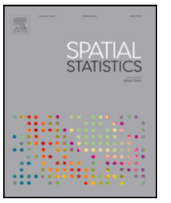
Spatial Statistics 59 (2024) 100796

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta



Computationally efficient localised spatial smoothing of disease rates using anisotropic basis functions and penalised regression fitting

Duncan Lee

School of Mathematics and Statistics, University of Glasgow, University Place, Glasgow, G12 8SQ, Scotland, United Kingdom

ARTICLE INFO

Keywords:
Anisotropic spatial basis functions
Big data disease mapping
Mental ill health
Ridge regression

ABSTRACT

The spatial variation in population-level disease rates can be estimated from aggregated disease data relating to N areal units using Bayesian hierarchical models. Spatial autocorrelation in these data is captured by random effects that are assigned a Conditional autoregressive (CAR) prior, which assumes that neighbouring areal units exhibit similar disease rates. This approach ignores boundaries in the disease rate surface, which are locations where neighbouring units exhibit a step-change in their rates. CAR type models have been extended to account for this localised spatial smoothness, but they are computationally prohibitive for big data sets. Therefore this paper proposes a novel computationally efficient approach for localised spatial smoothing, which is motivated by a new study of mental ill health across $N = 32{,}754$ Lower Super Output Areas in England. The approach is based on a computationally efficient ridge regression framework, where the spatial trend in disease rates is modelled by a set of anisotropic spatial basis functions that can exhibit either smooth or step change transitions in values between neighbouring areal units. The efficacy of this approach is evidenced by simulation, before using it to identify the highest rate areas and the magnitude of the health inequalities in four measures of mental ill health, namely antidepressant usage, benefit claims, depression diagnoses and hospitalisations.

1. Introduction

The Population-level risk of disease varies spatially between different communities, and this variation is most often estimated from disease count or disease rate data relating to a set of N non-overlapping areal units. Many important epidemiological questions are answered by modelling these data, including estimating the effects of an 'exposure' on disease rates (e.g., Djeudeu et al., 2020); quantifying the magnitude of health inequalities (e.g., Jack et al., 2019); and highlighting the locations of sub-regions that exhibit elevated disease rates (e.g., Wakefield and Kim, 2013). Modelling the spatial variation in disease rates is a research field known as *disease mapping*, and a recent review is given by MacNab (2022).

The spatial variation in the disease rates is most often modelled using a Bayesian hierarchical structure, which includes a set of N random effects, one for each area. The spatial autocorrelation in the data is modelled through the covariance structure of these random effects, and conditional autoregressive (CAR, Besag et al., 1991) priors are commonly used to capture this autocorrelation. The majority of these priors utilise a single overall level of spatial autocorrelation such as the models of Besag et al. (1991) and Leroux et al. (2000), which means that similar levels of spatial smoothing are enforced across the entire study region. However, disease rate surfaces often exhibit step changes from low to high values between geographically neighbouring units, which are known

E-mail address: Duncan.Lee@glasgow.ac.uk.

https://doi.org/10.1016/j.spasta.2023.100796
Received 17 November 2023; Accepted 22 November 2023
Available online 29 November 2023

as *boundaries* and were first identified by Womble (1951). These boundaries have led researchers to develop models that smooth disease rate surfaces locally rather than globally, which means that any boundaries are not smoothed over. For example, Brewer and Nolan (2007) propose a spatially varying variance parameter for CAR type models, Charras-Garrido et al. (2011) incorporate a piece-wise constant cluster component into the mean model, while Gao et al. (2022) specify a discrete probability distribution for the spatial random effects with an areally referenced Dirichlet process prior. The most common approach however is to add an additional modelling layer to CAR-type priors that estimate whether or not each pair of neighbouring random effects should be smoothed towards each other, with examples including Ma et al. (2010), Lee and Mitchell (2012, 2013), Rushworth et al. (2017), Berchuck et al. (2019), and Lee et al. (2021).

However, these localised smoothing approaches are extremely computationally demanding, because they can have large numbers of parameters (e.g., Ma et al., 2010 and Rushworth et al., 2017), require matrix operations at each Markov chain Monte Carlo iteration (e.g., Lee and Mitchell, 2012), or utilise iterative or complex model fitting algorithms (e.g., Lee and Mitchell, 2013 and Lee et al., 2021). Therefore, they have been restricted to relatively small spatial data applications with at most 1000 areal units, and are computationally infeasible for *big data* applications. Computationally efficient models for big disease data is currently a frontier in the field, with examples including the divide and conquer methodology proposed by Orozco-Acosta et al. (2021) and the reduced-dimensional representation of the random effects based on random projections proposed by Guan and Haran (2018). However, due to the computational challenges posed by the big data these models assume a single global level of spatial autocorrelation, which the localised smoothing literature has consistently shown to result in inferior estimation performance.

Therefore, this paper proposes the first computationally efficient approach for localised smoothing for big spatial data, which is close in computational efficiency to global smoothing models but has the improved estimation performance of localised smoothing models. The approach represents the disease rate surface with N anisotropic spatial basis functions, which can exhibit either smooth or step change transitions in values between neighbouring areal units. They are constructed using a geodesic distance measure and an ancillary data set, the latter having a similar spatial pattern to the study data. These basis functions are included in a computationally efficient generalised linear model, and to prevent overfitting a ridge regression penalty (Hoerl and Kennard, 1970) is applied to the corresponding regression parameters. This approach thus models the spatial trend and autocorrelation in the disease rates through the mean model, rather than via the covariance structure of a set of random effects as CAR models do.

This methodology is motivated by a new study of the spatial patterns and inequalities in mental ill health in mainland England using data from $N = 32{,}754$ small-areas, thus making it an order of magnitude bigger than existing disease mapping applications. The data and the epidemiological questions that motivate the study are presented in Section 2, while a brief review of disease mapping is given in Section 3. The novel methodology is presented in Section 4, while its efficacy is assessed by simulation in Section 5. The results of the motivating study are presented in Section 6, while the paper concludes with a discussion in Section 7.

2. Motivating study

2.1. Aim and questions of interest

This study aims to estimate the spatial distribution of mental ill health across mainland England in 2019 by answering the following questions.

