

Land use regression application to predict the PM₁₀ in São Paulo, Brazil

Mateus Habermann, *Faculdade de Medicina da Universidade de São Paulo, Brazil*

Nelson Gouveia, *Faculdade de Medicina da Universidade de São Paulo, Brazil*

Rogério Prado, *Faculdade de Medicina da Universidade de São Paulo, Brazil*

Background and Aims: land use regression is a method able to predict air pollution concentrations in a specifically location, combining variables related to the concentration and dispersion of pollutants that can influence the exposure. Therefore we aimed to construct a land use regression model to predict the environmental concentration of PM₁₀ in São Paulo, Brazil.

Methods: The study was based on the 2007 average of PM₁₀ measured in monitoring stations in São Paulo, southeast Brazil. We used GIS applications to obtain geographic, demographic, roads and land use data to compose the regression model of the monitoring locations vicinity in buffers from 250 to 1.000 meters. First we calculated linear regression between each independent variables and PM₁₀. We selected most robust variables ($p \leq 0.2$) without collinearity to compose the multivariate linear regression. The final model was used to predict particulate matter in the city and mapping using interpolation method.

Results: The municipality has 12 PM₁₀ monitoring stations, but in 2007 only 09 stations had sufficient valid measurements to construct the model. The annual average of PM₁₀ was 40.1 $\mu\text{g}/\text{m}^3$ ($\pm 4.6 \mu\text{g}/\text{m}^3$). For each location we obtained 113 independent variables. Four variables remained in the multiple regression model. The final model explained 63.8% of the MP₁₀ variance with only one predictor variable, which was the average of mild traffic until 250 m from the monitoring stations. The model was applied in a grid of points 1 x 1 km and interpolation to mapping the dispersion of PM₁₀ in the city.

Conclusions: This method is fast and easy to develop and may be used to predict air pollution in epidemiological studies in São Paulo when there is no available data from measurements. However, the results must be interpreted with caution because the model was based on measurements of only nine monitoring stations.