University of Glasgow

Statistical challenges in estimating the long-term health impact of air pollution

Duncan Lee

Medical University of South Carolina

Acknowledgements

- This is joint work with Alastair Rushworth from the University of Glasgow and Sujit Sahu from the University of Southampton.

- The work is funded by the EPSRC grants EP/J017442/1 and EP/J017485/1.

- The main part of the work in this talk will appear in Biometrics under the title *A Bayesian Localised Conditional Autoregressive Model for Estimating the Health Effects of Air Pollution*.

EPSRC
Engineering and Physical Sciences Research Council

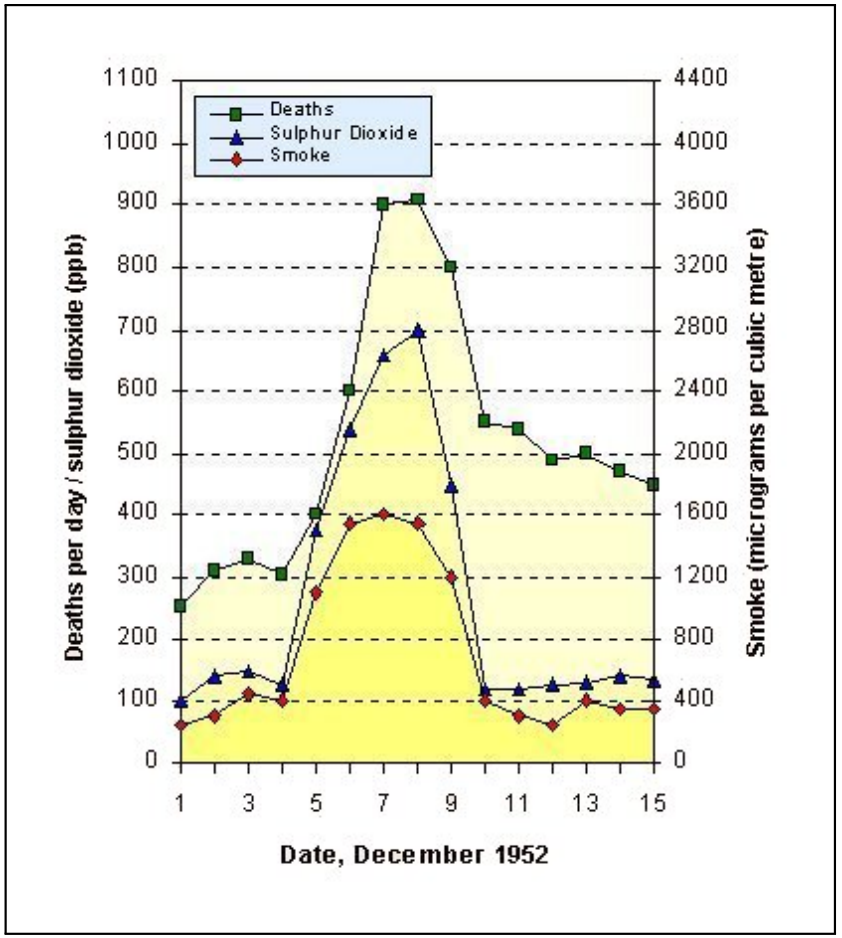
1. Background

- Air pollution has long been known to adversely affect public health, in both the developed and developing world.
- A recent report by the UK government estimates that particulate matter alone reduces life expectancy by 6 months, with a health cost of £19 billion per year.
- Epidemiological studies into the effects of air pollution have been conducted since the 1990s, with one of the first being that conducted by Schwartz and Marcus (1990) in London.
- Since 1990 a large number of studies have been conducted, which collectively have investigated the short-term and long-term health impact of air pollution.

A brief history

The relationship between air pollution exposure and mortality came to prominence during high air pollution episodes in:

- the Meuse Valley, Belgium in 1930;
- Donora, Pennsylvania in 1948; and
- London in December 1952.



High pollution

Government action

- Clean air acts in 1956, 1968 and 1993
  - Prohibited and regulated pollution sources.
  - Set up *'smoke control areas'* in which it was prohibited to emit smoke from buildings or chimneys.
- UK air quality strategy 1997, 2000, 2003 and 2007
  - Set target limits for annual or daily average concentrations for a number of common pollutants.
- Set up the Committee on the Medical Effects of Air Pollution (COMEAP).

Pollution is still a problem today

BBC Sign in | News | Sport | Weather | iPlayer | TV | Ra

NEWS UK

Home | World | UK | England | N. Ireland | Scotland | Wales | Business | Politics | Health | Education | Sci/En

3 April 2014 Last updated at 22:15

Share

Air pollution: Forecasters hope for cleaner air on Friday

People with lung and heart problems have been advised to avoid strenuous outdoor activity

Study designs

Pollution legislation continues to be informed by epidemiological studies investigating both the short-term and long-term health effects of air pollution exposure.

Acute studies investigate the effects resulting from a few days of high exposure.

- e.g. NMMAPS in the USA, Dominici et al (2002) and APHEA in Europe, Katsouyanni et al (2001).

Chronic studies investigate the cumulative effects of exposure over numerous years

- e.g. Dockery et al (1993) in six US cities, and Elliot et al (2007) in the UK.

Chronic studies

There are two main study designs when investigating the effects of long-term exposure to air pollution.

Cohort studies e.g. The Six Cities study by Dockery *et al* (1993) and the American Cancer study by Pope *et al* (2002), which relate average air pollution concentrations to the health status of a large pre-defined cohort of people.