(i). Which regions of England have the highest levels of mental ill health?
(ii). Where are the most prominent boundaries that separate neighbouring areas with high and low rates of mental ill health?
(iii). Are there socio-economic inequalities in the levels of mental ill health?
(iv). How robust are the answers to questions (i) to (iii) to how one measures the rates of mental ill health?

Answering (i) and (ii) allows policymakers to pinpoint the spatial extent of high rate clusters, allowing future health interventions to be targeted at communities who have the greatest need. Answering (iii) allows one to quantify the level of inequality between rich and poor communities, and the reduction of such inequalities is a government policy objective (Public Health England, 2021). Finally, answering (iv) provides a comprehensive and robust picture of the spatial distribution of mental ill health in England.

2.2. Study data

Mainland England is partitioned into $N = 32{,}754$ Lower Super Output Areas (LSOA), which have an average population of around 1700 people. LSOA-level mental ill health rates are available from the Place-Based Longitudinal Data Resource (https://pldr.org/) annually between 2011 and 2019, and in this study we focus on 2019. The data comprise 4 indicators of mental ill health, namely antidepressant usage, benefits claims for state aid due to mental ill health, doctor confirmed diagnoses of depression, and incidence of hospitalisations. A full description of each indicator is given in Section 1 of the supplementary material, along with a map of its spatial distribution. We note that these indicators are all crude (population size adjusted) rates of mental ill health unadjusted for population demography, and thus relate to the overall prevalence of mental ill health.

Fig. 1 displays the relationships between the individual indicators via scatterplots and correlation coefficients, with the latter ranging between 0.291 (benefit claimants and depression) and 0.706 (benefits and hospitalisations), suggesting that the four indicators capture somewhat different aspects of mental ill health. Additionally, the indicators are all on different numerical scales and are skewed to the right, suggesting that a transformation will be required before one can fairly compare their spatial

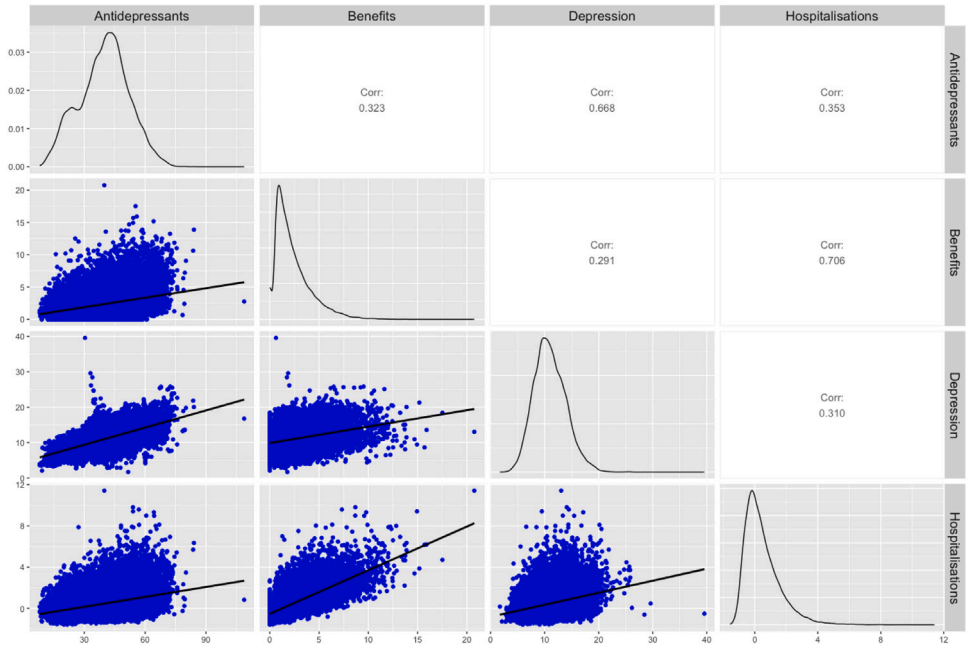


Fig. 1. Pairs plot showing the relationships between the individual indicators.

distributions. The level of spatial autocorrelation in these indicators is assessed using a permutation test based on Moran's I statistic (Moran, 1950), which is constructed using a neighbourhood matrix that determines which pairs of LSOAs are close together. Here we assume that two LSOAs are neighbours if they share a common border. Based on this matrix all four indicators exhibit substantial spatial autocorrelation, with statistically significant (at the 0.1% level) Moran's I values of: antidepressants - 0.948; benefits - 0.427; depression - 0.867; and hospitalisations - 0.375. Thus, while each indicator is spatially autocorrelated, antidepressant usage and depression diagnoses appear to exhibit much stronger levels of autocorrelation than benefit claims or hospitalisations.

3. Spatial disease mapping

3.1. A commonly used global smoothing model

The study region is partitioned into N non-overlapping areal units, $\mathcal{S} = \{\mathcal{A}_1, \ldots, \mathcal{A}_N\}$, and $\mathbf{Y} = \{Y(\mathcal{A}_1), \ldots, Y(\mathcal{A}_N)\}$ denotes a vector of disease rates or counts for each areal unit. In the latter case they will be accompanied by a set of population offsets $\mathbf{O} = \{O(\mathcal{A}_1), \ldots, O(\mathcal{A}_N)\}$, which adjusts for varying population demographics across the N areal units. Additionally, there may also be an $N \times p$ matrix of covariates $\mathbf{Z}$, whose kth row is given by $\mathbf{z}(\mathcal{A}_k) = \{1, z_2(\mathcal{A}_k), \ldots, z_p(\mathcal{A}_k)\}$. The following generalised linear mixed model is commonly used to model these data in a Bayesian setting.

$$\begin{aligned} Y(\mathcal{A}_k) &\sim f\{Y(\mathcal{A}_k)|\mu(\mathcal{A}_k), \sigma^2\} \\ g\{\mu(\mathcal{A}_k)\} &= O(\mathcal{A}_k) + \mathbf{z}(\mathcal{A}_k)^\top \boldsymbol{\alpha} + \phi(\mathcal{A}_k). \end{aligned} \tag{1}$$

