PM-10 Forecasting using Neural Networks Model

S.H. Yu*, Y.S. Koo**, E.Y. Ha* and H.Y. Kwon*
* *Dept. of Computer Engineering, Anyang University, Anyang-Shi, 430-714 KOREA*
* * *Dept. of Environmental Engineering, Anyang University, Anyang-Shi, 430-714 KOREA*
E-mail: hykwon@anyang.ac.kr

Abstract

PM-10 is one of major air pollutants which affect on human health. Since PM-10 comes from various emission sources and its level of concentration is largely dependent on meteorological and geographical factors of the local region, the forecasting of PM-10 concentration is of great interest to protect daily human health. In this study, the dependent variables on PM-10 concentration were derived from the correlation analysis between PM-10 and meteorological as well as environmental factors based on the observations at the monitoring stations. Using the potential variables on the PM-10 level, the neural network model was developed and tested. The root mean square errors of the prediction in test runs were 0.064 to 0.077 and the test results implied that the system could be used in real forecasting within 10% error rates.

1. Introduction

Air pollution is a contamination of the atmosphere by gaseous, liquid, or solid waste or by products that have a serious affect on human health and the biosphere, reduce visibility, and damage materials. One of the major pollutants throughout the country is PM(Particulate Matter). PM is a complex mixture of solid and liquid particles that vary in size and composition, and remain suspended in the air. The size of ambient air particles ranges from approximately 0.005 to 100 μm in diameter. PM-10 is defined as particulate matter with a diameter less than 10 μm. Over the past decades, many health effect studies have shown an association between exposure to PM-10 and increase in daily mortality and symptoms of certain illnesses such as asthma, chronic bronchitis, decreased lung function, and premature death. Sources of PM-10 are numerous; naturally occurring processes and human activities all contribute to PM-10 concentrations. Some sources are natural, such as dust from the earth's surface, sea salt in coastal area, and biologic pollen. Periodic events like forest fires and dust storm can produce large amount of PM-10. In cities, PM-10 is mainly a product of combustion from mobile sources such as cars, buses, ships, trucks, and construction equipments, and from stationary sources such as municipal incinerators, power plants, and factories. Some PM are emitted directly into the atmosphere as particles, while other PM are produced by chemical reactions from the air pollutants in the air.

The forecasting of PM-10 concentration provides the public with the information with which they can make daily lifestyle decisions to protect their health. This information allows people to take precautionary measures to avoid or limit their exposure to unhealthy levels of air quality.

The major developed countries such as USA and UK have developed PM forecasting system using the statistical methods.[1,2,3,4] The neural network forecasting models for Seoul metropolitan area were developed in this study. In order to determine the major variables as the model input, various environmental and meteorological factors were analyzed based on the measured data at national monitoring stations of ambient air pollutants and local meteorology in Seoul metropolitan area. The neural network model was tested under various modeling conditions and the optimal model for PM-10 forecasting was finally presented in this paper.

The remaining sections are organized as follows: Section 2 introduces the proposed neural network model. The experimental results are presented in section 3. Finally, the conclusion is described.

2. Neural Network Model for PM-10 Forecasting

978-0-7695-3514-2/08 $25.00 © 2008 IEEE
DOI 10.1109/CIMCA.2008.173

IEEE computer society

2.1. Environmental and Meteorological factors affecting PM-10 level

It is important to determine input variables of neural network model for the PM-10 forecasting system. In this study, the factors affecting the PM-10 level were selected using the measured data at the monitoring stations. The meteorological factors were wind direction and speed, temperature, humidity, atmospheric pressure, rain fall, mixing height, atmospheric stability and irradiation. The environmental factors were concentrations of SO_2, O_3, NO_2, CO, and PM-10. The correlation analysis between PM-10 level and various factors were carried out and the major variables were finally selected and listed in Table 1.

Table 1. Correlation analysis results

Variable	Correlativity	Variable	Correlativity
SO_2	0.49	Amount of rainfall SUM	−0.19
O_3	−0.06	Higher Wind direction	0.06
NO_2	0.48	Higher Wind speed	−0.05
CO	0.5	a.m 6 wind speed	−0.22
Wind Speed	−0.25	a.m 6 Humidity	0.04
Temperature	−0.06	Temperature − dew point	−0.06
Humidity	0.02		

