Air Pollution Analysis for Kannur City Using Artificial Neural Network

Nayana Vijayaraghavan¹, Gayathri S Mohan²

¹M. Tech Student, Department of Civil Engineering, K M C T College of Engineering for Women, Kozhikode, Kerala, India
²Assistant Professor, Department of Civil Engineering, K M C T College of Engineering for Women, Kozhikode, Kerala, India

Abstract: The pattern of economic growth that we are adopting is increasingly associated with environmental pollution. It is clear that in the developing or developed countries increased mechanization, transportation, populations, etc. causes air pollution which is a major and effective environmental problem. Present study focuses on analysis of air pollution in Kannur city of Kerala. In this study, prediction models are developed using non-linear autoregressive network with exogenous inputs (NARX Network) for the air pollutants. Several meteorological data such as ambient air temperature, relative humidity, and wind speed were given as input parameters while concentration of nitrogen dioxide (NO₂), sulphur dioxide (SO₂), respirable suspended particulate matter (RSPM) and suspended particulate matter (SPM) were considered as the output variable in this study. Also the four past values of output were fed back to the input. The performance of the developed model was assessed through a measure of minimum Mean Square Error (MSE). Result obtained shows that pollutant concentration at all two locations were found to be within the limit of NAAQS, India. Developed ANN model was found to be better model for future air pollution prediction with minimum MSE and it gives higher R².

Keywords: Air pollution; Artificial Neural Network; Back propagation neural network; Multilayer perceptron; Mean square error; NO₂; SO₂; RSPM; SPM

1. Introduction

Air pollution is an important issue nowadays, being a factor which influences both human health and activities. The growth of both an industrial and residential area is unplanned in many developing cities of India, thus, contributing to the air pollution problems. Air pollutants exert a wide range of impacts on biological, physical, and economic systems. Their effects on human health are of particular concern. A pollutant can be of natural origin or man-made. Based on the World Health Organization reports that 2.4 million individuals die annually from causes directly attributable to air pollution, 1.5 million of these from indoor air pollution.

Neural networks are parallel information processing systems with their architecture inspired by the structure and functioning of the brain. Neural networks can be trained to make classifications and predictions based on historical data. In ANN the neurons are grouped into layers as input layer, output layer and hidden layers between these two layers. The ability of ANN to learn by example tool to simulate dispersion phenomenon in complex environmental situations. ANN’s suitability for modelling complex system has resulted in their popularity and application in an ever increasing number of areas. Types of problems encountered by ANN include mathematical modelling, classification, clustering, forecasting, vector quantization, pattern association, and optimization problems.

2. Materials and Method

2.1 Data Sets

Two stations of different characterization (urban area, suburban background) were especially selected for use in the study. First station Kannur which is growing rapidly in population and second station Mangattuparamba which is within the premises of University of Kannur. The data used in this study are daily ambient air temperature, relative humidity, wind speed and daily concentration of SO₂, NO₂, SPM and RSPM for 2 years period from 2013 to 2014. Meteorological data were collected from Meteorological Department of Kerala and pollutant concentrations were collected from Pollution Control Board of Kannur. The data was divided into three sets which is learning set for ANN training, validation and testing set to verify the efficiency and correctness of the developed model.

2.2 ANN Model for Air Pollution Prediction

The NARX network uses the past values of the actual time series to be predicted and past values of other inputs to make predictions about the future value of the target series. The environment that was considered for the task is MATLAB Neural Network Toolbox. The NARX networks will have a linear input layer of neurons (default by MATLAB2014) for the hidden and the output layers were use the Transig and Purelin as transfer functions respectively and network was trained and tested by Levenberg - Marquardt (L-M) algorithm. 3-layer perceptron model used for the prediction of air pollutant concentration is as shown in Figure 2.1.
of NO$_2$, SO$_2$, RSPM and SPM. Also the four past values of output are fed back to the input.

The tool used for the analysis of data is “nntool” of Neural Network Toolbox. The “nntool” was run 20 times by varying the number of neurons in hidden layer from 1 to 20. The database was divided into three sections for early stopping. 75% of the data were used in training the networks, 15% were designated as the validation set, and the remaining 15% were employed in testing the networks.