Here, $\mu(\mathcal{A}_k) = \mathbf{E}\{Y(\mathcal{A}_k)\}$, $g\{.\}$ is a known invertible link function and σ^2 is an additional scale parameter if required. Two commonly used special cases of (1) are a Gaussian linear model when the disease data are rates, and a Poisson log-linear model when the disease data are counts. In either case the spatial variation in the fitted values is modelled by covariates $\mathbf{Z}$ with regression parameters $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_p\}$ and a vector of random effects $\boldsymbol{\phi} = \{\phi(\mathcal{A}_1), \ldots, \phi(\mathcal{A}_N)\}$, the latter capturing any residual spatial autocorrelation in the data through their covariance structure. Conditional autoregressive priors are commonly assumed for the random effects, and their spatial dependence structure is based on a (typically) binary $N \times N$ neighbourhood matrix $\mathbf{W}$. Here $w_{kj} = 1$ if areal units $\{\mathcal{A}_k, \mathcal{A}_j\}$ share a common border and hence are spatially close, with $w_{kj} = 0$ otherwise and $w_{kk} = 0$ for all k. CAR priors have been proposed by Besag et al. (1991), Leroux et al. (2000) and Riebler et al. (2016), with the one proposed by Leroux et al. (2000) being defined by the following set of full conditional distributions:

$$\phi(\mathcal{A}_k)|\boldsymbol{\phi}(-\mathcal{A}_k), \sim \mathrm{N}\left(\frac{\rho\sum_{j=1}^N w_{kj}\phi(\mathcal{A}_j)}{\rho\sum_{j=1}^N w_{kj} + 1 - \rho}, \frac{1}{\tau\left[\rho\sum_{j=1}^N w_{kj} + 1 - \rho\right]}\right) \quad k = 1, \ldots, N, \tag{2}$$

where $\boldsymbol{\phi}(-\mathcal{A}_k) = \boldsymbol{\phi} \setminus \{\phi(\mathcal{A}_k)\}$. Their level of spatial dependence is controlled globally by ρ, with $\rho = 0$ corresponding to spatial independence because $\phi(\mathcal{A}_k) \sim \text{N}(0, 1/\tau)$, while if $\rho = 1$ then (2) becomes the intrinsic CAR prior for strong spatial autocorrelation proposed by Besag et al. (1991). Bayesian inference for this model is typically achieved using either MCMC simulation (e.g., via `CARBayes`, Lee, 2013) or integrated nested Laplace approximations (INLA, Rue et al., 2009), and here we use INLA for its computational speed. The above model is completed by specifying weakly informative priors for the remaining parameters to let the data speak for themselves, which include: (i) $\alpha_j \sim \text{N}(0, 100000)$ for $j = 1, \ldots, p$; (ii) $\ln(\tau) \sim$ log-gamma$(1, 0.01)$; (iii) $\ln(\sigma^{-2}) \sim$ log-gamma$(1, 0.01)$; and (iv) $\ln\left(\frac{\rho}{1-\rho}\right) \sim \text{N}(0, 100)$.

3.2. Localised smoothing models

The general model outlined above is the predominant approach in the disease mapping literature, but it does control the amount of spatial smoothing globally by the single parameter ρ. This is because the partial autocorrelation between $[\phi(\mathcal{A}_k), \phi(\mathcal{A}_j)]$ implied by (2) is given by

$$\text{Corr}[\phi(\mathcal{A}_k), \phi(\mathcal{A}_j)|\boldsymbol{\phi}_{-kj}] = \frac{\rho w_{kj}}{\sqrt{\left(\rho \sum_{l=1}^{N} w_{kl} + 1 - \rho\right)\left(\rho \sum_{l=1}^{N} w_{jl} + 1 - \rho\right)}}, \tag{3}$$

where $\boldsymbol{\phi}_{-kj} = \boldsymbol{\phi} \setminus \{\phi(\mathcal{A}_k), \phi(\mathcal{A}_j)\}$. Thus all pairs of geographically neighbouring areal units with $w_{kj} = 1$ have partially autocorrelated random effects, with the strength of the autocorrelation controlled globally by ρ. In data applications ρ is typically close to one, which smooths disease rates between all pairs of neighbouring areal units and hence smooths over any boundaries in the disease rate surface. The most popular localised smoothing approach extends CAR type priors by modelling the elements of $\mathbf{W}$ corresponding to geographically adjacent areal units as random quantities, rather than being fixed at 1. From (3) this means that estimating $w_{kj} = 0$ corresponds to modelling $\{\phi(\mathcal{A}_k), \phi(\mathcal{A}_j)\}$ as conditionally independent and hence not enforcing spatial smoothing between them, while estimating $w_{kj} = 1$ means they are smoothed towards each other. Ma et al. (2010) and Rushworth et al. (2017) model these neighbourhood elements with random effects, Lee and Mitchell (2012) and Berchuck et al. (2019) estimate them based on ancillary data, Lee and Mitchell (2013) use an iterative algorithm to jointly estimate $\mathbf{W}$ and the remaining model parameters, while Lee et al. (2021) utilise a graph-based optimisation algorithm for estimating $\mathbf{W}$.

4. Methodology

This section presents a novel computationally efficient localised spatial smoothing model, which represents the spatial smoothness in disease rates through the mean function via N anisotropic basis functions. It thus differs from CAR models that capture spatial smoothness via the covariance structure of the random effects. The basis functions are constructed using the spatial structure of the areal units and a vector of ancillary data $\mathbf{V} = \{V(\mathcal{A}_1), \ldots, V(\mathcal{A}_N)\}$, the latter having a similar spatial pattern to the study data. A ridge regression penalty is applied to constrain the corresponding regression coefficients, and the model is denoted by **RAB** for Ridge regression model with Anisotropic Basis functions. A description of the model, parameter estimation and boundary identification is presented below, while the choice and construction of the ancillary data is discussed in Section 2.4 of the supplementary material.

