The INLA approach for Bayesian air quality models

Michela Cameletti
Università di Bergamo, Bergamo (I) michela.cameletti@unibg.it

Epidemiological, clinical and toxicological studies carried out in the last 10 years have shown a significative link between particulate matter (PM) exposure and adverse health outcome in human. The European Environmental Agency has designated 2013 as the year of air (see http://www.eea.europa.eu/highlights/2013-kicking-off-the-2018year) in order to stress the seriousness of the problem and the need to take proper actions that improve air quality. Bayesian spatio-temporal models have been extensively used for air quality modelling thanks to the possibility to specify a hierarchical structure on the data and/or parameters, allowing also to account for similarities based on the neighbourhood or on the distance, for area-level or geostatistical data, respectively. The main challenge in Bayesian models resides in the computational costs required for implementing Markov Chain Monte Carlo (MCMC) methods, especially in case of complex models or massive datasets. Recently, the Integrated Nested Laplace Approximation (INLA) approach has been developed as a computationally efficient alternative to MCMC. INLA is a deterministic algorithm designed for latent Gaussian models, a very wide and flexible class of models ranging from (generalized) linear mixed to spatial and spatio-temporal models. In this work we implement spatio-temporal models with the INLA approach for modelling PM concentration in Piemonte region (Italy) and assessing the risk of respiratory disease. In the first step we model the exposure, by using as response variables PM measurements coming from a monitoring network and including in the latent field a spatio-temporal continuous process. Note that at this stage we are able to deal with the change of support problem and to predict the PM exposure level at the same areal resolution of the health data. In the second step, we implement a disease mapping model using the health outcome as response variable and including in the latent field a spatially structured random effect, an unstructured component and a fixed effect for the exposure level estimated at the first step. By estimating the relative risk of mortality or hospitalization, we are able to detect high risk areas where air pollution is a serious concern for human health.

Key words: disease mapping, change of support, areal data, geostatistical data.