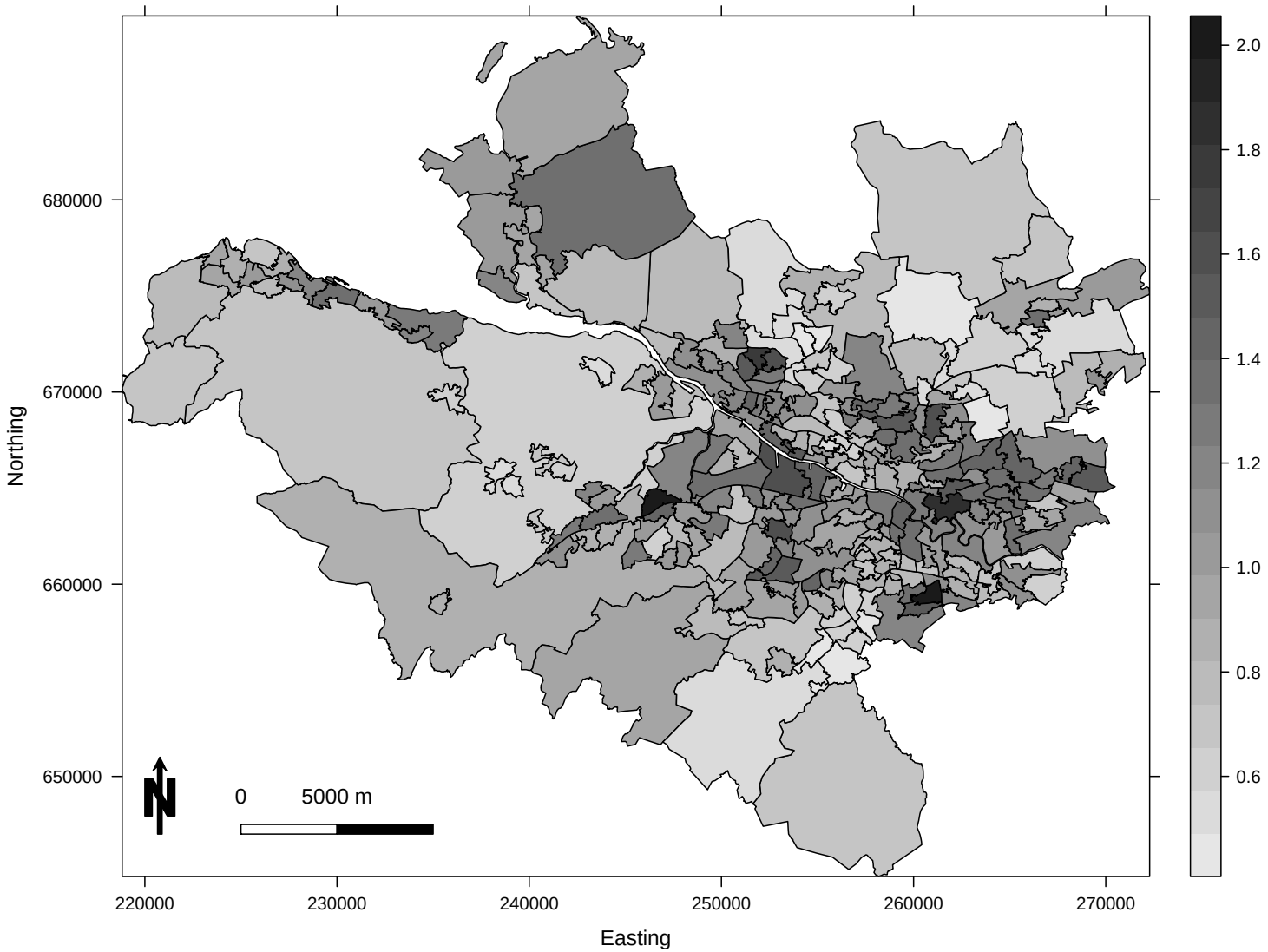
Ecological studies e.g. Elliot et al (2007) and Lee et al (2009), which relate average air pollution concentrations in contiguous small areas (such as electoral wards), against yearly numbers of health events from the population living in that area.

Ecological study design

- Small area studies have an ecological design, because the data relate to populations living in a set of n non-overlapping areal units, rather than to individuals.
- Examples of such studies include Jerrett *et al.* (2005), Elliott *et al.* (2007), Lee *et al.* (2009) and Greven *et al.* (2011).
- The health data are denoted by $\mathbf{Y} = (Y_1, \ldots, Y_n)$ and $\mathbf{E} = (E_1, \ldots, E_n)$, which are the observed and expected numbers of disease cases in each areal unit over a year.
- The expected numbers of cases are computed using external standardisation, based on age and sex specific disease rates.

Example - Greater Glasgow, Scotland

Respiratory hospitalisation risk - $\text{SIR}_k = Y_k/E_k$.



Covariate data

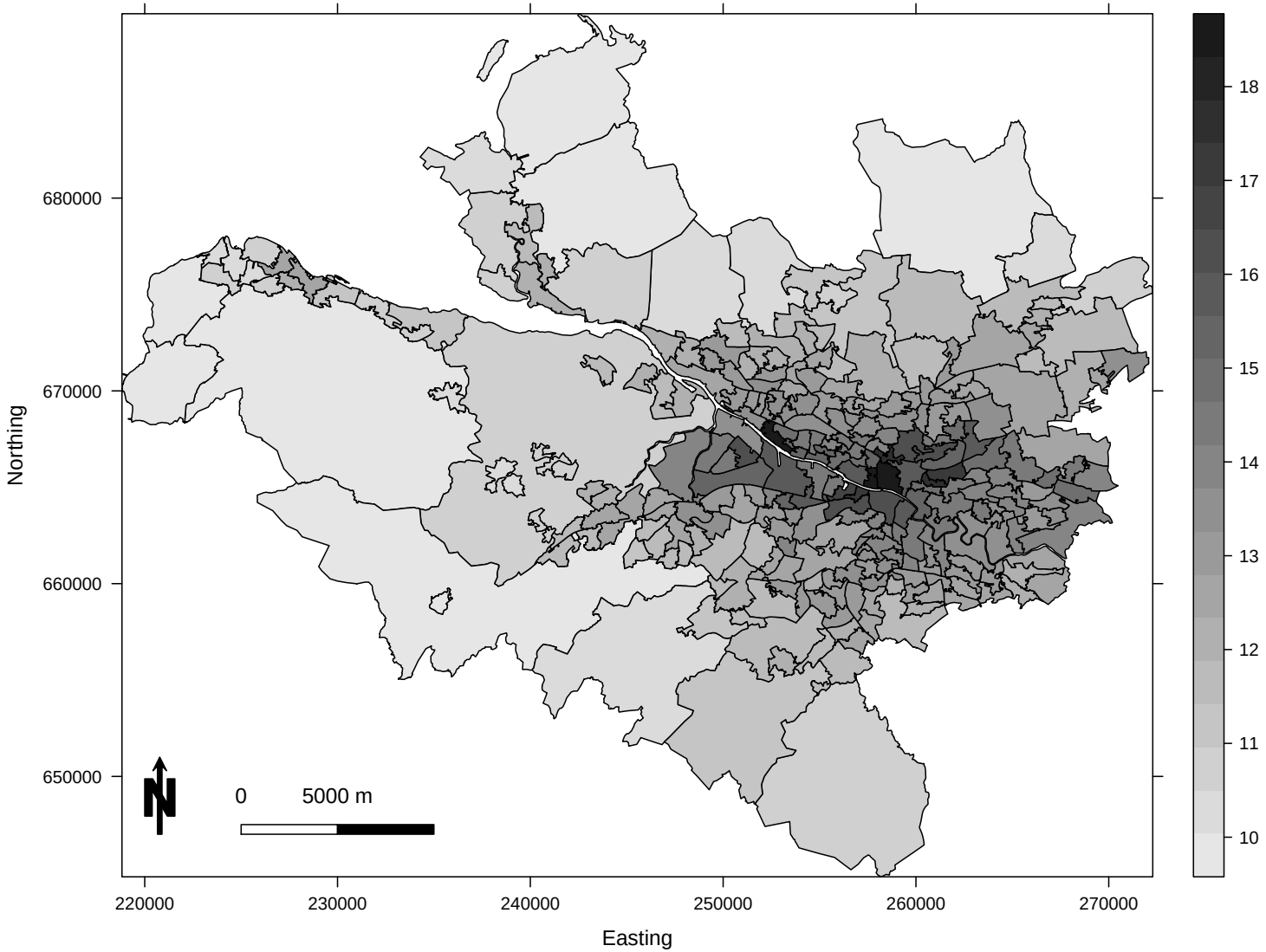
The covariates required for such a study are contained in an $n \times p$ matrix $X = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, and come in two main types:

1. Annual average air pollution concentrations for each small area. Typically, these are obtained from atmospheric computer models, as the network of monitoring sites is not dense at the small area scale.

