Geophysical Research Abstracts
Vol. 19, EGU2017-6358, 2017
EGU General Assembly 2017

Artificial neural network-based modelling and prediction of air pollutant concentrations in Sabah, Malaysia

Gissella Lebron (1), Erickson Fajiculay (1), Md Firoz Khan (2), Mohd Talib Latif (3), and Mikinori Kuwata (1)

(1) Earth Observatory of Singapore, Nanyang Technological University, Singapore (gissella.lebron@ntu.edu.sg), (2) Centre for Tropical Climate Change, Institute of Climate Change, Universiti Kebangsaan Malaysia, (3) School of Environment and Natural Resources, Universiti Kebangsaan Malaysia

Artificial neural networks, a sub-branch of machine learning, have been slowly but steadily gaining traction as a predictive modelling technique in atmospheric sciences. In particular, they have been proven to be a reliable tool to capture the highly complex and non-linear relationships between air pollutants and meteorological parameters.

In this ongoing study, feedforward and back-propagated neural network models are trained, cross-validated and tested to predict hourly concentrations of NO_2, NO_x and CO in an urban area located in Tawau, Sabah in Malaysia. Year-long hourly meteorological data (temperature, relative humidity, wind speed and wind direction) routinely collected by the Tawau monitoring station are used as predictor variables, along with the time of day, day of week and month of year time parameters. Sensitivity analysis carried out on the modeled networks reveals the relative influence of the predictor variables on the air pollutant concentrations. This allows for the removal of the least significant predictors resulting in neural networks with fewer input nodes. Performance-wise, the simplified networks predict the pollutant concentrations with accuracy comparable to the originally developed 7-predictor network models. The results also show that the neural network models perform better than conventional multilinear regression models in predicting the aforementioned air pollutant concentrations given the same input variables.