Development of multiple linear regression for particulate matter (PM10) forecasting during episodic transboundary haze event in Malaysia

ABSTRACT

Malaysia has been facing transboundary haze events every year in which the air contains particulate matter, particularly PM10, which affects human health and the environment. Therefore, it is crucial to develop a PM10 forecasting model for early information and warning alerts to the responsible parties in order for them to mitigate and plan precautionary measures during such events. Therefore, this study aimed to develop and compare the best-fitted model for PM10 prediction from the first hour until the next three hours during transboundary haze events. The air pollution data acquired from the Malaysian Department of Environment spanned from the years 2005 until 2014 (excluding years 2007–2009), which included particulate matter (PM10), ozone (O3), nitrogen oxide (NO), nitrogen dioxide (NO), carbon monoxide (CO), sulfur dioxide (SO2), wind speed (WS), ambient temperature (T), and relative humidity (RH) on an hourly basis. Three different stepwise Multiple Linear Regression (MLR) models for predicting the PM10 concentration were then developed based on three different prediction hours, namely t+1, t+2, and t+3. The PM10, t+1 model was the best MLR model to predict PM10 during transboundary haze events compared to PM10,.t+2 and PM10,t+3 models, having the lowest percentage of total error (28%) and the highest accuracy of 46%. A better prediction and explanation of PM10 concentration will help the authorities in getting early information for preserving the air quality, especially during transboundary haze episodes.

Keyword: Transboundary haze; Prediction; Multiple linear regression; Accuracy; Error; Malaysia