2. Data on confounding factors, such as socio-economic deprivation which acts as a proxy for risk inducing behaviour such as smoking.

Example - Greater Glasgow, Scotland

Modelled particulate matter concentrations (PM_{10}).



Modelling

A Poisson Generalised Linear Mixed Model is given by:

$$\begin{aligned} Y_k &\sim \text{Poisson}(E_k R_k), \\ \log(R_k) &= \mathbf{x}_k^{\mathrm{T}} \boldsymbol{\beta} + \phi_k + \theta_k, \end{aligned}$$

where

- R_k quantifies disease risk in area k, so $R_k = 1.2$ means a 20% increased risk of disease.
- $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)$ are random effects to model residual spatial autocorrelation not captured by the covariates.
- $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ are random effects to model non-spatial variation (overdispersion) where $\theta_k \sim \text{N}(0, \sigma^2)$.

A Bayesian approach is adopted, using MCMC simulation.

Modelling spatial correlation

Conditional Autoregressive (CAR, *Besag et al. (1991)*) models are typically specified to capture the spatial autocorrelation, and can be written as a set of n univariate full conditional distributions $f(\phi_k|\boldsymbol{\phi}_{-k})$ for $k = 1, \ldots, n$ as:

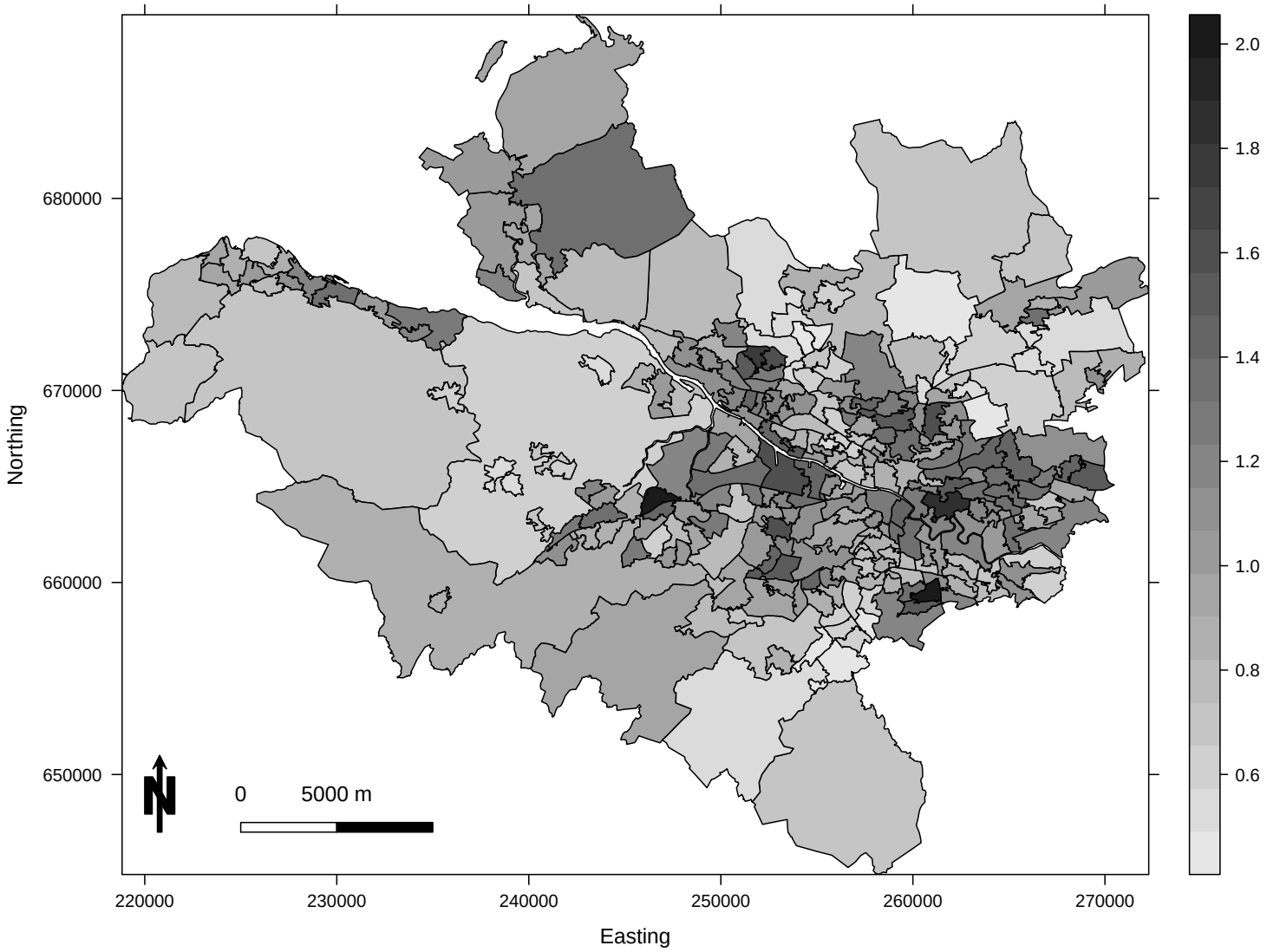
$$\phi_k|\boldsymbol{\phi}_{-k}, \tau^2, W \sim \mathrm{N}\left(\frac{\sum_{i=1}^{n} w_{ki}\phi_i}{\sum_{i=1}^{n} w_{ki}}, \frac{\tau^2}{\sum_{i=1}^{n} w_{ki}}\right).$$

Here $W = (w_{ki})$ is a binary $n \times n$ neighbourhood matrix, with $w_{ki} = 1$, denoted $k \sim i$ if areal units (k, i) share a common border and $w_{ki} = 0$ otherwise. The combination $\phi_k + \theta_k$ is known as the convolution or BYM CAR model.

Limitations - random effects

- The BYM CAR prior forces the sum $\phi_k + \theta_k$ to be globally spatially smooth, as the relative sizes of ϕ_k and θ_k are controlled globally by the relative sizes of (τ^2, σ^2).
- Therefore as health data are typically spatially correlated, the random effects sum $\phi_k + \theta_k$ is typically spatially smooth.
- This leads to problems of collinearity with covariates that are also smooth such as air pollution, as was illustrated by Reich *et al.* (2006) and Hughes and Haran (2013).
- Furthermore, the residual spatial structure is unlikely to be globally smooth, because the disease data (e.g. the SIR) are not globally smooth so the residuals after removing covariate effects are also unlikely to be.

SIR again - localised smoothness



Limitations - Pollution data

The pollution data are typically estimated concentrations from a dispersion model, which has the following limitations:

1. They are assumed to be true known measurements where as they are in fact subject to error and potential biases.
2. The uncertainty due to them being estimates rather than known data should be accounted for in the model.
3. They are estimated on a regular grid and ad-hoc approaches are typically used to re-align them to the areal units for which health data are available.

Thus spatio-temporal data fusion methods should be used to combine the modelled data with the available monitoring data.

