Geophysical Research Abstracts
Vol. 16, EGU2014-3743, 2014
EGU General Assembly 2014

Development and Application of Nonlinear Land-Use Regression Models

Alexandre Champendal (1), Mikhail Kanevski (1), and Pierre-Emmanuel Huguenot (2)

(1) Institute of Earth Surface Dynamics (IDYST), University of Lausanne, Switzerland (Mikhail.Kanevski@unil.ch), (2) Service de l'air, du bruit et des rayonnements non ionisants (SABRA), Geneva Switzerland

The problem of air pollution modelling in urban zones is of great importance both from scientific and applied points of view. At present there are several fundamental approaches either based on science-based modelling (air pollution dispersion) or on the application of space-time geostatistical methods (e.g. family of kriging models or conditional stochastic simulations). Recently, there were important developments in so-called Land Use Regression (LUR) models. These models take into account geospatial information (e.g. traffic network, sources of pollution, average traffic, population census, land use, etc.) at different scales, for example, using buffering operations. Usually the dimension of the input space (number of independent variables) is within the range of (10-100). It was shown that LUR models have some potential to model complex and highly variable patterns of air pollution in urban zones. Most of LUR models currently used are linear models.

In the present research the nonlinear LUR models are developed and applied for Geneva city. Mainly two nonlinear data-driven models were elaborated: multilayer perceptron and random forest. An important part of the research deals also with a comprehensive exploratory data analysis using statistical, geostatistical and time series tools. Unsupervised self-organizing maps were applied to better understand space-time patterns of the pollution.

The real data case study deals with spatial-temporal air pollution data of Geneva (2002-2011). Nitrogen dioxide (NO_2) has caught our attention. It has effects on human health and on plants; NO_2 contributes to the phenomenon of acid rain. The negative effects of nitrogen dioxides on plants are the reduction of the growth, production and pesticide resistance. And finally, the effects on materials: nitrogen dioxide increases the corrosion. The data used for this study consist of a set of 106 NO_2 passive sensors. 80 were used to build the models and the remaining 36 have constituted the testing set. Missing data have been completed using multiple linear regression and annual average values of pollutant concentrations were computed. All sensors are dispersed homogeneously over the central urban area of Geneva. The main result of the study is that the nonlinear LUR models developed have demonstrated their efficiency in modelling complex phrenomena of air pollution in urban zones and significantly reduced the testing error in comparison with linear models.

Further research deals with the development and application of other non-linear data-driven models (Kanevski et al. 2009).

References

Kanevski M., Pozdnoukhov A. and Timonin V. (2009). Machine Learning for Spatial Environmental Data. Theory, Applications and Software. EPLF Press, Lausanne.