December 2002. As seen from Figures 3 a–b, the predicted pollutants concentrations fit to the observed concentrations on most of the days. The best results were obtained in training data set. It was observed that the maximum SO_2 and PM concentrations were reached approximately at the period of winter season as shown in Figure 2 and 3. They have increased due to urban heating system, local meteorology and topography. The averages of the observed SO_2 concentrations are 100, 129 and 133 µg/m^3 in November, December and January, respectively. At this period, the averages of the observed PM concentrations are 145, 172 and 194 µg/m^3. The observed PM concentrations are higher than SO_2 at both stations. It can be attribute the coal mining, energy (power plant) and steel industry in the city and its surrounds.

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

(a) (b)

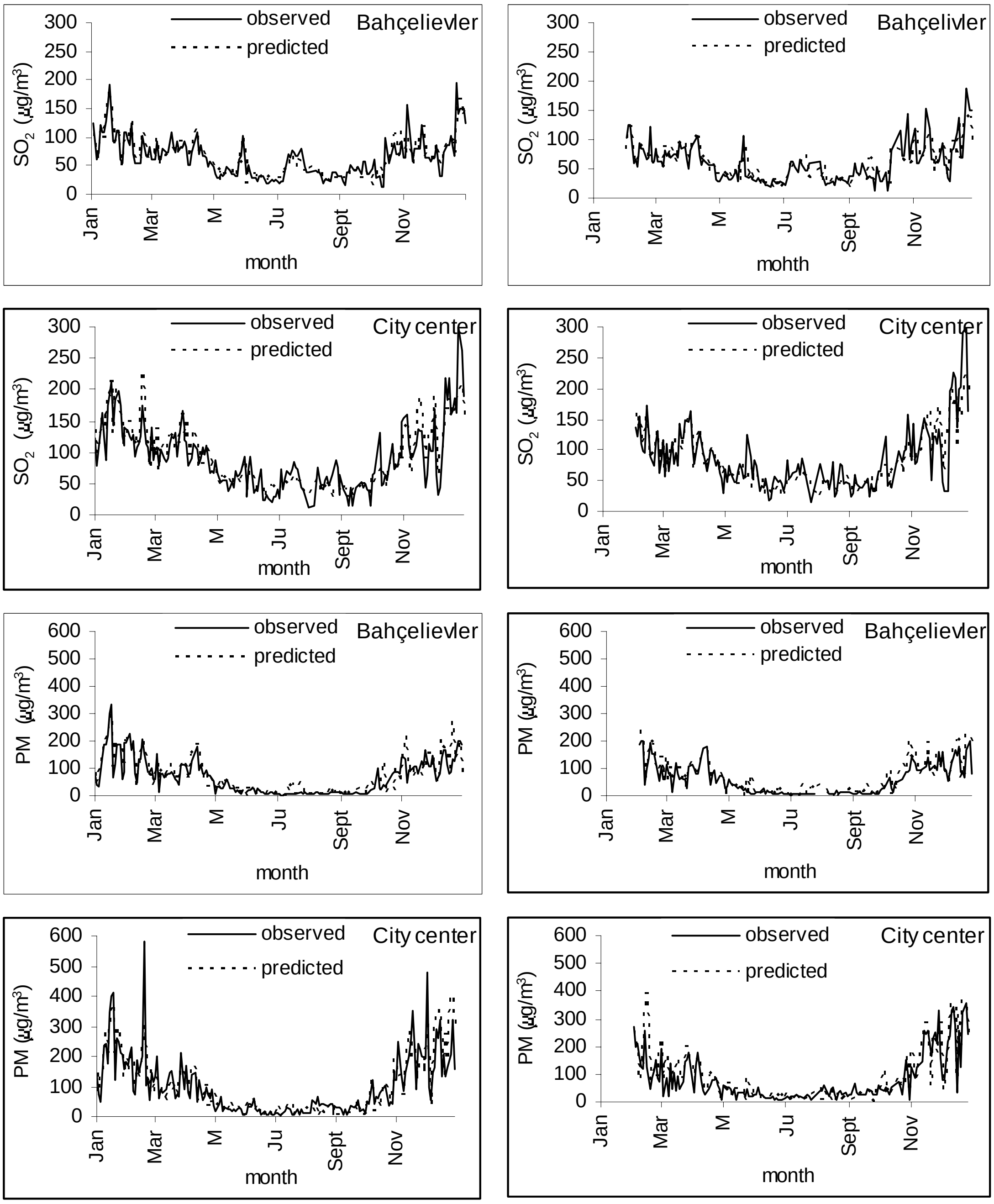

Figure 3. Results of neural network models with (a) training set; (b) testing set for two stations (January-December 2002 period.)

