BAYESIAN MODELS FOR SPATIO-TEMPORAL ASSESSMENT OF DISEASE

SU YUN KANG

Epidemiological data are often characterised by a spatial and/or temporal structure. To adequately account for spatial and temporal dependence in these data, there are point-based and area-based spatial and spatio-temporal models in the literature. However, there is a lack of knowledge about the impact of modelling at different spatial scales, temporal scales and spatial structures. This is of practical interest for diseases such as cancer that can display high and low intensities over a geographical region, can be subjected to a range of socio-economic and other risk factors and can change in spatial pattern over time with demographic and other changes. Given the importance for epidemiologists to take into account the spatial correlation in a disease dataset using spatial smoothing techniques, the choice of spatial and temporal smoothness priors is an acknowledged challenge that motivates the current research. In view of the fact that the spatial and spatio-temporal models are hierarchical models in which inference and estimation are not trivial, the research is conducted using Bayesian techniques to facilitate the inference.

This thesis aims to explore, assess and provide guidance on the suitability of different spatial scales, spatial smoothness priors and temporal scales in an original and comprehensive way. We focus on a rich and flexible class of Bayesian spatial and spatio-temporal models. This research endeavours to fulfil the aim by addressing the following objectives.

Firstly, we discuss and evaluate a number of spatial models and their suitability for analysing various structures of spatial point patterns at the grid level. The study confirms that different models may be more appropriate for different structures of point patterns due to their varying complexity and flexibility. Spatially complicated datasets generally require a spatial prior with greater flexibility.

Secondly, we evaluate the impact of spatial scales and spatial smoothness priors for various structures of point level binary data. We illustrate the importance of repeating the spatial analyses at multiple spatial scales for a spatial dataset. It is shown in the study that different spatial smoothness priors are applicable for different spatial structures. The intrinsic Gaussian Markov random field (IGMRF) prior is recommended for spatial smoothing in spatially dense and inhomogeneous point patterns due to the spatial dependence among first-order neighbours. The second-order random walk on a lattice prior is a reasonable choice to smooth spatially sparse point data regardless of the level of inhomogeneity in the data. The Matérn model is very sensitive to changing spatial scale and has great flexibility in modelling spatially clustered point data.

Thirdly, we investigate the impact of spatial scales for various structures of Poisson count data. Complicated spatial patterns such as inhomogeneous point patterns and spatially clustered patterns appear to be more sensitive towards the changing spatial scales. The study confirms the importance of repeating the spatial analyses at multiple spatial scales in order to determine the best scale to analyse the data.

Fourthly, we develop a spatial model for analysing point level disease data using a geographically more relevant scale for spatial smoothing. It is found that finer grid cells perform better than statistical local areas (SLAs) for spatially sparse data while similar performance between fine grid cells and SLAs is observed for spatially dense data based on the following criteria: (a) the overall goodness of fit of the multilevel model and the resulting model selection using deviance information criteria and logarithmic score; (b) the resulting posterior estimation and inference for linear predictor and the model parameters; and (c) the identification of spatial/localised disease risks clustering using image plots.

Fifthly, we demonstrate the selection of an optimum temporal scale by evaluating the impact of the choice of temporal scales for modelling individual disease outcomes. The study shows that the model goodness of fit, predictive power and precision of estimation depend on the scale of temporal aggregation, particularly for the nonparametric model formulation. The parametric time trend, however, was less susceptible to the changing scale compared to the nonparametric time effect.

Finally, we provide guidance on the choice of spatial scales and spatial smoothness priors based on the aims of spatial smoothing for various structures of spatial point patterns. The recommendations are as follows. If the aim of investigation is to identify clusters, the first-order IGMRF prior is a reasonable choice as it allows for less spatial smoothing compared to two other priors, and the preferred spatial scales are those that show some degree of clustering in the data. When the aim is to smooth the spatial surface, either the second-order IGMRF on a lattice or the Matérn model is recommended, depending on the desired degree of smoothing. These two priors are ideal for the estimation of the surface of regression effect as they impose a higher level of smoothing than the first-order IGMRF prior. With respect to this aim, spatial scales that show randomness or less clustering in the data are preferred.

Using a rich class of Bayesian spatial and spatio-temporal models, we address interesting and crucial issues that are relevant to the applications of spatial and spatio-temporal modelling. The overall contribution of this research is the advancement of knowledge in spatial and spatio-temporal modelling through the increased understanding of spatial scales, smoothness priors and temporal scales in terms of their methodology and applications. This research is of particular significance to researchers seeking to understand and employ a range of spatial scales, smoothness priors and temporal scales in various disciplines.

SU YUN KANG, Mathematical Sciences School,
Queensland University of Technology,
Gardens Point Campus, GPO Box 2434,
Brisbane, Queensland 4001, Australia
e-mail: s7.kang@qut.edu.au