Application of extensive air monitoring network: Development of National Land Use Regression for Air Pollution Exposure in China

Jin Li[1], Huan Liu[1], Chufan Wang[1]

[1] *School of Environment, Tsinghua University, Beijing, China*

Air pollution has increasingly become a hot focus by society and accurate assessment of air pollution exposure is required to understand its adverse health effects and improve public policy. Most previous assessment such as AOD model, air quality model or interpolation method have somewhat limitations for their appilication in higher resolution and precision. Land Use Regression (LUR) is an empirical statistical method combining air quality monitoring measurements and surrounding relevant information to estimate pollutants exposure for nonmeasurement locations and it has been recently applied throughout many regions typically for urban inner scale. Recently, some research demostrated national LUR was a powerful tool able to provide fine-scale estimates of pollution exposure scale over large geographic range. Ministry of Environmental Protection of China began to open the air quality data since January 2013 and rapid development of Chinese national monitoring network may provide necessary requirements of national LUR in China, but as for our knowledge, basically national LUR model across China has not yet been reported.

In this study, two typical pollutants ($PM_{2.5}$ and NO_2) were selected to represent air pollution. We collected hourly pollutants concentrations from all available national sites on January 1, 2013 to December 31, 2016 and several types of predictor variables including geographical position, emission inventory, satellite estimates, road network, socioeconomic indicators, land use, metorologic condition. A total of 213 independent variables (point type and buffer type) were finally generated to develop our model through a two-procedure LUR building approach consisting of Lasso variables selection and multiple stepwise forward regression. In addition, 10-fold cross-validation was used to test model's robustness. We also generated $PM_{2.5}$ and NO_2 exposure mapping based on our final model and analysed spatialtemporal variety of air pollution across China.

Our model has relatively better performance with adjust R^2 0.72 for $PM_{2.5}$ and 0.74 for NO_2. There are some differences between structures of these two models due to different formation methanisms. The variables kept in final model are satellite estimates, wind speed, precipitation, year and impervious_6000m for $PM_{2.5}$ and satellite NO_2 column, impervious_10000m, latitude, wind speed and road lenth_4000m for NO_2. Population-weight. Satellite data both play the most important roles in final model, demontrating that it could effectively solve the limitation of LUR constrained to individual cities. The relatively smaller reduction (2.8% for $PM_{2.5}$ and 3.1% for NO_2) of model performance results in cross validation suggests robustness of our model. From generated mapping based LUR model, pollution concentration areas are distributed in economically developed places including Beijing-Tianjin-Hebei (BTH) region, the Pearl River Delta, the Yangtze River Delta and the Sichuan Basin. Average population-weight concentration in 2013-2016 were 34.7 $\mu g/m^3$ for NO_2 and 62.2 $\mu g/m^3$ for $PM_{2.5}$. Our study estimates air pollution exposure in China based on national LUR model and could provide fundamental data for epidemiological research and air quality policy in China.

Key words: air pollution exposure, land use regression, air monitoring network.