Prediction of PM2.5 and PM10 parameters using artificial neural network: a case study in Kemayoran, Jakarta

A M M B Putra¹,², Martarizal¹* and R M Putra¹,²

¹Physics Department, University of Indonesia, Depok, Jawa Barat, 16424, Indonesia
²Indonesian Agency for Meteorological, Climatological and Geophysical

*martarizal@sci.ui.ac.id

Abstract. It was recorded that in August 2019 the case of acute respiratory infection in Indonesia had doubled compared to the previous months. This is in line with the increasing levels of PM10 and PM2.5 in several regions in Indonesia. In the end the public is increasingly aware of the importance of air quality information. Prediction of air quality will greatly help the public to anticipate the dangers of declining air quality. The use of Artificial Neural Network can be a solution in making daily air quality forecasts whose parameters are not linear. This research shows that utilization of historical data parameters of temperature, humidity, air pressure, rainfall, sun exposure and wind speed as well as BMKG's PM10 and PM2.5 data is able to produce forecasting modeling for PM2.5 and PM10 concentrations in the Kemayoran area, Jakarta by utilizing Artificial Neural Network Modeling. The result is success to make prediction of PM2.5 and PM10 and it will be better if more historical data applied.

1. Introduction

Acute respiratory infection is a leading cause of illness or death in infants and toddlers. This situation is closely related to various conditions underlying it such as malnutrition as well as environmental conditions both pollution in the home in the form of smoke and dust in the environment and so on [8]. to prevent this from continuing, early treatment and care is needed for the potential deterioration in air quality.

The application of neural network for forecast ambient PM2.5 and PM10 concentrations described in this paper. The fact that the city of Jakarta is the center of government and industry in Indonesia and presents one of the most polluted cities in the world on August 2019. So we need to improve the air quality forecast include PM2.5 and PM10 in term the impact of bad air quality. The case are modeled using neural network tool. Some research was successfully forecast PM2.5 and PM10 by using various and different method. The prediction of the concentration of PM2.5 and PM10 is also affected by the corresponding meteorological parameters [1], it means the weather condition affected the condition of air quality is good or worst. some literature show many different statistical modeling approaches used for analyzing the influence of meteorological parameters on the prediction of the concentration of PM2.5 and PM10 [2, 3, 8, 9].

Multi Linear Regression was used for PM2.5 and PM10 daily prediction. This method using linear representation of nonlinear models [11]. It means ANN often to use as the method of air quality prediction. Previous study show that the Regression Analysis was using as the techniques for temporal prediction of PM2.5 and PM10 in terms of Linear Regression [12], Multi Layer Perception (MLP)
and Radial Basis Function (RBF) [7]. The selection of dimensionality problems, input variables and computation was advantages by Each techniques. From previous literature we can see that ANN very useful technique in terms of forecast the concentrations of PM2.5 and PM10 particles and ANN also can improve the air quality monitoring system with non linear parameters.

2. Methodology

2.1. Research Data
Meteorological parameter have been proven to have an effect on PM2.5 and PM10 Concentration [17]. To value the prediction of PM2.5 anda PM10 with high accuracy we must guarantee the quality of Meteorological data were good. So filtering and check the input variables are very important for PM2.5 and PM10 forecasting model. Previous study show that climate has influence on air pollution and wind speed known as the major influencer for PM2.5 and PM10 transport. Air temperature is the secondary influencer for pollution formation [15]. And another predictors of PM2.5 and PM10 prediction are pressure and humidity [21].

This paper are analyzed case by using some input variables daily values of wind speed, humidity, temperature, pressure, rainfall, sun exposure and a value of PM2.5 and PM10 particles from BMKG since 2017-2018 in kemayoran, Jakarta.

2.2. Artificial Neural Network
Artificial Neural Network (ANN) is a paradigm of information processing that is inspired by the biological nervous system, such as the information process in the human brain. The key element of this paradigm is the structure of the information processing system which consists of a large number of interconnected processing elements (neurons), working together to solve certain problems. The way ANN works like the way humans work, namely learning through examples. The layers making up ANN are divided into 3, namely the input layer, the hidden layer, and the output layer.