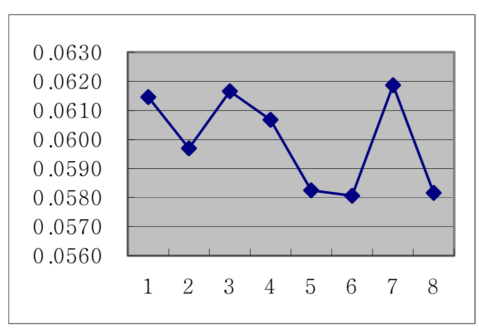
2.2. Neural Networks Model for PM-10 Prediction

Neural networks are non-parametric models which relate unknown inputs with corresponding outputs after learning by examples. In case, relations between inputs and outputs are unknown, like the PM-10 prediction problem, neural networks are effective and powerful tools for solving the problem. There are various models that have different learning rules and architectures as they are applied to different areas. For example, MLP (Multi-Layer Perceptron) model with EBP (Error Back Propagation) learning rule, which is a well-known neural network, is a supervised learning model and can be used in the case that the relations between inputs and outputs are explicit. There can be therefore various models as the number of the layers and the connection methods vary. Three or more layers MLP with two or more additional hidden layers can solve any nonlinear relation problem. A model capability depends on the connection topology between two layers or among neurons. We need therefore to find a neural model, a network architecture and a learning rule which are adequate for the PM-10 prediction problem.[5, 6]

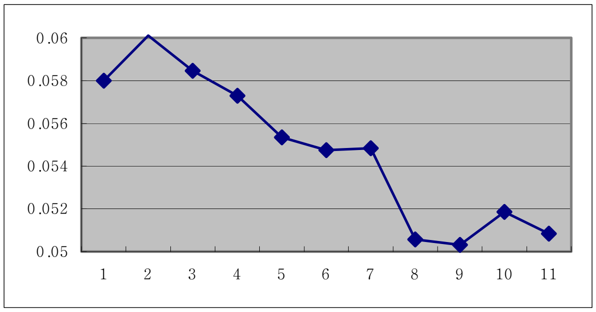
In this paper, we use MLP neural network with EBP learning rule which is the most popular and has the highest generalization capability. Considering that PM-10 prediction is a typical nonlinear problem, we construct the network with input, hidden(1), hidden(2) and output layers. The node numbers of each layers vary according to the experimental environments. The input values are normalized with 0.0 to 1.0. Though the output value can have discrete PM-10 levels, it has continuous value, 0.0 to 1.0, to represent a general prediction performance.

3. Experimental results

Three experiments are conducted. The first experiment is to find the effects of the network structures and learning to construct an optimal network. In the second one, we test the prediction capability of the model obtained from the former experiment. In the last one, to compare the result with a conventional approach, we conduct the prediction experiment with a linear regression model. Experimental data consist of the environmental and the meteorological ones observed in Seoul, Incheon and Suwon in the years 2004 to 2006. The environmental data are observed from 5 sites per city and the meteorological data are observed from a site per city. Input and Output is restricted from 0 to 260 and is normalized with 0.0 to 1.0. For level representation, we use modified level values which are details of AQI . Figure 1 shows the model's performances according to the various network parameters.



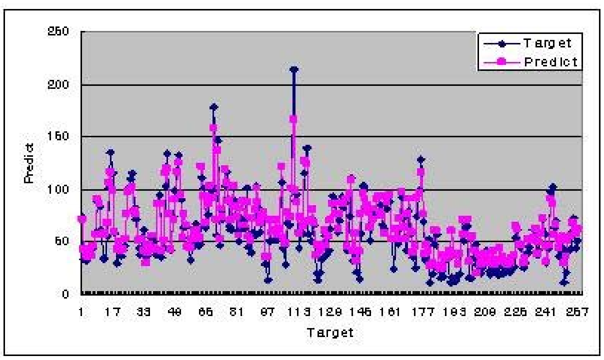
(a) by neurons numbers



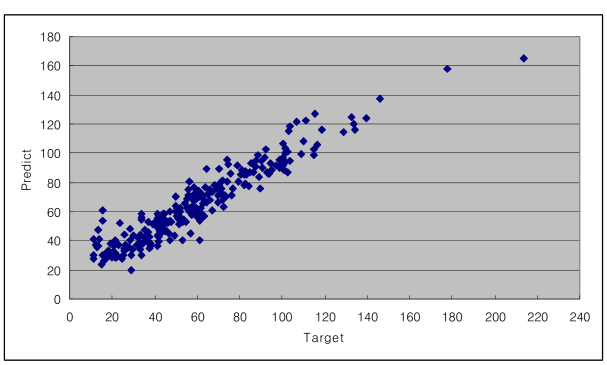
(b) by learning rates 0.21 to 0.0005.