Research aims

- In this talk we propose a solution to the first of these problems, namely the insufficient flexibility of the globally smooth BYM CAR prior.
- Our extension is an extension of CAR models to allow for localised spatial smoothness in the random effects surface, that is, subregions of spatial smoothness separated by step-changes.
- The other known problem with the BYM model is that only the sum of the two random effects $\phi_k + \theta_k$ are identifiable from the data and not the individual components. Thus we only utilise a single set of random effects ϕ in our methodology.

2. Methodology - motivation

The CAR model can also be written in a multivariate form as

$$\boldsymbol{\phi} \sim \mathrm{N}(\mathbf{0}, \tau^2 Q(W)^-),$$

where $Q(W) = \mathrm{diag}(W\mathbf{1}) - W$, is a singular precision matrix. Then from multivariate Gaussian theory we have that:

$$\mathrm{Corr}[\phi_k, \phi_j | \boldsymbol{\phi}_{-kj}, \mathbf{W}] = \frac{w_{kj}}{\sqrt{(\sum_{i=1}^{n} w_{ki})(\sum_{i=1}^{n} w_{ji})}}$$

So that if $w_{kj} = 1$ (corresponding to adjacent areal units) then the random effects (ϕ_k, ϕ_j) are partially correlated, while if $w_{kj} = 0$ (corresponding to non-adjacent areal units) then (ϕ_k, ϕ_j) are conditionally independent.

Motivation continued

- Therefore we treat the set $\mathcal{W} = \{w_{kj} | k \sim j, k > j\}$ as binary random variables, rather than w_{kj} being fixed at 1.
- This is because if $w_{kj} = w_{jk} = 1$ then (ϕ_k, ϕ_j) are correlated and are smoothed over, where as if $w_{kj} = w_{jk} = 0$ they are conditionally independent and are not smoothed over.
- This allows for smoothness in the random effects surface between some pairs of areal units, while between others there can be a step change.
- We follow the terminology of graphical models and refer to $w_{kj} \in \mathcal{W}$ as *edges*, and define any edge w_{kj} that equals zero as having been removed.

General approach

Our methodological innovation is a *Localised Conditional AutoRegressive (LCAR)* prior, which decomposes the joint distribution for an extended set of random effects $\tilde{\phi}$ and the set of edges $\mathcal{W}$ as

$$f(\tilde{\phi}, \mathcal{W}) = f(\tilde{\phi}|\mathcal{W})f(\mathcal{W})$$

Note, a standard CAR model consists of the first of these distributions $f(\phi|\mathcal{W})$, while $\mathcal{W}$ is assumed fixed and hence does not have a prior distribution.

Problem

The simplest modelling extension would be to specify independent Bernoulli priors for the edges in $\mathcal{W}$, that is

$$f(\mathcal{W}) = \prod_{k \sim j} \text{Bernoulli}(w_{kj}|p_{kj}).$$

However, this leads to overparameterisaton, as the number of partial correlation parameters in $\mathcal{W}$ is much larger than the number of data points n. For example, in the Glasgow motivating example, we have $n = 271$ and $|\mathcal{W}| = 718$.

Also, you are attempting to estimate a flexible spatial autocorrelation structure from only one realisation of the spatial process!

Existing solutions

This area of research is linked to the area of spatial epidemiology known as *Wombling*, whose aim is to identify boundaries (step changes) in the spatial pattern of disease risk.

- Lu *et al. (2007)* proposed a logistic regression model for the elements in $\mathcal{W}$, where the covariate measured the dissimilarity between areal units $(\mathcal{A}_k, \mathcal{A}_j)$.
- Lee and Mitchell *(2013)* proposed an iterative algorithm in which $\mathcal{W}$ is updated deterministically based on the joint posterior distribution of the remaining model parameters.

In contrast in an ecological regression context Hughes and Haran *(2013)* propose a smoothing model orthogonal to the covariates that does not treat $\mathcal{W}$ as random.

Proposed methodology

We describe our proposed methodology in two stages:

- We describe an extended CAR model $f(\tilde{\phi}|\mathcal{W})$ conditional on a fixed neighbourhood structure $\mathcal{W}$.
- We describe the prior distribution used for the neighbourhood structure, that is $f(\mathcal{W})$.

Prior distribution - $f(\tilde{\phi}|\mathcal{W})$

Recall the standard CAR prior is given by

$$\phi_k|\boldsymbol{\phi}_{-k}, \tau^2, W \sim \mathrm{N}\left(\frac{\sum_{i=1}^{n} w_{ki}\phi_i}{\sum_{i=1}^{n} w_{ki}}, \frac{\tau^2}{\sum_{i=1}^{n} w_{ki}}\right).$$

It is clearly inappropriate if $\mathcal{W}$ is random, because one could get $\sum_{i=1}^{n} w_{ki} = 0$, leading to an infinite variance.

It also corresponds to an improper joint distribution for $\boldsymbol{\phi}$ (singular precision matrix), which could lead to problems in updating $\mathcal{W}$ due to the computation of the determinant.

Therefore we introduce $\tilde{\phi} = (\phi, \phi_*)$, where ϕ_* is a global random effect that prevents any unit from having no edges. The corresponding $(n+1) \times (n+1)$ neighbourhood matrix is given by

$$\tilde{W} = \begin{bmatrix} W & \mathbf{w}_* \\ \mathbf{w}_*^{\mathrm{T}} & 0 \end{bmatrix},$$

where $\mathbf{w}_* = (w_{1*}, \ldots, w_{n*})$, $w_{k*} = \mathrm{I}[\sum_{i \sim k}(1 - w_{ki}) > 0]$.

Essentially, $w_{k*} = 1$ if at least one of the edges relating to areal unit k has been removed.

We propose the extended CAR prior $\tilde{\phi} \sim \mathrm{N}(\mathbf{0}, \tau^2 Q(\tilde{W}, \epsilon)^{-1})$, where

$$Q(\tilde{W}, \epsilon) \;=\; \mathrm{diag}(\tilde{W}\mathbf{1}) - \tilde{W} + \epsilon I.$$

This is a CAR prior for $\tilde{\phi}$, except for the addition of ϵI, which ensures the matrix is invertible and thus the determinant is non-zero when calculating the acceptance probability for $\mathcal{W}$. $\epsilon = 0.001$ is used here, but the model is robust to different choices.

The full conditional distributions are given by:

$$\phi_k|\tilde{\boldsymbol{\phi}}_{-k} \sim \mathbf{N}\left(\frac{\sum_{i=1}^{n} w_{ki}\phi_i + w_{k*}\phi_*}{\sum_{i=1}^{n} w_{ki} + w_{k*} + \epsilon}, \frac{\tau^2}{\sum_{i=1}^{n} w_{ki} + w_{k*} + \epsilon}\right).$$

and

$$\phi_*|\tilde{\boldsymbol{\phi}}_{-*} \sim \mathbf{N}\left(\frac{\sum_{i=1}^{n} w_{i*}\phi_i}{\sum_{i=1}^{n} w_{i*} + \epsilon}, \frac{\tau^2}{\sum_{i=1}^{n} w_{i*} + \epsilon}\right).$$

Prior distribution - $f(\mathcal{W})$

- The dimensionality of $\mathcal{W}$ is $N_{\mathcal{W}} = \mathbf{1}^T W \mathbf{1}/2$, and as each edge is binary the sample space has size $2^{N_{\mathcal{W}}}$.
- Previous studies have shown that modelling each element in $\mathcal{W}$ separately results in weakly identifiable parameters.
- Therefore we treat $\mathcal{W}$ as a single random quantity, and propose the following prior for $\tilde{W}$;

$$\tilde{W} \sim \text{Discrete Uniform}(\tilde{W}^{(0)}, \tilde{W}^{(1)}, \ldots, \tilde{W}^{(N_{\mathcal{W}})}).$$

- Candidate $\tilde{W}^{(j)}$ has j edges retained in the model (i.e. j elements in $\mathcal{W}$ equal 1), so $(\tilde{W}^{(0)}, \tilde{W}^{(N_{\mathcal{W}})})$ correspond to independence and CAR priors respectively.

