Title:

Risk mapping based on hidden Markov random field and variational approximations

Authors & affiliations:

David abrial, inra clermont-ferrand, lamiae azizi: inria Grenoble and inra clermont-ferrand, Myriam charras garrido: inra clermont-ferrand and Florence Forbes: inria Grenoble

Abstract: (Your abstract must use **Normal style** and must fit in this box. Your abstract should be no longer than 500 words. The box will 'expand' over 2 pages as you add text/diagrams into it.)

The analysis of the geographic variation of disease and its representation on a map is an important tool in epidemiology which allows the detection of clusters of geographical areas characterized by homogeneity in the estimated relative risks and enables to understand the mechanisms which underlie the spread of the disease. Traditional approaches to disease mapping have some deficiencies and disadvantages in presenting the geographical distribution. The use, for example, of the crude Standardized Mortality Ratio (SMR) (The epidemiological measure under consideration) which is the ratio of observed cases y over expected cases e, is becoming unpopular because of its instability, especially when rare diseases are investigated in an area with a small population. In such a case, both the observed and the expected value are low. As a result, an area with small population tends to present an extreme SMR. Alternatively, Poisson-based models that smooth the risk estimates are proposed to overcome these deficiencies.The most popular approach in animal epidemiology is based on hierarchical Bayesian approaches designed for the estimation of risk at each geographical unit. The risk classification is performing manually, in a second time, by animal epidemiologists, with the difficult task of defining the risk ranges for each class. We propose an approach based on finite mixture models with spatial constraints, in which the risk classification is performed automatically. We assume that the counts follow a Poisson model and introduce a finite mixture model for the Poisson rates. The allocation to the mixture components is modeled through a hidden Markov Random Field (MRF) using a Potts presentation. Instead of using MCMC, which would be very time consuming, we introduce the computational implementation of our model via the EM mean field algorithm. EM solution can highly depends on its starting position. In this work, we propose a way of initialization working well for most situations arising in practice. This initialization strategy consists in randomly drawing initial parameters in an appropriate space including all possible EM trajectories. The performance of our model is examined on synthetic data and real data set concerning the Bovin Spongiform Encephalopathy (BSE) in France.