ASSESSMENT AND PREDICTION OF SURFACE OZONE CONCENTRATION USING ARTIFICIAL NEURAL NETWORK MULTI LAYER PERCEPTRON IN JAKARTA CITY*

S.B, NUGROHO**, A. FUJIWARA**, J.ZHANG**

1. Introduction

Ozone concentrations are very difficult to model because of the different interactions between pollutants and meteorological variables (Sousa et al, 2007). The atmospheric chemistry and diffusion mechanism of surface ozone in the ambient air is quite complicated and depends on several parameters of a different nature, also including the horographic characteristic of the area of study. The combinations of those situations increase in particularly ozone at certain times which are strongly affected by the emission intensity of air pollutants, meteorological conditions and the presence of primary pollutants which reacts each other (G Nunnari, 1998). These complex and non-linear relationships of multiple variables can make statistical models awkward and complicated (Abdul-wahab, S.A,2006). Therefore, it is expected that they will under-perform when used to model relationship between ozone and the other variables that are extremely non-linear. Nowadays, artificial neural network (ANN) model have the potential to describe highly non-linear relationships such as those controlling ozone production. Therefore, the application of artificial neural network for ozone modeling has recently more popular than other regression model.

.

In forecasting with neural network, especially in the atmospheric studies which cite in this study, the most popular tool is provided by Multilayer Perceptrons (MLP). The multilayer perceptron (MLP) has been applied to a wide variety of tasks which can ber further categorized as prediction, function approximation, or pattern recognition. Prediction involves the forecasting of future trends in a time series of data given current and previous conditions. The MLP consist of a system of simple interconnected neurons or nodes which represent a nonlinear mapping between input and output parameters. The nodes are connected by weights and output signals which are a function of the sum of inputs to the node modified by a simple nonlinear transfer, or activation function. It is the superposition of many simple nonlinear transfer functions that enables the MLP to approximate extremely non-linear function. The output of the node is scaled by the connecting weight and fed forward to be an input to the nodes in the next layer of the ANN model. By selecting a suitable set of connecting weights and transfer functions, it has been shown that a MLP can approximate any smooth, measurable function between input and output vector (Gardner, 1998). Multilayer perceptrons have the ability to learn through training.

The work reported in this paper deals with the use of Artificial Neural Network model as a complimentary method combine with the previous statistical models to predict ozone concentration in Jakarta city. The model developed based on previous work on Structural Equation Model. The purpose of combining SEM and ANN model in this study was to describes and predict ozone as influence by emission source, meteorology and interaction of pollutants in the atmosphere. We also examine the relative percent contribution of each input parameter by analyze the sensitivity of the mean of input variables. This will help to the decision maker to understand the real world of ozone behavior in Jakarta city which will useful in the determining effective control strategies.

2. Research Methodology

The development of a Multilayer perceptron back propagation (MLP-BP) neural network model essentially involves a number of stages. First the variables to be used as the input parameters for neural network model have to be identified. This requires an understanding of the problem domain and may require insight from previous studies. To minimize the number of input parameters, previous statistical model results used to identify to most significant variables in the model. Regarding on the ozone prediction in Jakarta city, we used all parameter in the Structural Equation Model which explore interaction and cause-effects relationship among vehicular emission, meteorology and air pollutants. In

---

* Keywords: SURFACE OZONE, ROADSIDE, JAKARTA CITY, ARTIFICIAL NEURAL NETWORK
** Graduate School for International Development and Cooperation, Hiroshima University
totoksbn@hiroshima-u.ac.jp; TEL&FAX: 082-424-6919

this MLP-BP, we choose multi-input and multi output model structure to predict Nitrogen Oxide and Ozone concentration at time (t). We incorporated observed traffic data for vehicular emissions, meteorology data and pollutants data in our model. Totally, fifteen variables used in the first layer of input parameter of our MLP-BP ANN. Considering time series and serial correlation which already mentioned in our previous SEM model, four pollutants (PM_{10}, SO_2, CO, NO_2) variables used as the input parameter at time (t) and (t-1). Three traffic flow conditions as vehicle per road capacity per hour of Motorcycles, Passenger Car and high duty trucks variables used as the input parameter at time (t-2). Solar radiation (SR), ambient temperature (T), Humidity (RH), wind speed (WS), and wind direction -which in this study stated as sine and cosine of wind direction- variables used as the input parameter at time (t-1). The structure of our MLP-BP ANN for Jakarta city shown in figure 1

