Spatio-temporal modelling of crime using low discrepancy sequences

Paul Brown[1], Chaitanya Joshi[1], Stephen Joe[1], Nigel McCarter[2]

[1] Department of Mathematics and Statistics, University of Waikato, Hamilton, New Zealand
[2] NZ Police Intelligence, Waikato, New Zealand

E-mail for correspondence: `ptb2@students.waikato.ac.nz`

Abstract: We perform spatio-temporal modelling of burglary data in order to predict areas of high criminal risk for local authorities. We wish to account for several spatio-temporal factors as latent processes to make the model as realistic as possible, thus creating a model with a large latent field with several hyperparameters. Analysis of the model is done using Integrated Nested Laplace Approximations (INLA) (Rue et al. 2009), a fast Bayesian inference methodology that provides more computationally efficient estimations than Markov Chain Monte Carlo (MCMC) methods.

Keywords: Bayesian inference; Crime modelling; Integrated nested Laplace approximations; Low discrepancy sequences; Spatio-temporal modelling

1 Introduction

Efficient use of police resources is vital for taking preventative measures against crime, as opposed to reactive measures. As such, intelligence-led policing, where data, analysis, and criminal theory are used to guide police allocation and decision-making, are becoming ever more popular and necessary (Ratcliffe, 2012). Given that crime and its associated factors occur within a geographical context that include both space and time (Fitterer et al., 2014) spatio-temporal modelling can lead to more accurate prediction of crime than modelling which does not take these factors into consideration. Spatio-temporal modelling under the Bayesian paradigm also allows more information to be included through prior knowledge from local authorities and officials.

---

2 Data and Methods

2.1 Burglary Data

The main dataset consists of residential, petty (under NZ$5000) burglaries from 2010 to 2015 in the Hamilton City region, New Zealand. All locations are geo-coded using the New Zealand Transverse Mercator (NZTM) northings and eastings. The region is bounded by an 11 kilometre (km) $\times$ 13 km rectangle, and is partitioned up in 1 km $\times$ 1 km cells, giving 143 cells in total (see Figure 1).

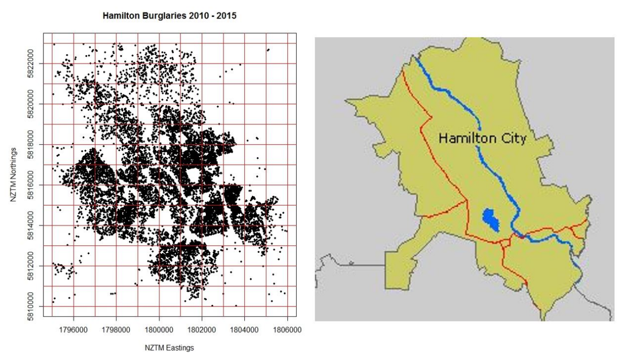

FIGURE 1. Hamilton burglaries from 2010 to 2015 and map of Hamilton, NZ.

There are several underlying factors involved with the spatial distribution of crime. Substantial research indicates that certain segments of the population and particular types of environments can generate high offender rates (Brown, 1982). The New Zealand Index of Deprivation (NZDEP) is a measure of socioeconomic status of a small geographical area, and includes a range of variables including income, employment and living spaces (see Atkinson et al. (2014) for full details). The social and physical environment is represented by off-licence liquor stores and graffiti respectively. Off-licence liquor stores, which may encourage a higher consumption of alcohol is also used as a measure of the perception of lawlessness and also of low anti-social behaviour. Incidence of graffiti has been used as a measure of lawlessness and environmental deterioration within an area.

2.2 INLA

INLA is a fast Bayesian inference methodology for latent Gaussian models that take the form

$$\eta_i = \beta_0 + \sum_{j=1}^{n_\beta} \beta_j z_{j,i} + \sum_{k=1}^{n_f} f^{(k)}(u_{k,i}) + \epsilon_i,$$

where the $\{\beta_k\}$'s are the linear effect on covariates $\mathbf{z}$, $\{f^{(\cdot)}\}$'s represent the unknown functions of the covariates $\mathbf{u}$, and ϵ_i's are the unstructured random errors. The latent Gaussian model is obtained by assigning all parameters in the latent field $\boldsymbol{\phi} = \{\beta_0, \{\beta_k\}, \{f^{(\cdot)}\}, \{\eta_i\}\}$ a Gaussian prior with mean $\mathbf{0}$ and precision matrix $Q(\boldsymbol{\theta})$, with hyperparameters $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_m\}$. Inference on $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ are made via the use of numerical integration, Laplace approximations and grid sampling strategies. For a full account of the inference stage, see Rue et al. (2009).