4.1. Localised spatial smoothing model

Let $\{Y(\mathcal{A}_k), O(\mathcal{A}_k), \mathbf{z}(\mathcal{A}_k)\}$ be defined as before, and let $\mathbf{b}(\mathcal{A}_k) = (b_{\mathcal{A}_1}(\mathcal{A}_k), \ldots, b_{\mathcal{A}_N}(\mathcal{A}_k))$ denote an $N \times 1$ vector of anisotropic spatial basis functions. Each basis function is centred at a different areal unit $\mathcal{A}_k$, so that $b_{\mathcal{A}_j}(\mathcal{A}_k)$ denotes the basis function centred at $\mathcal{A}_j$ and evaluated at $\mathcal{A}_k$. The basis functions replace the random effects in (1), and the proposed model is given by

$$\begin{aligned} Y(\mathcal{A}_k) &\sim f\{Y(\mathcal{A}_k)|\mu(\mathcal{A}_k), \sigma^2\} \\ g\{\mu(\mathcal{A}_k)\} &= O(\mathcal{A}_k) + \mathbf{z}(\mathcal{A}_k)^\top \boldsymbol{\alpha} + \mathbf{b}(\mathcal{A}_k)^\top \boldsymbol{\gamma}, \end{aligned} \tag{4}$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N)$ is an $N \times 1$ vector of regression parameters. The anisotropic basis functions $\{b_{\mathcal{A}_j}(\mathcal{A}_k)\}$ are a product of isotropic (directionally invariant) basis functions $\{\tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k)\}$ and a similarity measure $\{S_{\mathcal{A}_j}(\mathcal{A}_k)\}$, the latter being constructed from the ancillary data $\mathbf{V}$. The isotropic basis functions are built as monotonically decreasing functions of the geodesic (shortest path) distance $d(\mathcal{A}_j, \mathcal{A}_k)$ between $\{\mathcal{A}_j, \mathcal{A}_k\}$. We use geodesic rather than Euclidean distance because the units are geographical regions, hence there is not a single Euclidean distance between them. Note, geodesic distance on a two-dimensional plane satisfies the 3 standard properties of a distance measure, namely non-negativity, symmetry and the triangle inequality. We consider the following three distance-decay functionals for these isotropic basis functions, and their differing shapes are displayed in Section 2.1 of the supplementary material.

$$\begin{aligned} \textbf{Exponential}: \quad \tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k) &= \exp\left\{-d(\mathcal{A}_j, \mathcal{A}_k)\right\} \\ \textbf{Inverse}: \quad \tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k) &= \frac{1}{1 + d(\mathcal{A}_j, \mathcal{A}_k)} \\ \textbf{Linear}: \quad \tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k) &= 1 - \frac{d(\mathcal{A}_j, \mathcal{A}_k)}{\max_r\{d(\mathcal{A}_j, \mathcal{A}_r)\}}. \end{aligned} \tag{5}$$

In all 3 cases $\tilde{b}_{\mathcal{A}_j}(\mathcal{A}_j) = 1$, while $\tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k) \to 0$ as $d(\mathcal{A}_j, \mathcal{A}_k) \to \infty$. The basis function centred at $\mathcal{A}_j$ thus has a local influence on the model's fitted values for areal units close to $\mathcal{A}_j$, and has little influence on the remainder of the spatial surface. Note, we examined setting the basis functions equal to zero if $d(\mathcal{A}_j, \mathcal{A}_k) > C$ for some threshold C, but this led to poorer estimation accuracy. These basis functions exhibit gradual changes in their values between neighbouring areal units, and thus are suitable for capturing spatially smooth variation in disease rates. To account for boundaries we adjust these basis functions by making the final set anisotropic, by allowing $\{b_{\mathcal{A}_j}(\mathcal{A}_k), b_{\mathcal{A}_j}(\mathcal{A}_i)\}$ to be different if $d(\mathcal{A}_j, \mathcal{A}_k) = d(\mathcal{A}_j, \mathcal{A}_i)$. We achieve this by constructing a similarity measure $S_{\mathcal{A}_j}(\mathcal{A}_k)$ between $(\mathcal{A}_j, \mathcal{A}_k)$ from the ancillary data $\mathbf{V}$, which is used to construct the anisotropic basis functions $b_{\mathcal{A}_j}(\mathcal{A}_k)$ as follows:

$$\begin{aligned} S_{\mathcal{A}_j}(\mathcal{A}_k) &= 2\left[\frac{\max_{\mathcal{A}_r}\{|V(\mathcal{A}_j) - V(\mathcal{A}_r)|\} - |V(\mathcal{A}_j) - V(\mathcal{A}_k)|}{\max_{\mathcal{A}_r}\{|V(\mathcal{A}_j) - V(\mathcal{A}_r)|\}} - \frac{1}{2}\right] \\ b_{\mathcal{A}_j}(\mathcal{A}_k) &= \tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k) S_{\mathcal{A}_j}(\mathcal{A}_k). \end{aligned} \tag{6}$$

The final basis function $b_{\mathcal{A}_j}(\mathcal{A}_k)$ is a product of an isotropic basis function $\tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k)$ and the similarity measure $S_{\mathcal{A}_j}(\mathcal{A}_k)$, with the latter making it anisotropic and giving it the ability to capture localised spatial smoothness. It is clear from (6) that $S_{\mathcal{A}_j}(\mathcal{A}_k) \in [-1, 1]$ and $S_{\mathcal{A}_j}(\mathcal{A}_j) = 1$ for all $\mathcal{A}_j$, which also means that $b_{\mathcal{A}_j}(\mathcal{A}_j) = 1$ for all j. Now consider neighbouring areal units $(\mathcal{A}_j, \mathcal{A}_k)$ with $d(\mathcal{A}_j, \mathcal{A}_k) = 1$. If these units are very similar as measured by $\mathbf{V}$ (i.e., if $V(\mathcal{A}_j) \approx V(\mathcal{A}_k)$) then $S_{\mathcal{A}_j}(\mathcal{A}_j) \approx S_{\mathcal{A}_j}(\mathcal{A}_k)$, which means that $\{b_{\mathcal{A}_j}(\mathcal{A}_j), b_{\mathcal{A}_j}(\mathcal{A}_k)\}$ will have similar values and hence lead to a smooth transition between their fitted disease rates from (4). In contrast, if $\{V(\mathcal{A}_j), V(\mathcal{A}_k)\}$ are very different then $\{S_{\mathcal{A}_j}(\mathcal{A}_j), S_{\mathcal{A}_j}(\mathcal{A}_k)\}$ and hence $\{b_{\mathcal{A}_j}(\mathcal{A}_j), b_{\mathcal{A}_j}(\mathcal{A}_k)\}$ are also very different (and possibly have different signs), allowing for a boundary between the fitted values from (4). Section 2.1 in the supplementary material illustrates this localised smoothness property.

