Investigating the Stability of Public Confidence in the Police with Space-Time models.

Williams D[*1], Haworth J[†1], Blangiardo, M[**2] and Cheng T[‡1]

[1]Department of Civil, Environmental and Geomatic Engineering, University College London
[2]Department of Epidemiology and Biostatistics, Imperial College London

January 13, 2016

Summary

Public confidence in the police is essential for effective. Understanding the spatiotemporal trends in public confidence will better enable the police to carry out proactive policing interventions to improve it. Bayesian hierarchical methods provide a flexible framework for estimating public confidence in the police in spacetime. Public confidence in the police varies in space, time and interacts in spacetime. There is a need for further examination of whether these spacetime interaction patterns are stable. Adopting a mixture model approach to modelling the spacetime interactions will allow them to be studied closely and to be classified as stable or unstable.

KEYWORDS: space-time interactions, Bayesian hierarchical models, policing, public confidence

1. Introduction

Public confidence in the police is essential for effective policing (Jackson et al. 2013). Recent studies have established the spatiotemporal variability in public confidence in the police (Williams, Haworth, and Cheng 2015), as well as the usefulness of a Bayesian hierarchical approach to public confidence estimation and prediction (Williams, Haworth, and Cheng 2016). This study takes a step further in evaluating the usefulness of a mixture modelling approach to the Bayesian hierarchical modelling for exploring the spatiotemporal instability in public confidence in the police.

2. Bayesian Hierarchical modelling of spacetime instability

Our spatiotemporal Bayesian hierarchical model is comprised of a spatial component which estimates stable or "predictable" spatial patterns, a temporal component which estimates stable or "predictable" temporal patterns and a dynamic spatiotemporal interaction component. In this section we describe the general modelling framework as well as the mixture modelling approach which allows us to distinguish between the stable trends and the deviations from the stable patterns.

2.1. Study Area and Data

London is the capital city of the United Kingdom, located to the south-east of England. For administration, the city is divided into thirty-two boroughs. The Mayor's Office for Policing and Crime (MOPAC) and the Metropolitan Police (MET) have divided the city into 107 larger units of operational significance called borough neighbourhoods that consist of two or three wards (Mayor's Office for Policing and Crime 2016).This is the spatial unit used.

The Metropolitan Police Public Attitudes Survey (PAS) is a rolling cross-sectional survey conducted since 1983 (Mayor's Office for Policing and Crime, 2014b). It collects individual attitudes to the police

---

[*] dawn.williams.10@ucl.ac.uk

[†] j.haworth@ ucl.ac.uk

[**] m.blangiardo@imperial.ac.uk

[‡] tao.cheng@university.edu

and crime and is comparable to the Crime Survey for England and Wales. The survey is performed on a rolling basis, with results reported quarterly and annually. The data used consists of 36 quarters from April 2006 (Q5) to January 2015 (Q40). At the most recently released tally in September 2016, 69% of Londoners' think the MPS do a 'good job' in their area" (MOPAC 2016). Confidence in the police has been on the decline since January 2014.
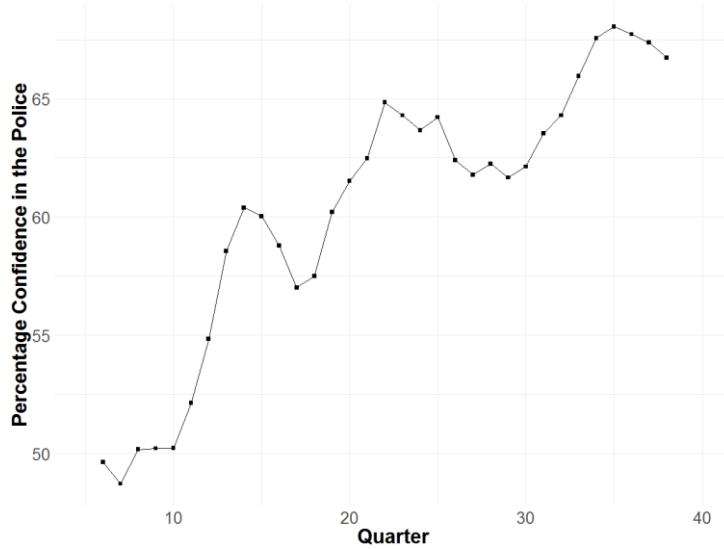