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

	SO_2	PM

Station	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
Citycenter	0.776	0.881	0.776	0.881	0.768	0.876	0.789	0.888

5. CONCLUSION

In this study, four forecast model based on past values of SO_2, PM concentrations and meteorological variables were developed using artificial neural networks. The neural networks are trained in the hidden layer which predicts the values by making use of the measured values.

The models explained the daily change of the pollutant concentrations and effectively predicted the observed values in Zonguldak city. We can conclude that knowledge of the sequence of past values of air pollutant concentrations is important in order to estimate its future values. Meteorological conditions at the time intended prediction have also an important effect. It is also shown that the neural network model provides a good agreement.

6. REFERENCES

[1] Monn, C., Alean-Kirkpatrick, P., Künzli, N., Defila, C., Peeters, A., Ackermann-Liebrich, U., Leuenberger, F. and SAPALDIA Team., 1999. Air pollution, climate and pollen comparisons in urban, rural and alpine regions in Switzerland (SAPALDIA study), *Atmospheric Environment,* 33(15), 2411-2416

[2] Williams, R., Creason, J., Zweidinger, R., Watts, R., Sheldon L., and Shy, C., 2000. Indoor, outdoor, and personal exposure monitoring of particulate air pollution: the Baltimore elderly epidemiology-exposure pilot study, *Atmospheric Environment,*34(24), 4193-4204.

[3] Alberini A. and Krupnick, A., 1998. Air Quality and Episodes of Acute Respiratory Illness in Taiwan Cities: Evidence from Survey Data', *Journal of Urban Economics*, 44(1), 68-92.

[4] Zanetti, P. (1998). Air Pollestion Modeling; Theories, Computational Methods and Available Software. 444 pp. The Looseleaf Compnay, Midsomer Norton, U.K.

[5] Pielke, R.A. and Uliasz, M. (1998). Use of Meteorological Models as Input to Regional and Mesoscale Air Quality Models – Limitations and Strengths. Atmospheric Environmental 32(8), 1455 – 1466.

[6] Kumar, N. and Russel, A. G. (1996). Comparing Prognostic and Diagnostic Meteorological Fields and Their Impacts on Photochemical Air Quality Modeling. Atmospheric Environmental 32, 1989-2010.

[7] Reynolds, S.D. and Seinfeld, J.H. (1975). Interm Evaluation of Strategies for Meeting Ambient Air Quality Modelling. Atmospheric Environmental 9(4), 433-447.

[8] Merz, P.H., Painter, L.J., and Ryason, P.R. (1972). Aerometric Data Analysis–Time Series Analysis and Forcast and an Atmospheric Smog Diagram, Atmospheric Env., 6,319–342.

[9] McCollister, G.M. and Wilson, K.R. (1975). Linear Stochastic Models for Forecasting Daily Maxima and Hourly Concentrations of Air Pollutants, Atmospheric Environment, 9, 417-423.

[10] Simpson, R.W. and Layton, A.P. (1983). Forecasting Peak Ozone Levels, Atmospheric Environment, 17(9), 1649–1654.

[11] Boznar M., Lesjak M., and Mlakar P., (1993), "A Neural Network Based Methot for Short Term Predictions of Ambient SO_2 Concentratıons in Highly Polluted Industrial Areas of Complex Terrain", Atm. Env., Vol.27B, No.2, pp.221-230.

[12] Mlakar P., Boznar M., Lesjak M., (1994), "Neural Networks Predict Pollution", Air Pollution Modelling and Its Application X, Plenum Press, pp.659-660, New York.

[13] Comrie A.C., (1997), "Comparing Neural Networks and Regression Models for Ozone Forecasting", J. Of the Air and Waste Management Assoc., Vol. 47, pp.653-663, June.

[14] Perautonis S.J., Vassilas N., Amanatidis G.T., et. al., (1994), "Neural Network Techniques for SO_2 Episode Prediction", Air Pollution Modelling and Its Application X, Plenum Press, pp.305-313, New York

[15] Mlakar P., Boznar M., (1997), "Perceptron Neural network–Based Model predicts Air pollution", Proceedings on Intelligent Information Systems, IIS'97, IEEE Computer Society, 0-8186-8218-3, pp.345-349.

[16] Beale, M., 1997. MATLAB User's Guide (5.1) Neural Network Toolbox, The Math Work Inc.

[17] Zonguldak Local Agenda 21, http://www.iula-emme.org/la21/cities/Zonguldak

[18] Tomaç N, Acun C, Demirel F, Ermiş B, Ayoğlu FN., 2002. Zonguldak ilinde astım ve diğer allerjik hastalıkların prevalansı ve bazı risk faktörlerinin araştırılması. X. Ulusal Allerji ve Klinik İmmünoloji Kongresi, 24-27 Eylül 2002, Adana.
[19] Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences. Atm. Environment, 32(14/15), 2627-2636.