The analysis of the performance of ANN models developed in this study was based on the comparison of predicted model values with actual observations. The mean square error was chosen as the statistical criteria for measuring the network performance.

2.3 Sensitivity Analysis

In ANN sensitivity analysis was carried out to know the relative importance of each of the input for prediction of pollutant concentration. In this study, sensitivity analysis was based on examining the effects of the input variables on the dependent variable following the so-called ‘leave-one-out’ method which corresponds to assessing changes in the network error that will be obtained if each input variable is removed at a time. Ultimately sensitivity analysis helps to identify the variable that most critically does affect the model.

3. Results and Discussion

3.1 ANN Modelling

Two ANN model were developed using back propagation algorithm to predict the concentration of air pollutants like sulphur dioxide (SO$_2$), nitrogen dioxide (NO$_2$), respirable suspended particulate matter (RSPM), and suspended particulate matter (SPM). The input and target values were normalized into the range of [0 1] in the processing phase. The weights and biases were adjusted based on gradient-descent back-propagation in the training phase. The mean square error was chosen as the statistical criteria for measuring of the network performance. The overview of the parameters and their values was shown in Table 3.1.

Table 3.1: Structure and Testing Results for the Neural Network Models

<table>
<thead>
<tr>
<th>Monitoring Station</th>
<th>Architecture</th>
<th>Training Function</th>
<th>R</th>
<th>R$^2$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kannur</td>
<td>19-6-4</td>
<td>Transig - purelin</td>
<td>0.98009</td>
<td>0.961</td>
<td>0.5738</td>
</tr>
<tr>
<td>Mangattuparamba</td>
<td>19-5-4</td>
<td>Transig - purelin</td>
<td>0.98178</td>
<td>0.964</td>
<td>0.3041</td>
</tr>
</tbody>
</table>

For Kannur, the best architecture has got a MSE of 0.5738 and regression value (R) of 0.98009 with 6 numbers of hidden neurons. For Mangattuparamba, the best architecture has got MSE of 0.3041 and regression value (R) 0.98178 with 5 numbers of hidden neurons. For Kannur, the value of coefficient of determination (R$^2$) is 0.961 and for Mangattuparamba, the R$^2$ value is 0.964.

3.2 Sensitivity Analysis in ANN

Sensitivity analysis has been carried out for both the selected locations Kannur and Mangattuparamba. The results of sensitivity analysis are shown in Table 3.2.

Table 3.2: Result of Sensitivity Analysis

<table>
<thead>
<tr>
<th>Monitoring Station</th>
<th>Parameters Excluded</th>
<th>Mean Square Error</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kannur</td>
<td>Ambient Air Temperature</td>
<td>0.6125</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Relative Humidity</td>
<td>0.6514</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Wind Speed</td>
<td>0.6063</td>
<td>3</td>
</tr>
<tr>
<td>Mangattuparamba</td>
<td>Ambient Air Temperature</td>
<td>0.3009</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Relative Humidity</td>
<td>0.3012</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Wind Speed</td>
<td>0.2998</td>
<td>3</td>
</tr>
</tbody>
</table>

For the monitoring station Kannur and Mangattuparamba, on removing the parameter relative humidity, the model MSE error of 0.6514 and 0.3012 were obtained respectively. Thus it has got the highest ranking among the three parameters for prediction of pollutant concentration.

4. Conclusion

In this paper, the study was carried out on modelling of air pollutants like SO$_2$, NO$_2$, RSPM, and SPM using Artificial Neural network. The study was focused at the estimation of the Mean Square Error (MSE) from the inputs and outputs which were given to ANN in the selected monitoring stations of Kannur district in Kerala.

For Kannur, the best architecture has got a MSE of 0.5738 and for Mangattuparamba; the best architecture has got MSE of 0.3041. From the sensitivity analysis it was clear that relative humidity has got highest MSE value than other two parameters for both Kannur and Mangattuparamba. For Kannur, MSE value was 0.6514 and for Mangattuparamba, MSE value was 0.3012.

References


Volume 5 Issue 10, October 2016