Figure 1 London-wide percentage confidence in the police

2.2. Spacetime Bayesian Hierarchical modelling of count data

The data consists of observed number of confident respondents y_{it}, the total number of respondents, n_{it} for a neighbourhood $i = 1, …, I$, an quarter $t = 1,…,T$. Conducting an attitude survey can be considered a Bernoulli trial, therefore a binomial model with a logistic link is appropriate. In the first level model of the model the binomial likelihood of the data is used to describe the within area variability of the counts conditional on the unknown risk parameters. The second level of the model specifies the space-time structure and parameterizes the unknown risk parameters with a prior distribution. In specifying the space-time structure, separate parameters are required to capture the stable space-time trends and the space-time interactions. A mixture model will be used to represent the space-time interactions to allow the atypical spatiotemporal interactions to be identified. The third level of the model specifies prior distributions on the hyperparameters, which are the unknown parameters from level 2.

At the first level of the model, we specify a binomial likelihood for the within neighbourhood variability of the counts:

$$y_{it} \sim binomial(n_{it}, \pi_{it}) \tag{1}$$

where π_{it} is the probability of a positive response in a particular neighbourhood i in quarter t. The second level of the model translates this to the logit scale and expresses it as an additive combination of the overall probability α, overall spatial random effects, λ_i, overall temporal random effects, ξ_t, and space-time interactions δ_{it}.

$$logit(\pi_{it}) = \alpha + \lambda_i + \xi_t + \delta_{it} \tag{2}$$

Prior distributions on the parameters are used here to describe how information can be borrowed across space and time to more accurately represent the underlying spatiotemporal risk structure. The Intrinsic Conditional Autoregressive (ICAR) prior (Besag and Kooperberg 1995) a commonly used structure which allows smoothing based on spatial or temporal adjacency was chosen for both the spatial and temporal random error components.

The convolution BYM model introduced by Besag, York and Mollié (1991) will be used to specify the structure of the predictable spatial and temporal trends. The convolution BYM model allows for a structured and unstructured random component as follows:

$$\lambda_i \sim N(\mu_i, \sigma_\lambda^2)$$
$$\mu_i \sim \text{CAR}(\boldsymbol{W}, \sigma_\mu^2)$$

As the BYM model is used to specify both the spatial and temporal random errors level 2 of the model is as follows:

$$logit(\pi_{it}) = \alpha + \lambda_i + \xi_t + \delta_{it}, i = 1, \dots, I; t = 1, \dots, T$$
$$\lambda_i \sim N(\mu_i, \sigma_\lambda^2), i = 1, \dots, I$$
$$\mu_i \sim \text{CAR}(\boldsymbol{W}, \sigma_\mu^2)$$
$$\xi_t \sim N(\gamma_t, \sigma_\xi^2), t = 1, \dots, T$$
$$\gamma_t \sim \text{CAR}(\boldsymbol{Q}, \sigma_\gamma^2) \tag{3}$$

An improper uniform prior on the whole real line was used for the intercept(α). Level 3 treats the variance parameters from level 2 as unknowns to be estimated. These unknowns, called hyperparameters, are given distributions. A $\text{Gamma}(a, b)$ distribution was used as the hyperprior distribution for σ_U^2, σ_S^2, σ_ν^2 and σ_δ^2 where $a = 0.5$ and b equal 0.0005. The Gamma distribution is a commonly used prior distribution for hyperparameters because it is very non-informative (Law, Quick, and Chan 2013). Additionally, for identifiability the sum of the vectors μ and γ was constrained to zero.

2.3. Investigating space-time instability

A mixture model will be used to represent the space-time interactions in a variation of the two-stage classification approach proposed by Abellan, Richardson, and Best (2008). The use of a mixture model allows random noise in the interactions to be separated from legitimate deviations from the norm (Abellan, Richardson, and Best 2008) i.e. the unstable interactions. The flexible prior structure of the mixture models makes them useful for classification problems such as this one (Richardson and Green 1997). The space-time interaction term can then be represented as follows:

$$\delta_{it} \sim \rho \text{Normal}(0, \tau_1^2) + (1 - \rho)\text{Normal}(0, \tau_2^2) \tag{4}$$

A uniform prior on [0,1] was used for ρ· while half-normal (i.e. the normal bounded by zero and positive infinity) hyperprior distribution was used for the standard deviations τ_1 and τ_2. τ_1 was kept small to enable shrinkage of noise while τ_2 allowed for a much larger range of values.

$$\tau_1 \sim \text{Normal}(0, 0.01)I_{(0,+\infty)}$$
$$\kappa \sim \text{Normal}(0, 100)I_{(0,+\infty)}$$
$$\tau_2 = \tau_1 + \kappa \tag{5}$$

A latent allocation variable, z_{it}, is then defined and used to determine whether the interactions are just noise captured by the $\text{Normal}(0, \tau_1^2)$ distribution or are a legitimate deviation from the overall trend captured by the $\text{Normal}(0, \tau_2^2)$ distribution. A posterior probability, ρ_{it}, of the interaction parameter, δ_{it}, being of either group can then be calculated using the allocation variable. These posterior probabilities can then be used to classify the interaction parameter point as either a "stable" or "unstable" space-time pattern.

