COVARIATE SELECTION IN MULTIVARIATE SPATIAL ANALYSIS OF OVINE PARASITIC INFECTIONS

Annibale Biggeri[1,2], Dolores Catelan[1,2], Corrado Lagazio[3], Emanuela Dreassi[1], Laura Rinaldi[4], Vincenzo Musella[4], Vincenzo Veneziano[4], Giuseppe Cringoli[4]

1 Department of Statistics "G. Parenti" – University of Florence
2 Biostatistics Unit – CSPO Institute for Cancer Prevention
3 Department of Statistical Science – University of Udine
4 Department of Pathology and Public Health, University of Naples "Federico II", Naples

Multivariate spatial analysis of parasitic infections is usually applied to geographical information system (GIS) with a large number of covariates from remote sensing. We frame the problem of covariate selection in ecological analysis in a Bayesian perspective. Bayesian variable selection methods are attractive since posterior probability of each model and marginal inclusion probabilities of each covariate can be easily derived.
Data come from a cross-sectional study carried out in the Campania Region (Southern Italy) with the aim of investigating the spatial distribution of four helminths in 121 ovine farms. The four helminths investigated were: Fasciola hepatica, Calicophoron (Paramphistomum) daubneyi and Dicrocoelium dendriticum and the Haemonchus contortus.
A GIS of the Campania region, was constructed utilizing as data-layers:
(i) the topographic base map of the Italian Geographic Military Institute (ii) the digital aerial photos of the Cartographic Office of the Campania region, (iii) the monthly and annual Normalized Difference Vegetation Index (iv) the elevation, slope and aspect of the study area obtained from the Digital Elevation Model (v) the land use, obtained from satellite imagines and divided into 34 classes.(vi) the geolithological map of the study area divided into 11 classes.
Data on each of these variables were then extracted for 'pasturing area' of the sheep farms, previously geocoded. For each of the 121 ovine farms we have information on a large set of covariates to be used in ecological regression models. We a priori selected 22 from all the available covariates: 18 referring to land use, the other to the lithological field characteristics.
We specified a series of Multivariate Hierarchical models with square root transformation of egg counts as Multivariate Normal response. An improper Multivariate CAR (MCAR) is used to capture spatially structured variability. Bayesian variable selection has been address via Stochastic Variable Search with mixture priors.
Results are consistent with theoretical knowledge, quantified impact of common intermediate hosts and supported evidence of previously unknown ecological components.