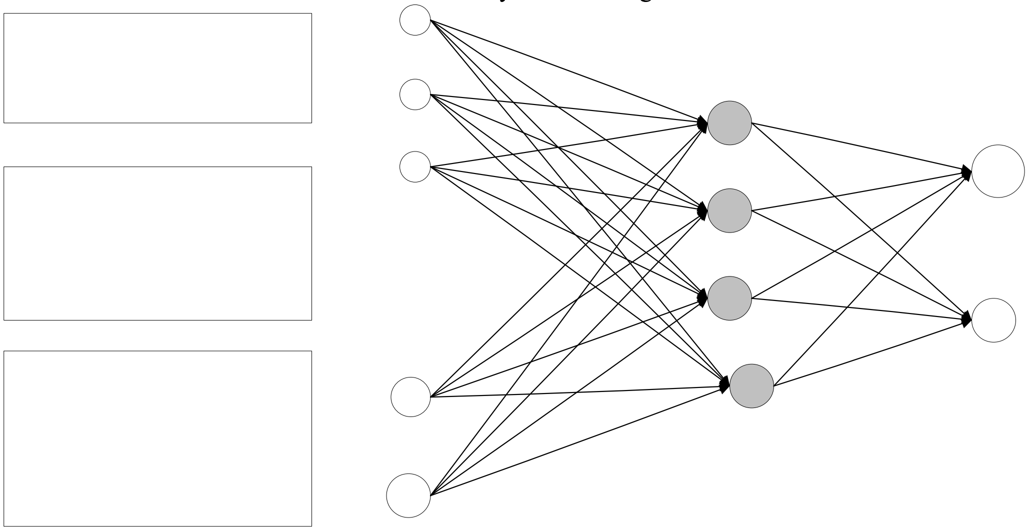

Figure 1. A Multi layer perceptron with four hidden layer

Usually about one-third of the data are used as the testing set (A.T.C Goh, 1995). For simulation, we use data collected in several major roads in Jakarta city. After discard the 'noise', total number data in the simulation is 647 cases. Those data consist of training data 471 cases (73%) and testing data set around 176 cases (27%). A critical review of the main application in atmospheric science which has been attempted by Gardner and Dorling (1998) whose comparison among models turn out to be rather unbalanced, since each model was trained with different kinds of data. In this study, in order to response these above matter, based on time observation basis, we select the data for training and testing data set. Then, we change in other both of training and testing data considering time observing and percentage of each category to observe the model performance in the different data set.

There is currently no rule for determining the optimal number of neurons in the second layer of our ANN model which consist of the hidden layer except through experimentation. Using too few neurons impairs the nural network and prevents the correct mapping of input to output. Using too many neurons impedes generalization and increases training time (A.T.C Goh, 1995). A common strategy and the one used in this study was to replicate the training several times, starting with two neurons and the increasing number while the monitoring the average sum square error. At the end of the training phase, the neural network should correctly reproduce the target output values for the training data provided the errors are minimal, i.e. convergence occurs. Training is carried out until there is no significant improvement in the error. In the third layer of our model, the pollutants NO and O_3 variables used as the output parameter at time (t). After several times trial and error by 2,3,4,5,6,7, and 8 hidden layer in our MLP-BP ANN we got the minimum of MSE at 4 hidden layer, and this 4 hidden layer use in this study. The testing set of patterns is then used to verify the performance of the neural network, on the satisfactory completion of the training. The testing phases assesses the quality of our MLP-BP ANN model and determines whether the neural network can generalize correct responses for patterns that only broadly resemble the data in the training set. In this case, we will use out model to predict the NO and ozone concentrations as the results of several policies which in practice will influence to the input parameter of our model. In this study, we used the conjugate gradient method for weight correction, as it was judged to be best for identifying the absolute minima of the error function. As the activation function of the single neurons, we choose the sigmoid function:

$$F(P) = \frac{A}{1 + \ell^{-K(p-s)}} \quad (11)$$

Where P is the potential, and s, K, A are threshold, the slope and the amplitude of the activation function, respectively. The model's behavior in both, training and testing steps, was evaluated calculated the following statistical parameters: correlation coefficient (R), mean bias error (MBE), mean absolute error (MAE), and root mean squared error (RMSE).

4. SUMMARY OF DATA

The data collected from 5 mobile ambient roadside air quality monitoring stations in Jakarta which is located 5-10 meter from main roads. We use a set of data on weekdays and weekend on several days in April, May, September and October 2005. All sampling points selected to describes the air quality situations of major roads in five districts of Jakarta. The monitoring stations operated at Thamrin road in central Jakarta, Fatmawati road in South Jakarta, Perintis Kemerdekaan road in East Jakarta, Yos Sudarso road in North Jakarta and Daan Mogot road in West Jakarta. The Fatmawati station is a sub-urban area of Jakarta city, and the others stations described the urban areas in Jakarta.