- The size and dimensionality of the sample space for $\mathcal{W}$ is large, and Li *et al. (2011)* have shown that treating all of these elements as binary random quantities leads to them being weakly identifiable.
- Therefore some form of data reduction is required, and the

$$\tilde{W} \sim \text{Discrete Uniform}(\tilde{W}^{(0)}, \tilde{W}^{(1)}, \ldots, \tilde{W}^{(N_{\mathcal{W}})}),$$

prior reduces the sample space to size $N_{\mathcal{W}} + 1$ rather than $2^{N_{\mathcal{W}}}$, and its dimensionality reduces from $N_{\mathcal{W}}$ to 1.

- This leads to the question of how should the sample states $\tilde{W}^{(0)}, \tilde{W}^{(1)}, \ldots, \tilde{W}^{(N_{\mathcal{W}})}$ be chosen?

Eliciting $(\tilde{W}^{(0)}, \tilde{W}^{(1)}, \ldots, \tilde{W}^{(N_W)})$

- We propose eliciting $(\tilde{W}^{(0)}, \tilde{W}^{(1)}, \ldots, \tilde{W}^{(N_W)})$ from disease data prior to the study period, because it should have a similar spatial structure to the response.
- Let $((\mathbf{Y}_1^p, \mathbf{E}_1^p), \ldots, (\mathbf{Y}_r^p, \mathbf{E}_r^p))$ denote disease data for the r years prior to the study.
- The study data have expectation $\mathbb{E}[\mathbf{Y}] = \mathbf{E}\exp(X\boldsymbol{\beta} + \boldsymbol{\phi})$, which is equivalent to $\ln(\mathbb{E}[\mathbf{Y}]/\mathbf{E}) = X\boldsymbol{\beta} + \boldsymbol{\phi}$. Thus as $\boldsymbol{\phi} \sim \mathrm{N}(\mathbf{0}, \tau^2 Q(\tilde{W}, \epsilon)_{1:n}^{-1})$ we make the approximation:

$$\phi_j^p = \ln\left[\frac{\mathbf{Y}_j^p}{\mathbf{E}_j^p}\right] \approx \ln\left[\frac{\mathbf{Y}}{\mathbf{E}}\right] \sim_{approx} \mathrm{N}(X\boldsymbol{\beta}, \tau^2 Q(\tilde{W}, \epsilon)_{1:n}^{-1}).$$

Elicitation algorithm

1. Start at $\tilde{W}^{(N_{\mathcal{W}})}$ which has all edges retained in the model ($w_{kj} = 1$) and corresponds to the ordinary CAR prior for strong spatial smoothing.

2. For $j = N_{\mathcal{W}}, \ldots, 1$ move from $\tilde{W}^{(j)}$ to $\tilde{W}^{(j-1)}$ by removing a single edge from $\mathcal{W}$ (i.e. by setting an element in $\mathcal{W}$ equal to zero). This corresponds to localised spatial smoothing

3. When $j = 0$ $\tilde{W}^{(0)}$ contains no edges ($w_{kj} = 0$), and corresponds to non-spatial smoothing.

Moving from $\tilde{W}^{(j)}$ to $\tilde{W}^{(j-1)}$

At step j compute the joint approximate Gaussian log-likelihood for $(\phi_1^p, \ldots, \phi_r^p)$ given by

$$\begin{aligned}
\ln[f(\phi_1^p, \ldots, \phi_r^p | \tilde{W}^{(*)})] &= \sum_{j=1}^{r} \ln[\mathrm{N}(\phi_j^p | X\hat{\beta}, \hat{\tau}^2 Q(\tilde{W}^*, \epsilon)_{1:n}^{-1})], \\
&\approx \frac{r}{2} \ln(|Q(\tilde{W}^*, \epsilon)_{1:n}|) - \frac{nr}{2} \ln(\hat{\tau}^2) \\
&\quad - \frac{1}{2\hat{\tau}^2} \sum_{j=1}^{r} (\phi_j^p - X\hat{\beta})^{\mathrm{T}} Q(\tilde{W}^*, \epsilon)_{1:n} (\phi_j^p - X\hat{\beta}),
\end{aligned}$$

for all matrices $\tilde{W}^{(*)}$ that differ from $\tilde{W}^{(j)}$ by having one additional edge removed. Then set $\tilde{W}^{(j-1)}$ equal to the value of $\tilde{W}^{(*)}$ that maximises the log-likelihood.

Overall model

The overall model is given by

$$
\begin{aligned}
Y_k | E_k, R_k &\sim \text{Poisson}(E_k R_k) \quad \text{for } k = 1, \ldots, n, \\
\ln(R_k) &= \mathbf{x}_k^{\mathrm{T}} \boldsymbol{\beta} + \phi_k, \\
\tilde{\boldsymbol{\phi}} &\sim \text{N}(\mathbf{0}, \tau^2 Q(\tilde{W}, \epsilon)^{-1}), \\
\tilde{W} &\sim \text{Discrete Uniform}(\tilde{W}^{(0)}, \tilde{W}^{(1)}, \ldots, \tilde{W}^{(N_{\mathcal{W}})}), \\
\beta_j &\sim \text{N}(0, 1000) \quad \text{for } j = 1, \ldots, p, \\
\tau^2 &\sim \text{Uniform}(0, 1000).
\end{aligned}
$$