2.3. Basic Approach of ANN
ANN Technique have some approach [3]: First step ANN system will receive raw data or the output from previous neuron, then process them and value an output Second Step. determine the amount of quantity hidden layer on the problem to be solved, the layer type and also layer factor. Third Step 3 ANN consist of:
- Inputs
- Outputs
- Weight
- Weighted summation
  \[ u_j = \sum_{i=1}^{n} W_{ij} o_i \]  \hspace{1cm} (1)
  where: \( w_{ij} \) is the weight between \( i^{th} \) input and \( j^{th} \) neuron layer and \( o = [o_1, ..., o_n, 1] \) are n inputs.
- Activation function
  The most popular is sigmoid and purelin function defined as:
  \[ y = \frac{1}{1 + e^{-u_j}} \]  \hspace{1cm} (2)
  \[ y = x \]  \hspace{1cm} (3)
- Learning method – Neural network learns from previous experience.

2.4. Analytic of The result
The network learning method is traindx which has a learning error calculation using Mean Squared Error (MSE), so that the success of the network conducting training and network testing is determined
how much the Mean Squared Error (MSE) and the resulting correlation coefficient. Table 1 is the success rate of the network correlation coefficient.

| Correlation Interval | Correlation Level |
|----------------------|-------------------|
| 0                    | No Correlation    |
| >0-0.25              | Weak              |
| >0.25-0.5            | Enough            |
| >0.5-0.75            | Strong            |
| >0.75-0.99           | Very Strong       |
| 1                    | Perfect           |

### 3. Result and Discussion

#### 3.1. Layer Architecture

Figure 1 show the structure of artificial neural network used on this research. This research using the MLP artificial Neural Network method to predict PM10 and PM2.5 with 1 hidden layer [5]. improvisation is done on the number of inputs, the number and number of neurons. Difference from previous study, this research need 100 neuron for the hidden layer to get strong correlation. This study using two activation were sigmoid and purelin.

![Figure 1. The structure of the MLP layer](image)

#### 3.2. Layer Training

Table 2 and Table 3 shows the result of training regression . Both parameter show bad result if the structure only using 15 neuron and become strong when using 100 neuron. table 2 show that the layer with 15 neuron have $R=0.55$ (enough) and after add 100 layer the result is $R=0.89$ or very strong. The graphic of the training result shown in figure 2.

| Number Of neurons | R    |
|-------------------|------|
| 15                | 0.5  |
| 100               | 0.89 |

![Figure 2. The graphic of training result of PM2.5](image)
Table 3 show that the layer for PM10 prediction training with 15 neuron have R=0.607 (enough) and after add 100 layer the result is R=0.91 or very strong is shown at Table 3. The result of data training is shown at figure 3. The graphic showed various result for the initial data and gradually become better after all.

| Number Of neurons | R    |
|-------------------|------|
| 15                | 0.60 |
| 100               | 0.91 |

**Figure 3.** The graphic of PM10 training result

3.3. Particulate Matter Prediction
The result of prediction test is described at figure 4. Prediction of PM2.5 and PM10 levels is done using network weights that have been given in network training and that have been tested in network testing.

Analytic of the result using NRMSE and Coefficient correlation formula showing different level. PM10 prediction have NRMSE=0.4 and r=0.60. That result shows PM10 have a high correlation with meteorological parameter. While, PM2.5 prediction have NRMSE=0.2 and r=0.4. That result shows PM2.5 doesn’t have really strong correlation with meteorological parameter at all.

**Figure 4.** PM forecasting Result Graphic

4. Conclusion
The results of the study concluded that PM10 and PM2.5 levels can be predicted based on weather parameters with Multi Layer Perceptron artificial neural networks. With a network correlation that is categorized as "Very Strong" that is 0.91 for PM10 and 0.89 for PM2.5. The Error Value also show a good result which is for PM10 NRMSE= 0.43, Coefficient Correlation=0.60 then for PM25 NRMSE=0.29 , Coefficient correlation only 0.43. This shows that weather parameters can be used as input parameters to predict PM10 and PM2.5 levels using artificial neural network methods but still need improvement or another methods for PM2.5.