Two rules for classification of the patterns a "stable" or "unstable" were used in this study:
- Rule 1: An area is considered unstable if it is deemed likely to be unstable for 40% of the time periods in the study
- Rule 2: An area is considered unstable if it is unstable for a at least three consecutive time periods

2.4. Implementation

The three models were implemented in WinBUGS, Windows based freeware package for building and analysing Bayesian probability models with Markov chain Monte Carlo techniques (Lunn et al. 2000). In all cases, two Monte Carlo chains were used, and 10,000 iterations were performed after convergence. Convergence of the chains was checked with three measures: visually with trace plots and the visual representation of the Gelman-Rubin $\hat{R}$ statistic as well as by ensuring that the Monte Carlo error was less than 5% of the posterior standard deviation.

2.5. Evaluation framework

Over-parameterization of the model can lead to loss of precision and incorrectly estimating the stable spatial and temporal patterns (Abellan, Richardson, and Best 2008). Correlation analysis was conducted to determine whether the stable spatial and temporal patterns estimated were comparable to a purely spatial and purely temporal model. The Deviance Information Criterion (DIC), a likelihood based measure of model complexity (Spiegelhalter et al. 2002), was also used to evaluate over-parameterization.

3. Results & Conclusion

Figures below show the stable spatial and temporal trends estimated. This maps shows that areas in south-west of London are more likely than the London wide average to be confident in the police. The empirical temporal trend is well captured.
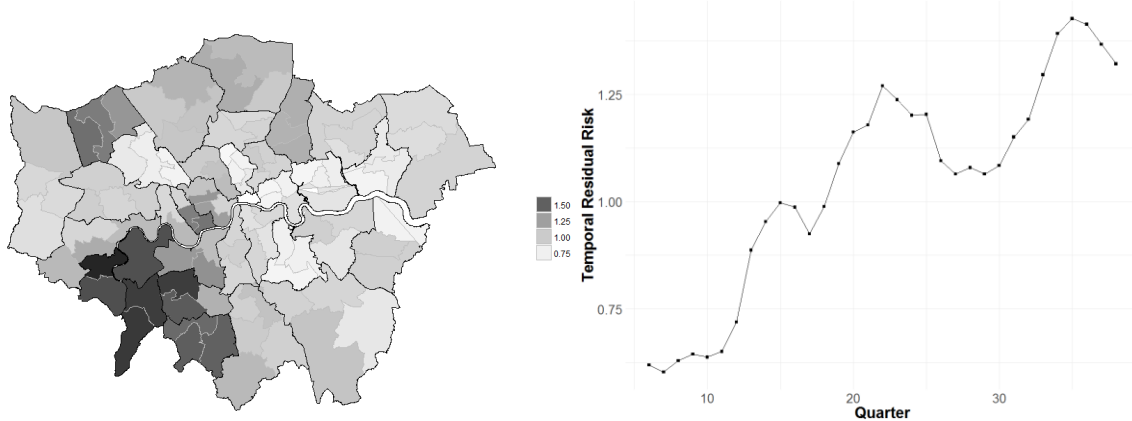

Figure 2 Map of stable spatial 'risk' of confidence (left), graph of stable temporal 'risk' of confidence (right)

Correlation analysis confirmed that the stable spatial and temporal patterns estimated were highly correlated with those estimated by the purely spatial and purely temporal models. This suggests that the additional model complexity does not lead to loss of power in determining the stable trends. The DIC values are presented in table 1. p_D represents the effective number of parameters and is an indicator of model complexity, while $\bar{D}$ is an indicator of model fit. The lower DIC value for the complex space-time mixture model indicates greater parsimony despite the increased complexity.