4.2. Parameter estimation

Let $(\mathbf{Y}, \mathbf{O})$ respectively denote the $N \times 1$ vectors of disease data and population offsets, while $\mathbf{X} = (\mathbf{Z}, \mathbf{B})_{N\times(p+N)}$ denotes the combined matrix of covariates and anisotropic spatial basis functions. We assume that $\mathbf{B}$ is fixed when fitting the model, which is the standard approach used in other spatial basis function approaches such as Ugarte et al. (2012). Attempting to additionally estimate the basis functions, for example via an errors-in-variables model for $\mathbf{V}$, would lead to a computationally prohibitive algorithm for big data that negates the entire motivation and contribution of the paper. Therefore, the parameters to be estimated comprise the regression parameters $\boldsymbol{\beta} = (\boldsymbol{\alpha}, \boldsymbol{\gamma})$ and possibly a scale parameter σ^2 depending on $f\{Y(\mathcal{A}_k)|\mu(\mathcal{A}_k), \sigma^2\}$. Here we suppress the scale parameter estimation for brevity, but provide full details for Gaussian data in Section 2.2 of the supplementary material.

As there are N data points and N regression parameters $\boldsymbol{\gamma}$, we penalise the latter using a ridge regression shrinkage penalty (Hoerl and Kennard, 1970) to prevent the model from overfitting. Note, the covariate regression parameters $\boldsymbol{\alpha}$ are not penalised. We utilise a ridge penalty rather than a LASSO or elastic net penalty because this results in faster parameter estimation and allows for easy uncertainty quantification (see Sections 2.2 and 2.3 of the supplementary material). Let $\mathbf{D}$ be the $(p+N)\times(p+N)$ diagonal matrix whose first p elements equal 0 and last N elements equal 1. Then the penalised maximum likelihood estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}}(\lambda) = \underset{\boldsymbol{\beta}}{\operatorname{argmax}} \left\{ \mathcal{L}(\boldsymbol{\beta}|\mathbf{Y}, \mathbf{O}, \mathbf{X}) - \lambda \boldsymbol{\beta}^\top \mathbf{D} \boldsymbol{\beta} \right\}, \tag{7}$$

where $\mathcal{L}(\boldsymbol{\beta}|\mathbf{Y}, \mathbf{O}, \mathbf{X}) = \ln\left[\prod_{k=1}^{N} f\{Y(\mathcal{A}_k)|\mu(\mathcal{A}_k)\}\right]$ denotes the joint log-likelihood function from (4). This estimator depends on the ridge penalty parameter λ, and in a prediction context it would be estimated by optimising out-of-sample predictive performance via a cross-validation approach. However, here the inferential goal is to explain the spatial pattern in the observed data, and the specific role of the basis functions is to capture the remaining spatially autocorrelated variation in the data after covariate adjustment. Therefore, we choose λ by minimising the level of spatial autocorrelation in the linear predictor scale residuals $\{r(\mathcal{A}_k, \lambda)\}$, because it is the scale at which the basis functions operate. Thus, if the data are Gaussian and do not contain population offsets then $\{r(\mathcal{A}_k, \lambda) = Y(\mathcal{A}_k) - \mathbf{x}(\mathcal{A}_k)^\top \hat{\boldsymbol{\beta}}(\lambda)\}$, while if a Poisson log-linear model with population offsets is used then $\{r(\mathcal{A}_k, \lambda) = \ln[Y(\mathcal{A}_k)] - O(\mathcal{A}_k) - \mathbf{x}(\mathcal{A}_k)^\top \hat{\boldsymbol{\beta}}(\lambda)\}$. We measure residual spatial autocorrelation by Moran's I statistic, and hence estimate λ by

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmin}} \left\{ \left| \frac{N \sum_{k=1}^{N} \sum_{j=1}^{N} w_{kj} [r(\mathcal{A}_k, \lambda) - \bar{r}(\lambda)][r(\mathcal{A}_j, \lambda) - \bar{r}(\lambda)]}{\left(\sum_{k=1}^{N} \sum_{j=1}^{N} w_{kj}\right) \sum_{k=1}^{N} [r(\mathcal{A}_k, \lambda) - \bar{r}(\lambda)]^2} \right| \right\}, \tag{8}$$

the value that minimises its absolute value, where $\bar{r}(\lambda) = (1/N)\sum_{k=1}^{N} r(\mathcal{A}_k, \lambda)$. Thus we estimate $\boldsymbol{\beta}$ by $\hat{\boldsymbol{\beta}}(\hat{\lambda})$, from (7) and (8). Finally, we also choose the optimal distance decay functional from (5) used to construct $\{\tilde{b}_{\mathcal{A}_j}(\mathcal{A}_k)\}$ as the one that minimises the level of residual spatial autocorrelation using an analogous approach to (8). The ridge regression model is fitted using the R package `glmnet` (Friedman et al., 2010), and further details of the parameter estimation and confidence interval construction in the cases of Gaussian and Poisson data likelihood models are given in Sections 2.2 and 2.3 of the supplementary material. Finally, the methodology is available for others to use via the `S.RAB()` function in the R package `CARBayes` (Lee, 2013), and a simulated data example similar to those in the simulation study in the next section is presented in Section 2.5 of the supplementary material.

4.3. Identification of boundaries

Following Lu and Carlin (2005) we measure the size of the boundary between $\{\mathcal{A}_k, \mathcal{A}_j\}$ by the absolute difference in the estimated disease rates:

$$\hat{\Delta}_{kj} = |\hat{\theta}(\mathcal{A}_k) - \hat{\theta}(\mathcal{A}_j)|. \tag{9}$$

If the data comprise disease rates then $\hat{\theta}(\mathcal{A}_k) = \hat{\mu}(\mathcal{A}_k)$ from (1), while if they are disease counts then $\hat{\theta}(\mathcal{A}_k) = \hat{\mu}(\mathcal{A}_k)/\exp[O(\mathcal{A}_k)]$. In either case if $\hat{\Delta}_{kj} = 0$ then no boundary occurs, while as $\hat{\Delta}_{kj}$ increases the size and importance of the boundary increases. However, for ease of visual presentation on a map we do not attempt to visualise the continuous measure $\{\hat{\Delta}_{kj}\}$ for every neighbouring pair of areas directly, but instead highlight the largest values so their locations can be easily observed. Specifically, we highlight those in the percentile ranges $[90, 95)$ and $[95, 100]$ in different colours, so that boundaries of different magnitudes can be observed. We note these groupings are somewhat arbitrary, but are chosen as a trade-off between the amount of boundary information to show and the clarity of the resulting disease map. Finally, we note that the ancillary data suggest the locations of likely boundaries through the construction of the basis functions $\{b_{\mathcal{A}_j}(\mathcal{A}_k)\}$. However, the estimated boundary sizes $\{\hat{\Delta}_{kj}\}$ and hence whether they are above the 90th percentile is determined by the study data when fitting the model, because a step change in a basis function would be nullified by a near-zero regression coefficient.

