222

Combining data from new and emerging data streams -- towards a general statistical modelling approach

This talk discusses statistical methodology that make optimal use of newly emerging and existing data streams by jointly modelling data acquired from several data sources. The approach operates across different spatial scales and resolutions by indexing the data in continuous space. Data collected at different spatial scales can be used within the same model through projecting it into continuous space and projecting it back into the spatial scale of interest.

The approach uses hierarchical Bayesian statistical models that combine data from different sources at potentially different spatial scales through a joint modelling approach. The approach operates within the general and flexible spatial statistics modelling framework INLA that is computationally efficient and hence fitting complex spatial models becomes feasible within a realistic time frame. We consider multi-layer models with several responses and associated likelihoods for the different types of data linked through one or several terms shared between the levels. Computationally, the approach allows a flexible representation of spatial and spatio-temporal structures by employing a spatially continuous approximation through the stochastic partial differential equation (SPDE) approach. In addition, observation processes specific to particular study are accounted for through methodology developed for the software package inlabru.

We will discuss the relevance of the methodology to a range of contexts by discussing applications to fisheries, epidemiological, land management and citizen science data.