All monitoring stations were operated automatically and those are capable to measure CO, NO, NO_2, SO_2, PM_{10}, and O_3. In-situ meteorological data (solar radiation-SR, temperature-T, relative humidity-RH, wind speed-WS and wind direction-WD) were also recorded by using the basic meteorological sensors, which attached and installed at 10 meter height above the ground. The interval of measurement data of the average pollutants concentration and meteorological data are 30 minutes. Besides, traffic is monitored by video recording at the same locations. From video images, vehicles in the traffic are classified into 9 categories which consist of motorcycles, three wheeler (bajaj), passenger car, taxi, mini and medium bus, bus, small truck, truck and heavy truck. In this study, we categorize data into 3 categories which will adopt in the MLP-BP ANN. Those categories consist of motorcycle and three wheeler (MC), passenger cars (PC), high duty trucks (HDT) which includes all diesel vehicles. As input for the model, all observed traffic data converted to the traffic volume per road capacity (V/C) ratio for each road. Altogether observation data consist of 715 time points obtained from 5 stations. After careful inspection of the data for checking disorders in measurements, the original data size is reduced to 672 time points.

5. MODEL ESTIMATION OF TRAINING AND TESTING DATA

All about the diurnal and relationship among variables used in the MLP-BP ANN model already discussed in chapter 5. And so, we didn't discuss again in this chapter and focus on the categorized observed data for training and testing data set.

At the training stage, after 19983 epoch the best network was achieved with the minimum square error 0,0009998. For parameter NO, the R^2 is 0,9758 and the percent correct to measured data around 99,2147 %. For parameter O_3, the R^2 is 0,9883 and the percent correct to measured data around 97,7528 %. The estimation results shows in table 1 & table 2. Based on the sensitivity analysis of mean value of input parameter to output parameters, we got several results:

a. The most sensitive input parameter for NO output variable at time (t) is concentration of NO at time (t-1) follows by ambient temperature (T) at time (t-1) and concentration of CO at time (t).
b. The most sensitive input parameter for O_3 output variable at time (t) is concentration of NO_2 at time (t) follows by ambient NO_2at time (t-1) and concentration of O_3 at time (t-1).

All above information shown in figures 2, 3 and 4

Table 1 Overall best network of model for training data

Best Network	***Training***
Epoch #	19983
Minimum MSE	0.000999882
Final MSE	0.000999882

Table 2 Overall performance of model for training data

Performance	***LNNO/10(t)***	***O3/100(t)***
MSE	0.000235	0.000870
NMSE	0.024146	0.011667
MAE	0.011150	0.021608
Min Abs Error	0.000022	0.000021
Max Abs Error	0.099666	0.130791
r^2	0.975880	0.988345
Percent Correct	99.21466	97.75281

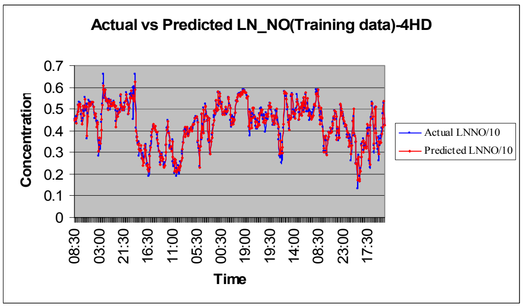

Figure 2. Actual and Predicted NO
for training data

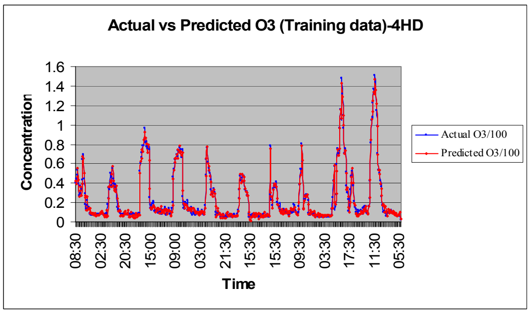

Figure 3 Actual and Predicted O_3 for training data

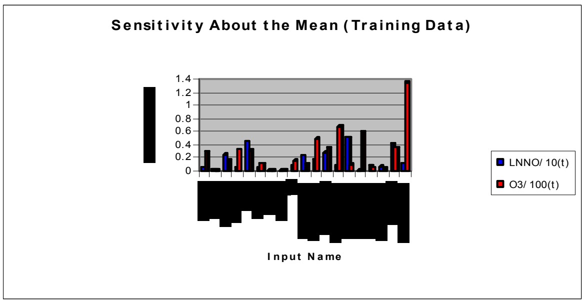

Figure 4 Sensitivity of the mean of training data

After the training stage, to obtain the model performance, we simulate testing data. For testing data for the output parameter NO, the R^2 is 0,6657 and the percent correct to measured data around 91,0448 %. For parameter O_3, the R^2 is 0,8056 and the percent correct to measured data around 88,0952 %. The estimation results shows in table 3 and table 4. Based on the sensitivity analysis of mean value of input parameter to output parameters, we got several results:

a. The most sensitive input parameter for NO output variable at time (t) is concentration of NO at time (t-1) follows by ambient temperature (T) at time (t-1) and concentration of CO at time (t).
b. The most sensitive input parameter for O_3 output variable at time (t) is concentration of NO_2 at time (t-1) follows by ambient NO at time (t-1), concentration of O_3 at time (t-1) and PM_{10} at time (t)