5. Simulation study

This study compares the speed and inferential accuracy of the proposed RAB methodology against existing global and localised spatial smoothing models. It is based on Gaussian disease rate data with no covariates to match the motivating application, while two additional simulation studies that consider Poisson count data and the impact of covariates are presented in Section 3 of the supplementary material.

5.1. Data generation

LSOA-level observed disease rates are generated as $\{Y(\mathcal{A}_k) = \mu(\mathcal{A}_k) + \epsilon(\mathcal{A}_k)\}$, where the errors $\{\epsilon(\mathcal{A}_k)\}$ are independent zero-mean Gaussian noise with standard deviation σ that is varied in the simulation design. The true disease rates are generated as $\{\mu(\mathcal{A}_k) = \phi(\mathcal{A}_k) + \delta(\mathcal{A}_k)\}$, a linear combination of a spatially smooth surface $\{\phi(\mathcal{A}_k)\}$ to induce spatial autocorrelation, and a piece-wise constant surface $\{\delta(\mathcal{A}_k)\}$ to induce boundaries. The former is simulated from a mean-zero multivariate Gaussian distribution with marginal standard deviation $\tau = 0.1$, whose correlation matrix is represented by a geostatistical style exponential autocorrelation structure. The range of the autocorrelation is fixed so that the pair of LSOAs whose distance is the 5th percentile of all pairwise distances has a correlation of 0.75.

The boundary surface $\{\delta(\mathcal{A}_k)\}$ only takes two values $(-\lambda, \lambda)$, and LSOAs with high disease rates have $\delta(\mathcal{A}_k) = \lambda$ while those with low rates have $\delta(\mathcal{A}_k) = -\lambda$. Thus boundaries occur where two spatially adjacent LSOAs have different $\{\delta(\mathcal{A}_k)\}$ values. The sizes of the boundaries are determined by λ, and its value is varied in the simulation design. To ensure the simulated data are similar to the study data the allocation of LSOAs to $\{-\lambda, \lambda\}$ is achieved by clustering the real disease rates using the k-means algorithm with 2 clusters. Finally, the ancillary data $\{V(\mathcal{A}_k)\}$ are generated as an error-prone version of the true spatial structure $\{\mu(\mathcal{A}_k)\}$, where the error is independent Gaussian noise with mean zero and standard deviation ν. Thus, ν determines the accuracy of the ancillary data, and is varied in the simulation design to assess the impact of the quality of the ancillary data on model performance.

5.2. Study design

The study is based on the set of $N = 3486$ LSOAs that make up the West Midlands region of England, which is chosen because it contains both rural areas and densely populated urban areas including the city of Birmingham. We do not use all of England ($N = 32{,}754$) for this study because it would be computationally prohibitive to apply the competitor localised spatial smoothing model to hundreds of simulated data sets of that size. In this study one hundred data sets are generated under each of 8 scenarios, which include all pairwise combinations of the following 3 factors:

- **Boundary size** - The size of the disease rate boundaries are varied by considering $\lambda = 0.2$ (small boundaries) and $\lambda = 0.4$ (large boundaries).
- **Data noise** - The amount of random noise in the data $\{Y(\mathcal{A}_k)\}$ is varied by considering $\sigma = 0.05$ (small noise), $\sigma = 0.1$ (large noise).
- **Ancillary data quality** - The accuracy of $\{V(\mathcal{A}_k)\}$ as an estimate of the true disease rate surface is varied by setting $\nu = 0.05$ (high accuracy), and $\nu = 0.1$ (low accuracy).

The Gaussian RAB model is compared against both global and localised spatial smoothing alternatives. A global smoothing model is considered because it is the most commonly used approach in disease mapping, thus benchmarking our methodology against the 'standard' modelling approach. We also compare against a localised smoothing model because it is a direct competitor to the RAB methodology. Global spatial smoothing is represented by the Gaussian variant of the CAR-based random effects model proposed by Leroux et al. (2000), which is denoted here by **GCAR** (for global CAR) and described in Section 3.1. The model is fitted in a

Bayesian paradigm using computationally efficient INLAs, and does not use the ancillary data $\{V(\mathcal{A}_k)\}$. Localised spatial smoothing is represented by a Gaussian variant of the model proposed by Lee et al. (2021), which is also based on the CAR prior proposed by Leroux et al. (2000) but uses the ancillary data to adjust the neighbourhood matrix to allow for localised smoothing. The model is denoted by **LCAR**, and is summarised in Section 2.6 of the supplementary material. This particular localised smoothing model is chosen because it is computationally fast to implement compared with the alternatives as it uses INLAs, and is appropriate for a range of data likelihood models.

These models are compared in three main ways, (i) computational speed; (ii) in-sample fit to the observed data; and (iii) out-of-sample predictive performance. In-sample fit to the data is the most relevant to the motivating study because the aim is to estimate the true disease rates from the noisy data, whilst out-of-sample predictive performance allows us to summarise the ability of the model to estimate the true disease rates for areal units with missing observations. Both of these are measured by the bias and root mean square error (RMSE) of the estimated (predicted) disease rates, while we also assess the coverage probabilities (CP) and mean interval width (MIW) for the 95% uncertainty intervals for the in-sample fitted values. These latter statistics are not included for the predictions for brevity. Finally as this paper focuses on boundary detection, we also present the RMSE of the estimated disease rate boundaries $\{|\hat{\mu}(\mathcal{A}_k) - \hat{\mu}(\mathcal{A}_j)|\}$ for all spatially adjacent pairs of LSOAs.