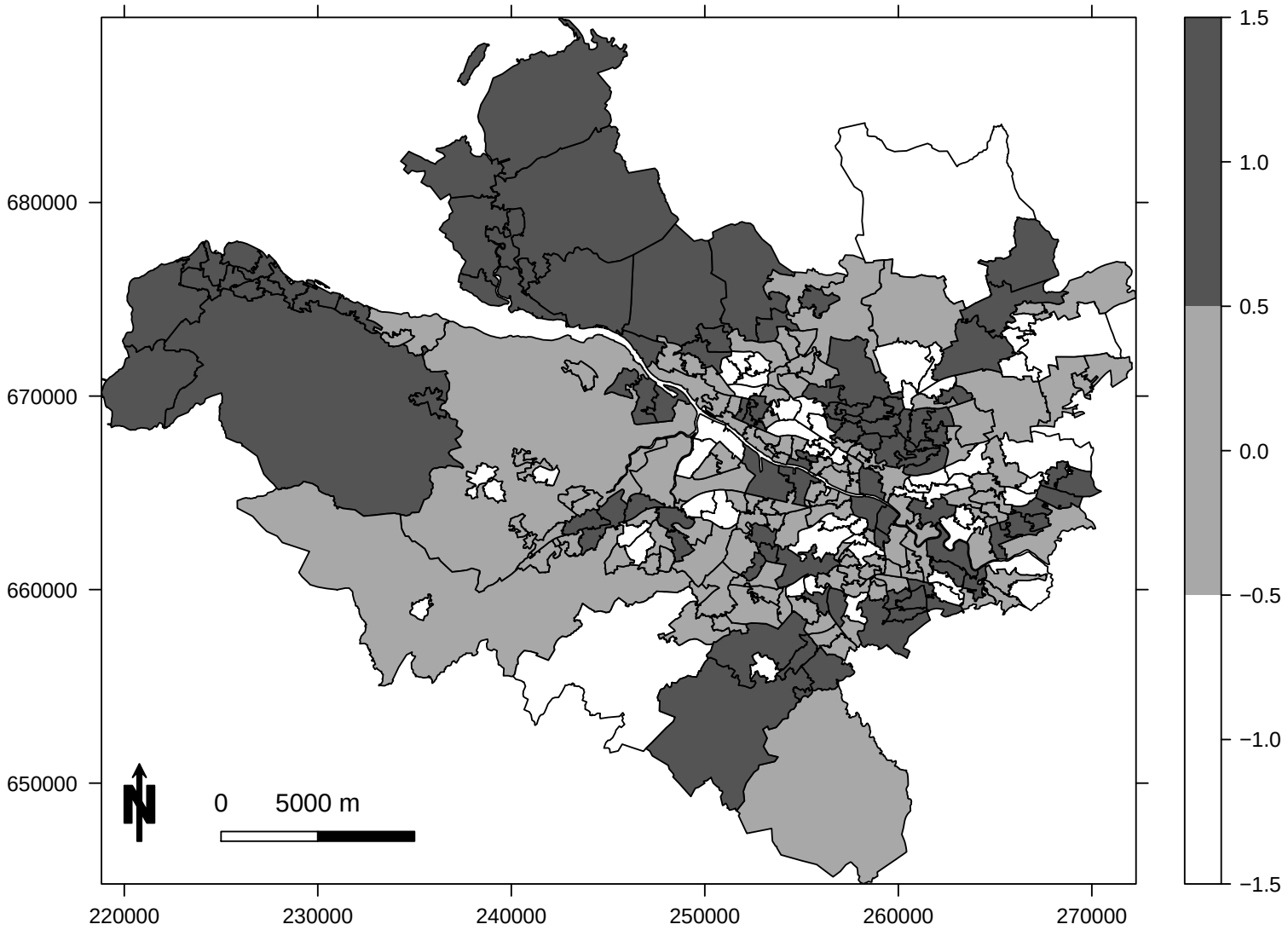
This approach thus produces a random effects model for localised spatial smoothing, which by construction will likely favour spatial structures that are not collinear to the covariates.

3. Simulation study

- Five hundred data sets were generated for Greater Glasgow under a number of different scenarios.
- Each data set consisted of study data and three years of prior data for the LCAR model.
- For each data set the log-risk surface was generated as a linear combination of a spatially smooth covariate (representing air pollution) and localised residual spatial structure (to be modelled by the random effects).
- The localised residual spatial structure was generated from a multivariate Gaussian distribution with a spatially smooth variance and a piecewise constant mean, the template for which is shown on the next slide.

Generating localised spatial structure

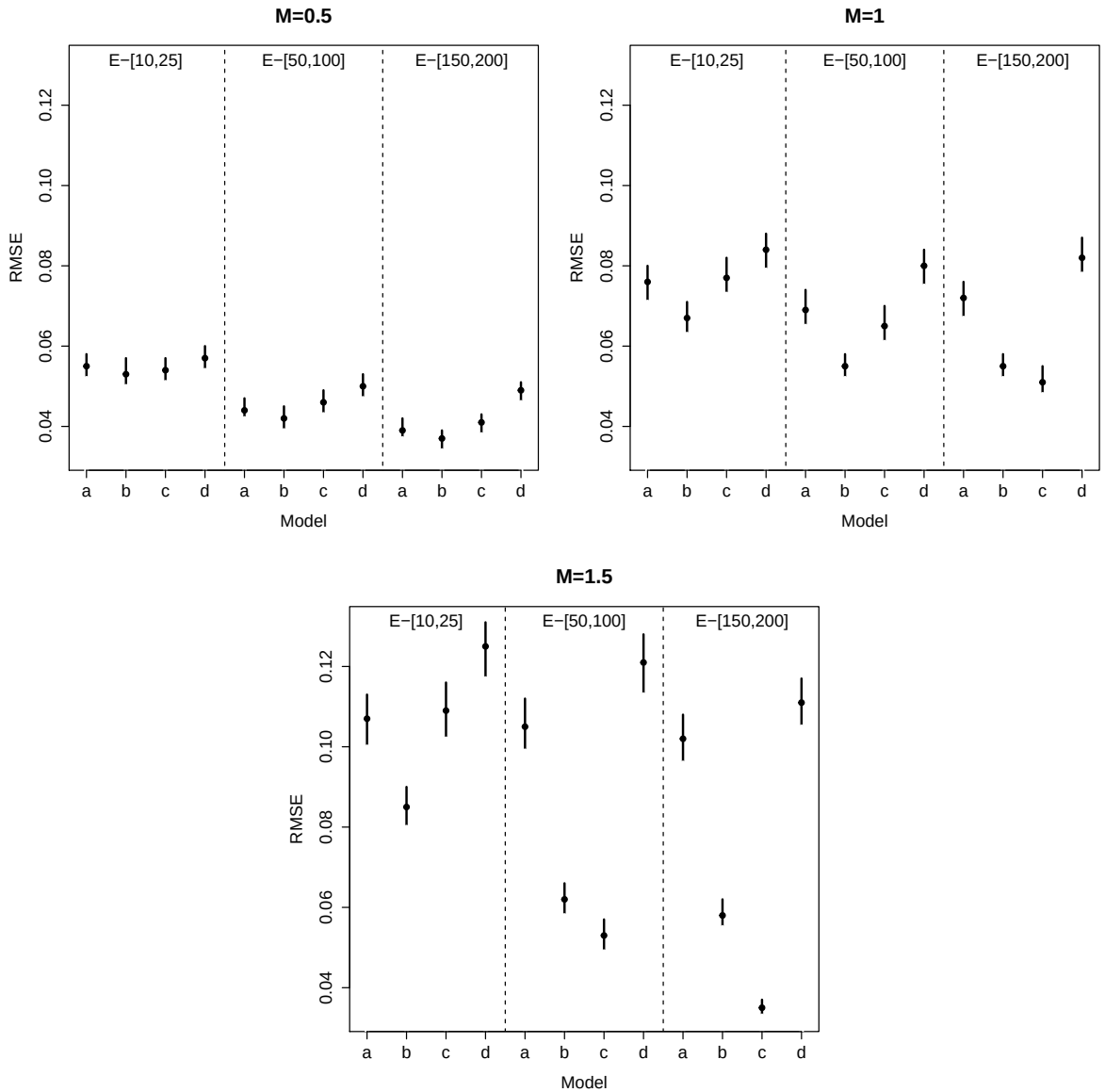
The piecewise constant mean below is multiplied by M.