Table 1 Model evaluation statistics

Model	$\bar{D}$	p_D	DIC
Space-time Mixture Model	18,408	3,256	21,664
Purely spatial benchmark model	34,227	258	34,485
Purely temporal benchmark model	34,291	167	34,458

Figure 3 maps the number of unstable interaction terms over the time series. The prevalence of unusual spatiotemporal interactions is in general higher in West London and in areas with a geodemographic classification of "aspirational bustle".
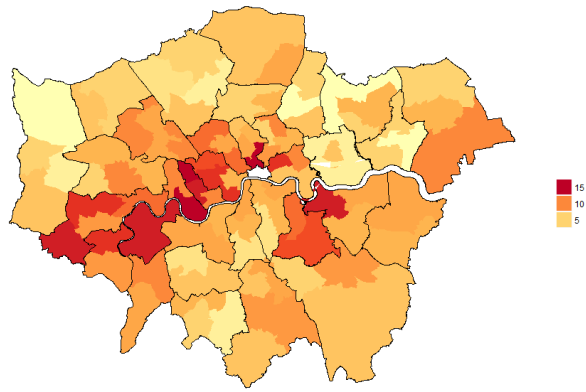
Figure 3 Map of the number of time periods classed as likely to be unstable

Figure 4 maps the output of our classification. Rule 1 selected neighbourhoods with more than 30% of time periods unstable while rule 2 deemed neighbourhoods unstable if they were deemed likely to be unstable for three consecutive time periods. Figure 5 maps presents a graph of the interactions for the unstable neighbourhoods with a stable neighbourhood, "Waltham Forest – Central" used as a benchmark.
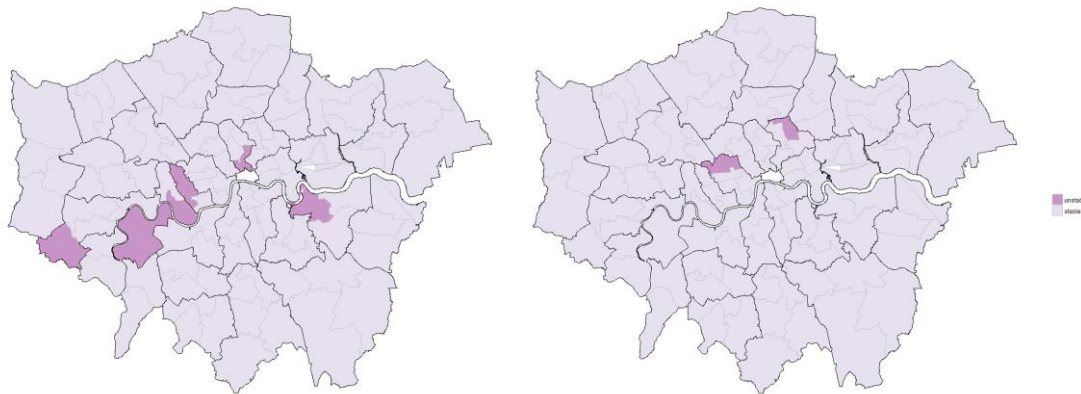

Figure 4 Neighbourhoods likely to be unstable (darker) as classed by rule 1 (left) and rule 2 (right)
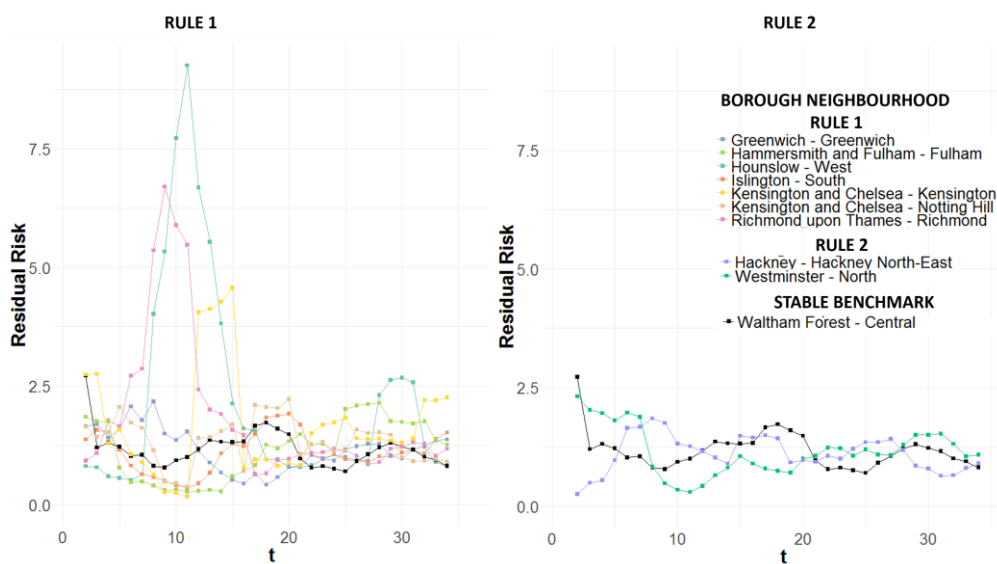