2.3 Modelling

Let $y_{i,t}$ be the count of burglaries in cell i at time t. Assume that $y_{i,t} \sim \text{Poisson}(\lambda_{i,t})$, and $\log(y_{i,t}) = \eta_{i,t}$. We have a generalised additive model

$$\eta_{i,t} = \beta_0 + \beta_1 \text{NZDEP}_i + r_i + g_i + l_i + Y_t + \epsilon_{i,t},$$

where β_0 is the overall mean, β_1 is a fixed linear parameter for the covariate NZDEP. The term r_i is the spatial effect of each cell, and is modelled as a Gaussian Markov Random Field (GMRF) with unknown precision τ_r.This specification, also known as conditionally autoregressive (CAR) prior, was introduced by Besag et al. (1991) and is used extensively in disease mapping. The terms g_i and l_i are incidence of graffiti and number of liquor stores respectively and are both modelled similarly to r_i, with hyperparameters τ_g and τ_l. The term Y_t represents the yearly time effect and is modelled as a random walk of order 1 (RW1) and has a hyperparameter τ_Y. Hence our latent parameters are $\boldsymbol{\phi} = \{\{\eta_{i,t}\}, \beta_0, \beta_1, \{r_i\}, \{g_i\}, \{l_i\}, \{Y_t\}\}$ with hyperparameters $\boldsymbol{\theta} = \{\tau_r, \tau_g, \tau_l, \tau_Y\}$.

3 Preliminary Results

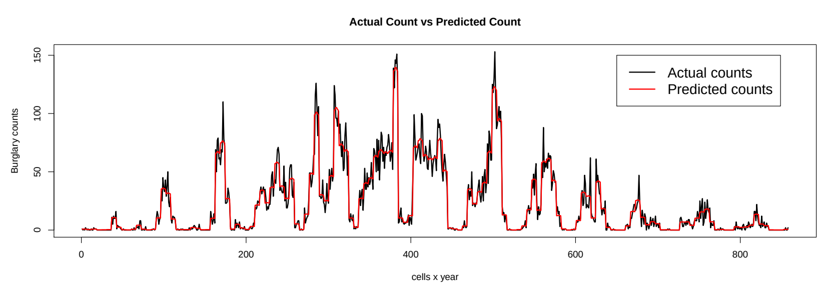

FIGURE 2. Actual count of burglaries per year vs. predicted count using INLA.

Figure 2 shows that the model predicts the actual counts well. Note that there are 143 cells, each with six years of data, giving the total of 858 cells for prediction. There is some variability between the years for each cell. There may be other temporal factors, such as seasonality, that may play an important role. Adding and developing the model, as well as model performance is currently being worked on.

4 Ongoing Work

There are many factors that may be involved in burglaries that we have not considered yet. Many of these factors may have some spatial or temporal effects, thus adding to the number of latent and hyper-parameters in the model. INLA could lose its computational efficiency for models with a large number of hyperparameters, therefore we need a methodology that can provide accurate and efficient inference on such a model. Recently, a paper by Joshi et al. (2016) has proposed using low discrepancy sequences instead of grids at the hyperparameter stage to increase computational gains and accuracy. In ongoing work, we are building a larger model with many hyperparameters using this approach.

References

Atkinson, J., Salmond, C., and Crampton, C. (2014). *NZDEP2013 Index of deprivation*. NZ Ministry of Health, http://www.health.govt.nz

Besag, J., York, J., and Mollië, A. (1991). Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, **43**, 1 – 59.

Brown, M.A. (1982). Modelling the spatial distribution of suburban crime. *Economic Geography*, **58:3**, 247 – 261.

Fitterer, J., Nelson, T.A., and Nathoo, F. (2014). Predictive crime mapping. *Police Practice and Research: An International Journal*, **16:2**, 121 – 135.

Joshi, C., Brown, P.T., and Joe, S. Fast Bayesian inference using low discrepancy sequences. *Bayesian Analysis*, (Submitted).

Ratcliffe, J. (2012). *Intelligence-led policing*. Portland, OR: Willian Publishing.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B*, **71:2**, 319 – 392.