References
[1] S Avdakovic, N Dautbasic, M M Dedovic and J Dizdarevic 2016 The influence of wind speed humidity, temperature and air pressure on pollutants concentrations of
PM10 – Sarajevo case study using wavelet coherence approach (Bosnia: XI International Symposium on Telecommunications (BIHTEL)).

[2] G Grivas and A Chaloulakou 2006 Artificial neural network models for prediction of PM10 hourly concentrations, in the Greater Area of Athens, Greece (Atmos. Environ vol 40) chapter 7 pp 1216-1229.

[3] M Alkasassbeh, A F Sheta, H Faris and H Turabieh 2013 Prediction of PM10 and TSP Air Pollution Parameters Using Artificial Neural Network Autoregressive, External Input Models: A Case Study in Salt, Jordan, (Middle-East Journal of Scientific Research volume 14) chapter 7 pp 999-1009.

[4] I G McKendry 2002 Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2.5) forecasting (J. Air Waste Manag. Assoc volume 52) pp 1096–1101.

[5] M M Dedovic, S Avdakovic, I Turkovic, N Dautbasic, T Konjic 2016 Forecasting PM10 concentrations using neural networks and system for improving air quality (Sarajevo, Bosnia and Herzegovina : XI International Symposium on Telecommunications (BIHTEL)).

[6] E Stadlober, S Hormann, B Pfeiler 2008 Quality and Performance of PM10 Daily Forecasting Model (Atmospheric Environment volume 42) pp 1098–1109.

[7] Zheng Haiming, Shang Xiaoxiao 2013 Study on prediction of atmospheric PM2.5 based on RBF neural network (Fourth International Conference on Digital Manufacturing & Automation) Quebec pp 1287-1289.

[8] Ministry of Health 2006 guidelines for controlling ARI in children (Jakarta)

[9] H Taheri Shahraiyni and S Sodoudi, 2016 Statistical Modeling Approaches for PM10 Prediction in Urban Areas (A Review of 21st-Century Studies, Atmosphere chapter 7) pp 15.

[10] E Stadlober, S Hörmann and B Pfeiler 2008 Quality and performance of a PM10 daily forecasting model (Atmos. Environ. volume 42) pp 1098–1109.

[11] M Cai, Y Yin and M Xie 2009 Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach (Transp. Res. Part D Transp. Environ. volume 14) pp 32–41.

[12] G Corani 2005 Air quality prediction in Milan: Feed-forward neural networks, pruned neural networks and lazy learning (Ecol. Model) pp 513–529.

[13] I G McKendry 2002 Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2.5) forecasting (J. Air Waste Manag. Assoc. volume 52) pp 1096–1101.

[14] A Chaloulakou, G Grivas, N Spyrellis 2003 Neural network and multiple regression models for PM10 prediction in Athens: A comparative assessment (J. Air Waste Manag. Assoc. Volume 53) pp 1183–1190.

[15] P Perez 2012 Combined model for PM10 forecasting in a large city (Atmos. Environ volume 60) pp 271–276.

[16] W Z Lu, W J Wang, X K Wang, SH Yan and J C Lam 2004 Potential assessment of a neural network model with PCA/RBF approach for forecasting pollutant trends in Mong Kok urban air (Hong Kong Environ. Res. volume 96) pp 79–87.

[17] E Stadlobera, S Hormannb, B Pfeiler 2007 Quality and performance of a PM10 daily forecasting model (Atmo.envir volume 55) pp 1098-1108

[18] Y Aprianto, Nurhasanah, I Sanubary 2018 Prediksi Kadar Particulate Matter (PM10) untuk Pemantauan Kualitas Udara Menggunakan Jaringan Syaraf Tiruan Studi Kasus Kota Pontianak (positron) pp 15-20.