Study design

- The LCAR model was compared against the commonly used BYM model, as well as the localised smoothing proposal of Lee and Mitchell *(2013)* and the orthogonal smoothing proposal of Hughes and Haran *(2013)*.
- The constant M is set to $M = 0.5, 1, 1, 5$, where larger values represent more localised rather than global spatial smoothness.
- Disease prevalence $\mathbf{E}$ is also changed from (a) [10-25], (b) [50-100], (c) [150,200].
- Model performance was compared by the Root Mean Square Error (RMSE) of the estimated regression parameter and the coverage probability of its 95% credible interval.

RMSE values for β by $(M, \mathbf{E})$



a - BYM, b - LCAR, c - Lee *et al.* (2013), and d - Hughes *et al.* (2013)

Coverage probabilities for β by $(M, \mathbf{E})$

E	M	**Model**			
		a	**b**	**c**	**d**
$[10, 25]$	0.5	94.2	92.2	92.8	73.8
	1	94.2	92.8	91.0	53.0
	1.5	94.4	93.0	80.0	32.8
$[50, 100]$	0.5	92.6	90.2	86.6	46.2
	1	94.0	89.8	73.8	28.0
	1.5	90.8	92.8	79.0	20.4
$[150, 200]$	0.5	94.2	89.6	78.2	31.4
	1	90.2	85.8	67.0	18.0
	1.5	92.4	93.0	81.4	12.6

a - BYM, b - LCAR, c - Lee *et al.* (2013), and d - Hughes *et al.* (2013)

4. Case study

- The methodology was motivated by a study estimating the effects of air pollution on hospitalisation due to respiratory disease in Greater Glasgow, Scotland in 2010.
- The prior distribution for the spatial structure of the LCAR model was elicited using three years of respiratory hospitalisation data between 2007 and 2009.
- Modelled concentrations of nitrogen dioxide (NO_2), and particulate matter ($PM_{2.5}$ and PM_{10}) were available for 2009, along with a measure of income deprivation, a major confounder in spatial ecological studies.
- The pollutants were included in separate models to avoid issues of collinearity, and in all cases inference was based on 150,000 samples obtained from 3 Markov chains.

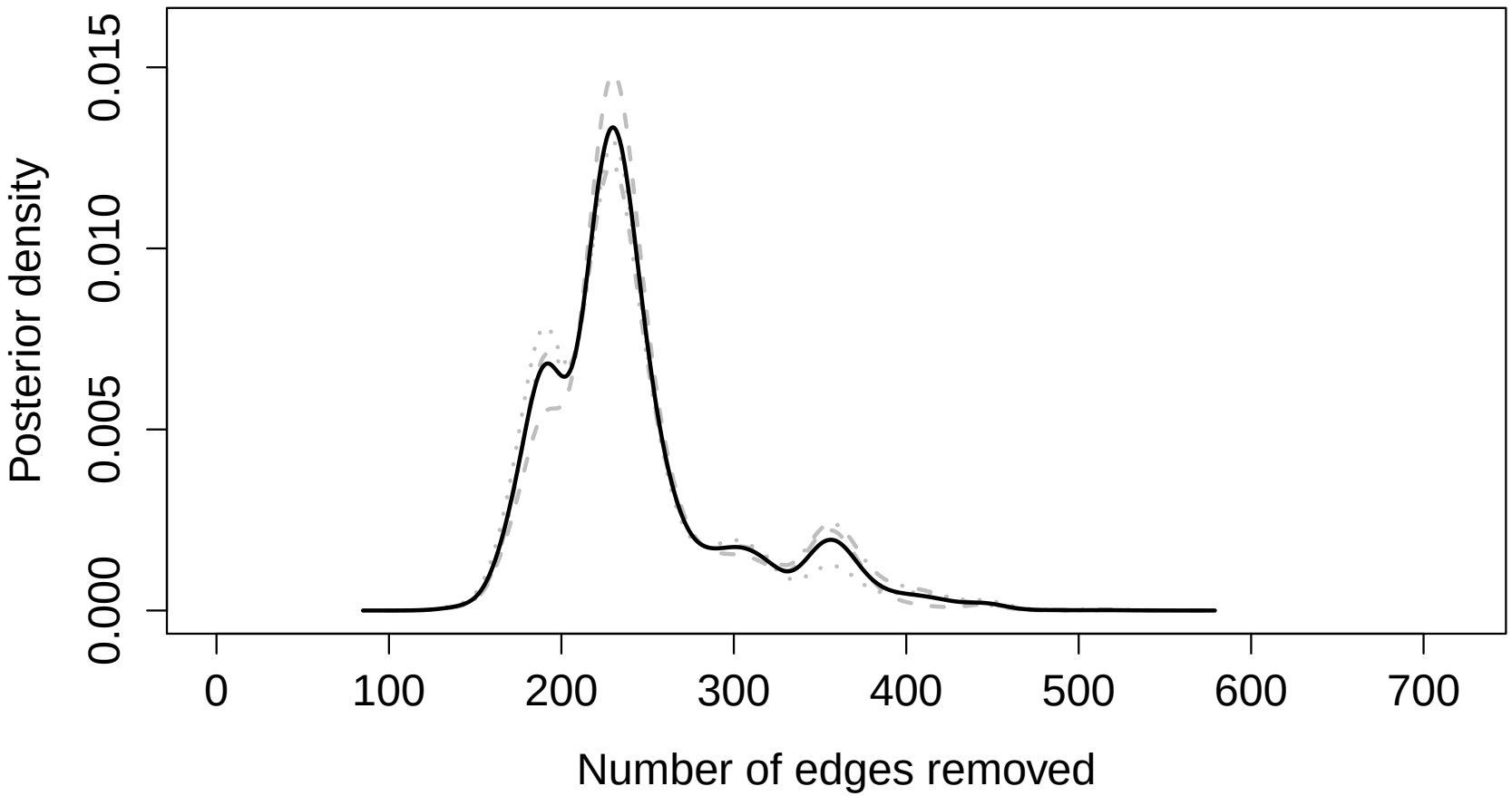
Results 1

The table shows the DIC values as well as relative risks for a one standard deviation increase in the yearly average concentrations, which are: CO (0.0076 mgm^{-3}), NO_2 (5.0μgm^{-3}), $PM_{2.5}$ (1.1μgm^{-3}), PM_{10} (1.5μgm^{-3}).

	Model			
	a	**b**	**c**	**d**
DIC	2124.0	2112.4	2115.8	2467.6
CO	0.997 (0.954, 1.038)	1.011 (0.973, 1.045)	0.998 (0.959, 1.036)	1.021 (1.006, 1.035)
NO_2	1.036 (0.998, 1.072)	1.040 (1.012, 1.067)	1.033 (1.003, 1.065)	1.043 (1.028, 1.059)
$PM_{2.5}$	1.029 (0.991, 1.067)	1.039 (1.007, 1.071)	1.026 (0.989, 1.063)	1.035 (1.021, 1.050)
PM_{10}	1.032 (0.994, 1.071)	1.040 (1.007, 1.073)	1.028 (0.993, 1.064)	1.034 (1.021, 1.048)

Results 2 - Localised spatial structure

The posterior distribution for the number of edges removed (i.e $w_{kj} \in \mathcal{W}$ that equal zero).