Fig. 1. Performances for the various network parameters (Y-axis means total squared error sum, TSS. X-axis of (a) means the neuron numbers of hidden layer, 5 to 33 by 4. X-axis of (b) means the learning rate, 0.21, 0.14, 0.07, 0.035, 0.007, 0.005, 0.003, 0.001, 0.0009, 0.0007 and 0.0005.)

To measure the performance of the constructed neural network model, yearly data of 2004 and 2005 from 3 cities and its average are used as learning data and the yearly data of 2006are used as prediction data. Figure 2-a shows daily prediction trace and 2-b shows the dispersion rate.



(a) Transition

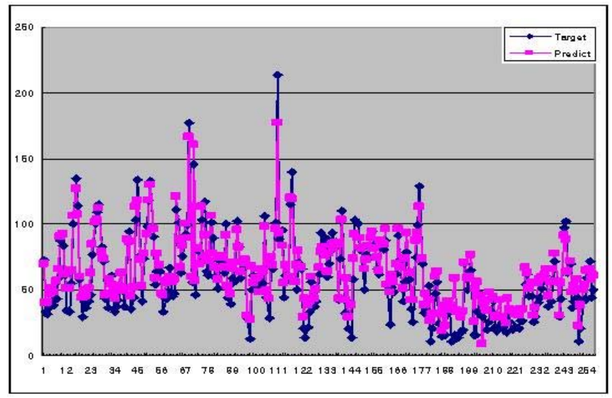


(b) Dispersion

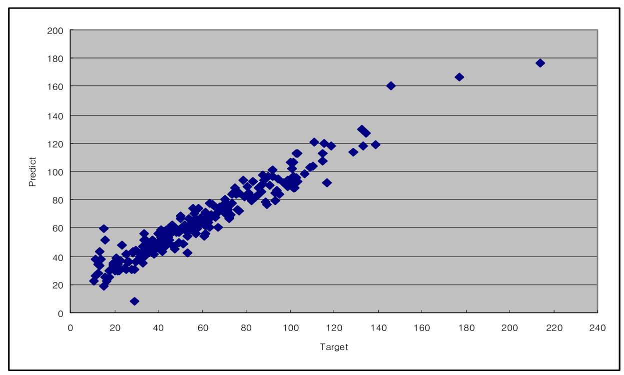
Fig. 2. Prediction Seoul area using neural network model.

(x-axis of (a) means 250 days, y-axis is PM-10 values)

To compare the result with that of a conventional model, linear regression model is applied to the same data. The results are as follows:



(a) Transition



(b) Dispersion

Fig. 3. Prediction Seoul area using linear regression model.

(x-axis of (a) means 250 days, y-axis is PM-10 values)

5. Conclusions

In this paper, we study on the factors to affect the PM-10 pollution and develop a PM-10 prediction model using MLP neural network model. The final results show that the accuracy and RMSE of the proposed neural network model are 0.064 and 0.077(20), and those of the regression model are 0.053 and 0.084(22). Numbers in the parenthesis are non-normalized PM-10 values. The proposed model can predict, therefore, PM-10 in 5%~10% error rate. Especially, neural model has an advantage that there doesn't need to analyze the input data before the data are used, like regression model. To improve the performance of the model, it needs to shorten the learning period from year to quarter month and to learn and predict PM-10 with multiple networks according to the PM-10 levels.

6. References

[1]T. S. Dye, D. S. Miller, C. B. Anderson, C. P. MacDonald, C. A. and Knoderer, B. S. Thompson: 'PM2.5 Forecasting Method Development and Operations for Salt Lake City, Utah', 2003 National Air Quality Conference, U.S. EPA, pp 1-18, 2003.

[2] M. Benjamin and J. Rousseau: 'Winter INFO-SMOG Program Forecast for the Greater Montreal Area', 2003 National Air Quality Conference, U.S. EPA, pp 19-23, 2003.

[3] Use of Time-Series Analysis to Examine the Link Between Photochemistry and PM Concentrations in Chicago, http://capita.wustl.edu/NEARDAT/WebLinks /pmupdate.htm.

[4] Air Pollution Forecasting in the UK, http: //www.airquality.co.uk/archive/reports/ list.php.

[5] Ian G. McKendry: 'Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM10 and PM2.5) Forecasting', Journal. of Air & Waste Management Association, Sep., 2002.

[6] J.L.McClelland, D.E.Rumelhart, Parallel Distributed Processing, Vol 1: Foundations, Cambridge, MA, MIT Press, 1986.