Figure 5 Graph of neighbourhoods likely to be unstable as classed by rule 1 (left) and rule 2 (right)

The difference between neighbourhoods which are frequently unstable and neighbourhoods consecutively unstable is quite marked. Rule 2 allows neighbourhoods with less extreme deviations from stability to be identified and can be useful in selecting neighbourhoods for interventions. These results confirm that this framework is flexible and fit for purpose. Further classification rules can be developed for use in various circumstances and explanatory covariates can be added to supply more context to the instability.

4. Acknowledgements

5. Biography

Dawn Williams is a PhD student at the SpaceTimeLab for Big Data Analytics (http://www.ucl.ac.uk/spacetimelab) at University College London. Her research interests include spatiotemporal data mining, Bayesian methods, visualization and big data analysis.

James Haworth is a lecturer in spatio-temporal analytics at SpaceTimeLab for Big Data Analytics (http://www.ucl.ac.uk/spacetimelab) at University College London. His main interests lie in the analysis, modelling and forecasting of spatio-temporal data using machine learning methods.

Marta Blangiardo is a senior lecturer in Biostatistics in the Department of Epidemiology and Biostatistics and part of the MRC-PHE Centre for Environment and Health at Imperial College London. Her reserach interests are the methodological aspects of environmental exposure estimation and on spatial and spatio-temporal models for disease mapping and for risk assessment. She has a special interest in novel computational methods for Bayesian inference and has recently written a book on the Integrated Nested Laplace approximation (INLA) for spatial and spatio-temporal applications.

Tao Cheng is a Professor in GeoInformatics and Director of the SpaceTimeLab for Big Data Analytics (http://www.ucl.ac.uk/spacetimelab) at University College London. Her research interests span network complexity, geocomputation, integrated spatio-temporal analytics and big data mining (modelling, prediction, clustering, visualisation and simulation), with applications in transport, crime, health, social media, and environmental monitoring.

References

Abellan, Juan Jose, Sylvia Richardson, and Nicky Best. 2008. "Use of Space–Time Models to Investigate the Stability of Patterns of Disease." *Environmental Health Perspectives* 116 (8): 1111–19. doi:10.1289/ehp.10814.

Besag, Julian, Jeremy York, and Annie Mollié. 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics." *Annals of the Institute of Statistical Mathematics* 43 (1): 1–20. doi:10.1007/BF00116466.

Jackson, Jonathan, Ben Bradford, Elizabeth A. Stanko, and Katrin Hohl. 2013. *Just Authority: Trust in the Police in England and Wales*. 1st ed. New York, NY: Routledge.

Law, Jane, Matthew Quick, and Ping Chan. 2013. "Bayesian Spatio-Temporal Modeling for Analysing Local Patterns of Crime Over Time at the Small-Area Level." *Journal of Quantitative Criminology* 30 (1): 57–78. doi:10.1007/s10940-013-9194-1.

Lunn, David J., Andrew Thomas, Nicky Best, and David Spiegelhalter. 2000. "WinBUGS - A Bayesian Modelling Framework: Concepts, Structure, and Extensibility." *Statistics and Computing* 10 (4): 325–37. doi:10.1023/A:1008929526011.

MOPAC. 2016. "Confidence Dashboard | London City Hall." https://www.london.gov.uk/what-we-do/mayors-office-policing-and-crime-mopac/data-and-research/confidence/confidence-dashboard.

Richardson, Sylvia., and Peter J. Green. 1997. "On Bayesian Analysis of Mixtures with an Unknown Number of Components (with Discussion)." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 59 (4): 731–92. doi:10.1111/1467-9868.00095.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde. 2002. "Bayesian Measures of Model Complexity and Fit." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 (4): 583–639. doi:10.1111/1467-9868.00353.

Williams, Dawn, James Haworth, and Tao Cheng. 2015. "Exploratory Spatiotemporal Data Analysis and Modelling of Public Confidence in the Police in Central London." In . Leeds, UK: The University of Leeds. http://leeds.gisruk.org/abstracts/GISRUK2015_submission_60.pdf.

———. 2016. "Predicting Public Confidence in the Police with Spatiotemporal Bayesian Hierarchical Modelling." In . Greenwich, London, UK.