All above information shown in figures 5,6, and 7

Table 3 Overall best network of model for testing data

Output / Desired	LNNO/10(t)	O3/100(t)
LNNO/10(t)	122	5
O3/100(t)	12	37

Table 4 Overall performance of model for testing data

Performance	LNNO/10(t)	O3/100(t)
MSE	0.002388492	0.014450819
NMSE	0.376405861	0.23335301
MAE	0.03590879	0.083554292
Min Abs Error	5.20475E-05	0.000832212
Max Abs Error	0.207691896	0.450451075
R^2	0.665723116	0.805615744
Percent Correct	91.04477692	88.09523773

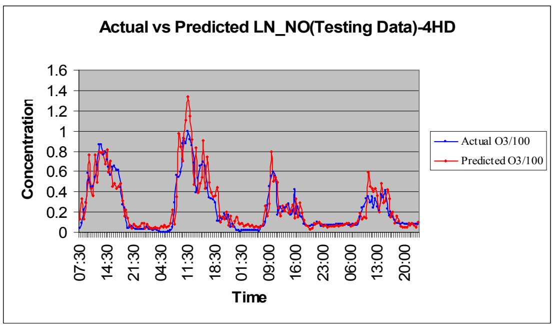

Figure 2. Actual and Predicted NO
for testing data

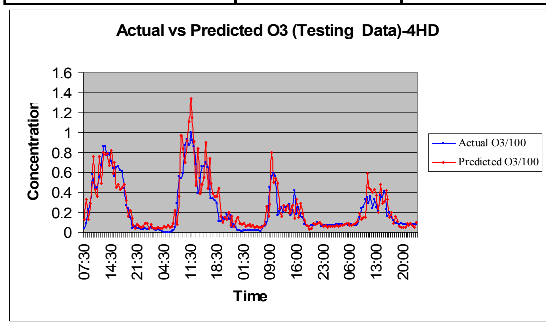

Figure 3 Actual and Predicted O_3 for testing data

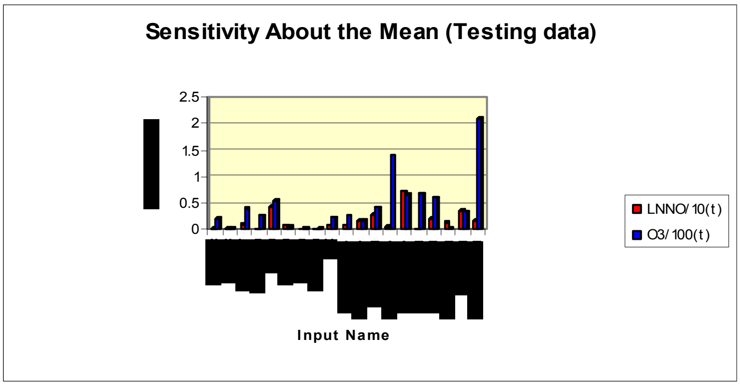

Figure 4 Sensitivity of the mean of testing data

6. CONCLUSION

Using data collected from roadside monitoring stations in Jakarta city, we confirmed the effectiveness of the proposed prediction of ozone concentration by ANN model. By the model, we found that there are several factors which sensitive to affects to the ozone concentrations.

REFERENCES

Chang, S.C., Chung, T.L. (2006) ‘Ozone variations through vehicle emission reductions based on air quality monitoring data in Taipei City, Taiwan, from 1994 to 2003’. *Atmospheric Environment*, 40, pp. 3513-3526.

Wang, W., Lu, W., Wang, X. and Leung, A.Y.T. (2003) ‘Prediction of maximum daily ozone level using combined neural network and statistical characteristics’. *Environmental International*, 29, pp. 555-562.

Gardner, M.W. and Dorling, S.R. (1998) ‘Artificial Neural networks (the multilayer perceptron)-a review of applications in the atmospheric sciences’, *Atmospheric Environment*, 32, pp. 2627-2636

Nugroho, S.B., Fujiwara, A., Zhang, J. (2006) ‘Analysis of Roadside Air Quality in Jakarta City: A Structural Equation Approach’, *JSME International Journal Series B*, 49, pp. 8-18.

S.A. Abdul-wahab & S.M. Al-alawi (2002) ‘ Assessment and prediction of thropospheric ozone concentration level using artifical neural networks’, Environmental Modeling and Software, 17, pp 219-228.