ABSTRACT

ANALYSIS AND PREDICTION OF PARTICULATE MATTER CONCENTRATIONS IN ROADSIDES OF NATIONAL CAPITAL REGION, PHILIPPINES USING FEED-FORWARD BACK-PROPAGATION ARTIFICIAL NEURAL NETWORK (FFBP ANN)

Begie C. Perdigones
University of the Philippines, 2019

Adviser: Dr. Mylene G. Cayetano
Reader: Dr. Fredegusto Guido P. David

The efforts to monitor and forecast particulate matter (PM) concentrations have been attracting attention in recent years due to PM's impact on air quality and climate change. As part of these efforts, this study focused on analyzing the spatial and temporal variations of PM in roadsides of National Capital Region (NCR), Philippines to identify the controls on its abundances and developing Feed-forward Back-propagation Artificial Neural Network (FFBP ANN) models to predict 3-hourly PM concentrations in the study area. A year-round hourly data (2016/10/01 – 2017/09/30) of PM10 and PM2.5 concentrations acquired by the Air Quality Monitoring System platform (AQMS, www.airtoday.ph) from four stations (Lung Center of the Philippines, LCP, Ayala Avenue, AYA, Edsa-Muñoz, MUN, and UST-España, UST) together with meteorological parameters were used. The highest (lowest) annual mean of roadside PM10 and PM2.5 concentrations were from AYA (LCP) with 31.05 ± 14.40 (mean ± 1σ) μg/m3 and 19.15 ± 10.15 μg/m3 (12.11 ± 9.32 μg/m3 and 10.92 ± 8.66 μg/m3), respectively. Pronounced peaks during the morning and the evening rush hours (07:00- 09:00 and 20:00 - 21:00) as well as higher PM concentrations during weekdays(peaked on Thursday and/or Friday) and lower concentrations on weekend were observed. These diurnal variations and day-of-the-week patterns indicate the importance of roadside emissions as source of PM in the study area. Moreover, seasonal variation was also evident, drawn from LCP and MUN; PM concentrations were high during dry season and low in wet season. However, this trend was not clearly seen from AYA and UST. Weak to moderate negative correlations were obtained between PM concentrations and meteorological variables such as temperature and wind speed and positive correlations with relative humidity and pressure. Furthermore, the ratios between PM2.5 and PM10, alongside with their dependence on temperature, indicate that majority of observed PM2.5 came from primary emissions which likely outweigh the signature of secondary formation via photochemical processes in the atmosphere. Based on the results, it is inferred that PM concentrations in selected sites in NCR exhibit spatial and temporal variations and are heavily controlled by roadside emissions and moderately influenced by meteorological factors. This analysis served as basis for developing models that could predict 3-hourly PM concentrations. ANN models of feed-forward back-propagation approach (FFBP) were developed for predicting PM10 and PM2.5 concentrations in each site. To see the impact of choosing the right input variables being fed to the network, performance of the models with different input variables were considered. Results showed that best performance of the models were obtained when 3-hourly data of meteorological parameters, previous PM concentrations, and time-scale variables used as traffic proxy - hour of the day (HoD), day of the week (DoW), and month of the year (MoY) were selected as network inputs. These were further trained, validated, and tested using different number of hidden neurons and activation functions. Performance of the models were evaluated using five performance indicators, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), Coefficient of Determination (r2), and Index of Agreement (IA). The best network topologies (input neurons-hidden neurons-output neuron) for PM10 prediction in each station were 10-17-1 (LCP), 10-16-1 (AYA), 10-8-1 (MUN), and 10-8-1 (UST), while for PM2.5, 10-9-1, 10-14-1, 10-8-1, and 10-11-1, with tangent sigmoid and pure linear combination of transfer functions. Except for AYA, ANN models for the other three AQMS had good prediction results. These were compared to the commonly used Multiple Linear Regression (MLR) models. Consistently, for all sites, both PM10 and PM2.5 prediction ANN models outperformed the linear regression models, obtained reduced RMSE (2.17% - 11.39%), MAE (5.98% - 10.47%), and lower underestimation and overestimation of observed values. Prediction ability was also better, denoted by higher values of r2, with 0.47 to 0.65 (3.23% - 38.24%), and IA that ranged from 0.78 to 0.89 (1.25% - 7.23%). It could be deduced that roadside PM concentrations could be predicted, and results suggest that FFBP ANN is a potentially viable approach to predict 3-hourly PM10 and PM2.5 concentrations.

Keywords: *diurnal variation, day-of-the-week pattern, roadside emissions, three-hourly prediction*