5.3. Results

The study results are displayed in Table 1 with the exception of the biases in the fitted values/predictions, which are all close to zero and hence not shown for brevity. The bottom left panel displays the average time taken to fit each model on a Dell Precision R7920 4.0 GHz server. The global smoothing (GCAR) and ridge regression (RAB) models are not too dissimilar in absolute terms, being around 3–4 s for GCAR and 9–11 s for RAB, while the localised smoothing (LCAR) model is much slower taking around 80 s to fit. The GCAR model is faster than the RAB model due to the matrix algebra required by the latter to construct 95% confidence intervals for the true disease rates.

Focusing on in-sample fit to the data, the top left panel displays the RMSE of the disease rates estimates (`Fitted`). The GCAR model almost always has the highest RMSEs, with the best localised smoothing model having reductions in RMSE of between 15% (row 2) to 50% (row 7). Thus, localised smoothing outperforms simpler global smoothing in the presence of boundaries. Secondly, the RAB model either outperforms the LCAR model or produces nearly identical results in almost all cases, which when combined with its vastly improved computational speed suggests that it is a faster approach to localised spatial smoothing that provides improved or comparable disease rate estimation. The only situation where this is not the case is row 2, which corresponds to ancillary data of relatively low quality, small boundaries and low data noise. The RMSEs of the estimated boundary sizes are displayed in the top right panel of the table, and the results mirror those for the fitted disease rates discussed above, with the GCAR model performing the worst as it smooth over these boundaries while the RAB model generally performs the best.

The appropriateness of the 95% uncertainty intervals for the estimated disease rates are summarised by their coverage probabilities (middle left panel) and mean interval widths (middle right panel), and these intervals are credible intervals for the GCAR and LCAR models and confidence intervals for the RAB model. The GCAR and RAB models provide coverage probabilities (CP) in the vicinity of 0.95 albeit slightly conservative, whilst for 2 scenarios the LCAR model has CPs of 0.877 and 0.799 suggesting that it cannot always capture uncertainty appropriately. The uncertainty intervals from the RAB model are narrower on average than those from the GCAR model (middle right panel), which combined with the above results suggests that the former produces more precise inference.

Finally, out-of-sample predictive performance was assessed by running the study a second time, where 20% of each simulated data set was removed at random to form the test set. Each model was then fitted to the training set comprising the remaining 80% of the observed data, and was subsequently used to predict the disease rates for the 20% test set. The RMSEs of these predictions are presented in the bottom right panel of Table 1, which follows a similar pattern to the in-sample RMSE results. Specifically, the GCAR model has the largest RMSEs in all cases, having values that are between 36% and 65% higher than those from the best performing localised smoothing model. This is likely to be because the two localised smoothing models utilise the ancillary data, which informs on the values of these missing observations to some extent. Additionally, the RAB model generally has better prediction performance than the LCAR model, although the differences between the two models is generally small.

6. Results of the mental ill health study

This section presents the study results, focusing on identifying regions with the highest rates of mental ill health, the locations of boundaries, and the levels of inequality across England.

6.1. Model fitting

The four mental ill health indicators have different numerical scales (see Fig. 1), so to allow a meaningful comparison they are re-scaled via a square root transformation to reduce the skew, and then standardised to have a mean of 0 and a standard deviation of 1. Thus a value of 0 corresponds to an area with an average level of mental ill health across England, while positive (negative) values respectively denote LSOAs with above (below) average rates of mental ill health. As each indicator relates to 2019 the average values of the transformed indicators over the period 2016–2018 comprise the ancillary data. These data are highly correlated with the study data, with Pearson's correlation coefficients of: antidepressants - 0.99; benefits - 0.94; depression - 0.95;

Table 1
Results of the simulation study, including the speed of model fitting, and the root mean square error (RMSE) of the in-sample estimated disease rates (`Fitted`), the corresponding boundary sizes and the out-of-sample predictions (`Predicted`). Additionally, coverage probabilities (CP) and mean interval widths (MIW) of the 95% uncertainty intervals are presented for the fitted (in-samples) disease rates.

Boundary	Data	Ancillary	Model					
size	noise	data quality	GCAR	LCAR	RAB	GCAR	LCAR	RAB
			RMSE - Fitted			RMSE - Boundary size		
small	small	high	0.047	0.035	0.036	0.061	0.039	0.040
small	small	low	0.047	0.040	0.054	0.061	0.047	0.066
small	large	high	0.076	0.051	0.041	0.094	0.053	0.044
small	large	low	0.076	0.062	0.063	0.093	0.070	0.074
large	small	high	0.061	0.039	0.036	0.086	0.044	0.039
large	small	low	0.061	0.044	0.040	0.086	0.052	0.043
large	large	high	0.094	0.054	0.047	0.125	0.057	0.049
large	large	low	0.094	0.067	0.055	0.125	0.075	0.058
			CP - Fitted			MIW - Fitted		
small	small	high	0.967	0.929	0.979	0.211	0.125	0.182
small	small	low	0.968	0.920	0.974	0.210	0.140	0.283
small	large	high	0.941	0.966	0.968	0.294	0.213	0.200
small	large	low	0.942	0.951	0.949	0.294	0.242	0.267
large	small	high	0.987	0.877	0.991	0.358	0.118	0.215
large	small	low	0.987	0.799	0.992	0.359	0.112	0.282
large	large	high	0.966	0.972	0.968	0.418	0.219	0.212
large	large	low	0.966	0.937	0.974	0.418	0.243	0.267
			Computational time (seconds)			RMSE - Predicted		
small	small	high	3.4	78.0	9.6	0.108	0.052	0.048
small	small	low	3.5	77.6	10.2	0.108	0.070	0.066
small	large	high	3.6	77.7	10.6	0.113	0.062	0.048
small	large	low	3.6	77.7	11.0	0.113	0.081	0.072
large	small	high	3.5	79.6	9.1	0.210	0.073	0.075
large	small	low	3.5	78.3	9.1	0.210	0.085	0.079
large	large	high	3.4	79.6	9.6	0.213	0.078	0.081
large	large	low	3.3	79.3	9.3	0.213	0.095	0.088

and hospitalisations - 0.89. We take a 3-year average because this should reduce the noise in the ancillary data, compared with just using the values from 2018. However, in Section 4.2 of the supplementary material we assess the robustness of the results by instead using: (i) the 2-year average over 2017–2018; and (ii) 2018 only.