5. Conclusions

1. The LCAR prior proposed here has the flexibility to capture both sub-regions of spatial correlation and step changes in the random effects surface.

2. By construction the candidate localised smoothing structures are unlikely to be collinear to the fixed effects, which should eliminate the problems identified by Reich *et al. (2006)*.

3. The improvements in the estimation of the fixed effects can be substantial, as the percentage reductions in RMSE between the BYM and LCAR models ranged between 4.5% and 45.8% in the simulation study presented here.

4. The models proposed by Lee and Mitchell *(2013)* and Hughes and Haran *(2013)* have issues of poor coverage.

Further work

- Future work will extend this model into the spatio-temporal domain, which with replication of the spatial process over time will enable modelling to be undertaken without such vast dimension reduction.
- The concept of localised spatial smoothing is also useful in the related fields of disease mapping and Wombling, whose aims are to identify clusters of high-risk areas and locations of step-changes in disease risk.
- We also have monthly air pollution and health data for England by local authority for 10 years, which will be one of the largest scale studies into the long-term health impact of air pollution.

A spatio-temporal extension - tentative steps

Letting t denote year, a spatio-temporal extension to the Poisson log-linear model we are currently working on is given by

$$\begin{aligned} Y_{kt} &\sim \text{Poisson}(E_{kt}R_{kt}), \\ \log(R_{kt}) &= \mathbf{x}_{kt}^{\mathrm{T}}\boldsymbol{\beta} + \phi_{kt}, \\ \boldsymbol{\phi}_t = (\phi_{1t}, \ldots, \phi_{nt}) &\sim \mathrm{N}(\alpha\boldsymbol{\phi}_{t-1}, \tau^2 Q(W)^{-1}), \\ \boldsymbol{\phi}_1 &\sim \mathrm{N}(\mathbf{0}, \tau^2 Q(W)^{-1}), \end{aligned}$$

where $Q(W) = \text{diag}(W\mathbf{1}) - W + \epsilon I$.

- In this spatio-temporal context we assume W is constant over time, so that the temporal replication is used to estimate W without having to make the vast simplifying assumptions used in the purely spatial case.
- Also, we model each element in $\mathcal{W}$ as a continuous quantity in the range $(0, 1)$ rather than as a binary quantity, as the normalising constant of the model below is easy to compute.
- Then a second stage localised smoothing model is specified for the transformed weights

$$\nu_{kj} = \log\left(\frac{w_{kj}}{1 - w_{kj}}\right)$$

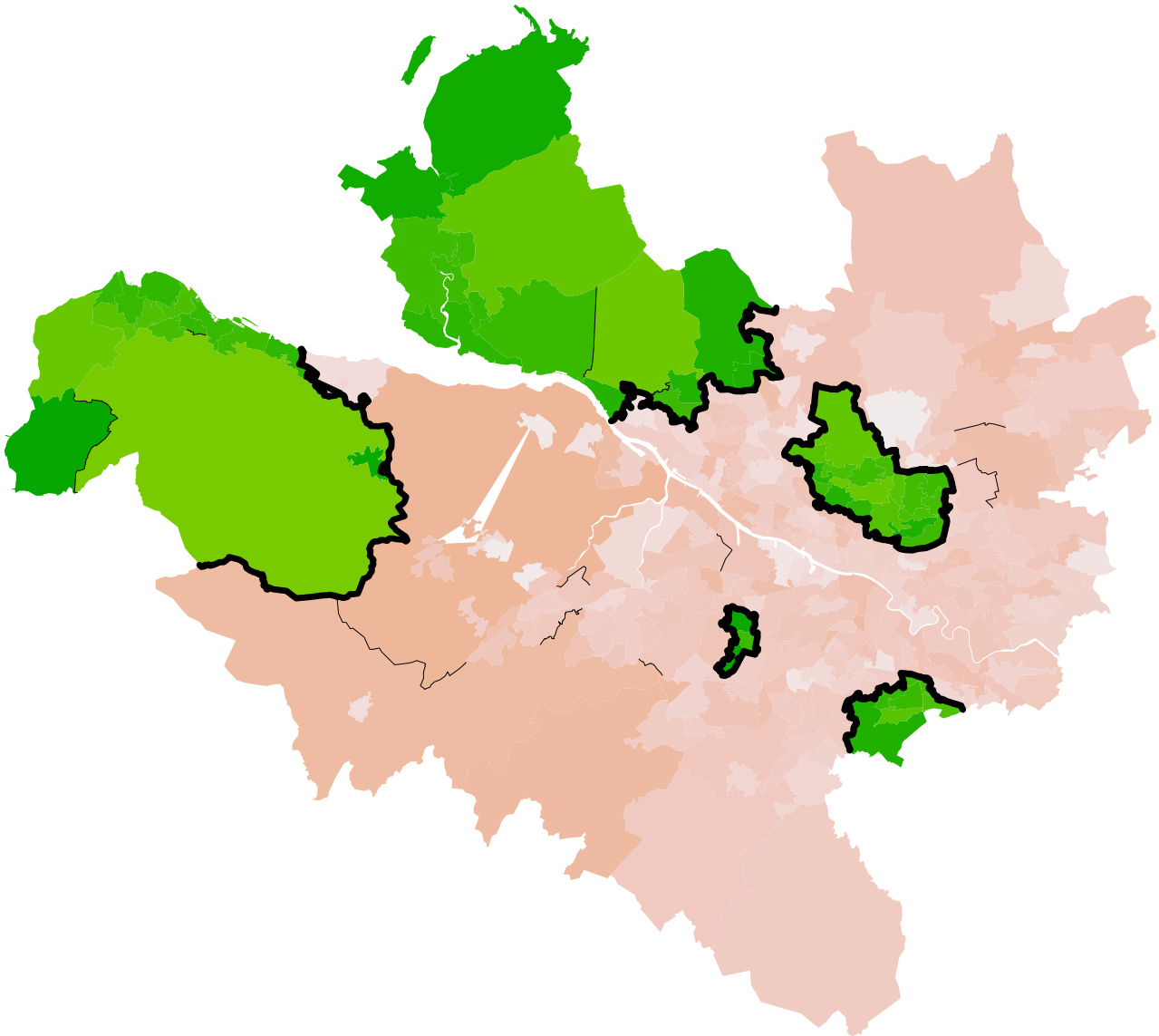
The CAR prior proposed by Leroux *et al. (1999)* is proposed for $\boldsymbol{\nu}$, the vector comprising $\{\nu_{kj}|k \sim j\}$, which has the general form

$$\boldsymbol{\nu} \sim \mathrm{N}(\mathbf{0}, \sigma^2[\rho(\mathrm{diag}(M\mathbf{1})) - (1-\rho)I]^{-1})$$

where M is the neighbourhood matrix for the set of edges. Thus

1. If $\rho = 1$ the ν_{kj} are spatially smooth hence the locations of step changes are spatially smooth.
2. If $\rho = 0$ the ν_{kj} are independent hence the locations of step changes exhibit no spatial structure.

Example - Model behaviour $\rho = 0$



Example - Model behaviour $\rho = 1$