The Gaussian variant of the RAB model is used as the data are continuous disease rates, and details of the optimal choice of the spatial basis functions and the value of the ridge penalty are given in Section 4.1 of the supplementary material. Additionally, Section 4.3 of the supplementary material provides a comparison between the ridge regression model and the Leroux CAR model, and shows that the latter appears to overfit the data for two of the four indicators. As a result, the remainder of this section presents results from the best fitting RAB model. Finally, as the LCAR model is not computationally feasible for $N = 32{,}754$ areal units, it is compared to the RAB model in terms of its boundary quantification in Section 4.4 of the supplementary material on the $N = 3486$ LSOAs that make up the West Midlands region of England.

6.2. Results - Elevated rate regions

To identify which regions of England exhibit elevated rates of mental ill health, the LSOA-level fitted values are aggregated to the set of $N = 307$ local authority districts (LAD), by computing the mean value in each LAD. Note, LSOAs nest exactly within LADs. Those districts in the top 10% (high rate — shaded red) and bottom 10% (low rate — shaded blue) for any of the four indicators are highlighted in Fig. 2. Here, 68 LADs are in the highest 10% and 80 LADs are in the lowest 10% for at least one indicator, with 36 (highest) and 39 (lowest) being in these groups for more than 1 indicator of mental ill health. This suggests that while the 4 indicators capture somewhat different aspects of mental ill health, it is perhaps not surprising that largely the same areas exhibit elevated rates across multiple indicators. The other striking finding from Fig. 2 is the clear north-south divide in mental ill health rates, with the majority of the lowest rate areas being in the south between Worcester in the west and London in the east, while most of the highest rate areas are in the north in the vicinity of Liverpool, Manchester, Sheffield and Newcastle. The two clear exceptions to this are the ring of high rates coastal areas in East Anglia in the south-east, and the ridge of low rate areas in the north centred around Ripon in Yorkshire.

6.3. Results - Boundary locations

The locations of the largest disease rate boundaries are presented in Fig. 3 for the capital city of London, with the whole of England not shown because the boundaries would be impossible to see due to the large number of LSOAs. Each panel displays

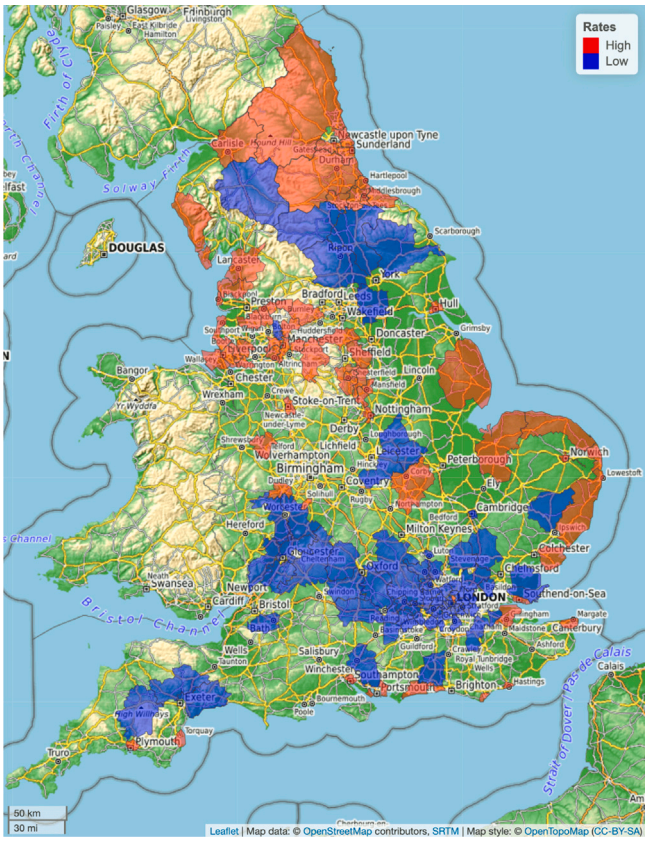


Fig. 2. A map highlighting the local authority districts (LAD) that have mental ill health rates in the highest 10% (red) and lowest 10% (blue) of all LADs in England for at least 1 indicator. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a different mental ill health indicator, and while the estimated disease rates are represented by the yellow to red colour scale, boundaries are depicted by black and grey lines. Boundaries are identified as those areal unit borders that are between the 90th and 95th (grey) and above the 95th percentiles of the set of boundary sizes $\{\hat{\Delta}_{kj}\}$ across the city. The figure shows that the antidepressant and depression indicators exhibit similar spatial trends (correlation of 0.70) and are visually spatially smooth, while the benefits and hospitalisations indicators are also similar (correlation of 0.75) but exhibit more localised spatial structure. Thus, these groupings exhibit very different spatial trends to each other, having a maximum correlation of only 0.35 between antidepressants and hospitalisations.

These differences in spatial structure can be seen by the types of boundaries that are observable in the maps, which mainly surround single LSOAs for benefits and hospitalisations, but form longer connected boundary segments for antidepressants and depression. For benefits and hospitalisations the most prominent connected boundary segment separates the low rate affluent west-end of London including Kensington and Westminster in the centre of the map from their higher rate neighbours. In contrast, for antidepressants and depression diagnoses the main connected boundary segments separate Sutton (high rate) from Croydon (low rate) in the south, Romford (high rate) from Ilford (low rate) in the north-east and follow part of the eastern portion of the river Thames.

6.4. Results - Health inequalities

The level of socio-economic inequality in mental ill health is quantified by comparing the fitted values for each of the four indicators to the 2019 English index of multiple deprivation (IMD). The index is comprised of 7 different domains (crime, education, employment, health, housing, income and living environment) that are combined to create an overall score, and further details can be found at https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019. The results are displayed in Fig. 4, which presents boxplots for each indicator partitioned by deciles of the IMD score. Here decile 1 is the most affluent and decile 10 is the most deprived.

The figure shows that all four indicators exhibit socio-economic inequalities, with LSOAs that are more deprived (higher deciles) having higher levels of mental ill health than LSOAs that are more affluent (lower deciles). These relationships are consistent across almost all deciles, as the median levels of benefit claims and hospitalisations monotonically increase with increasing deprivation,

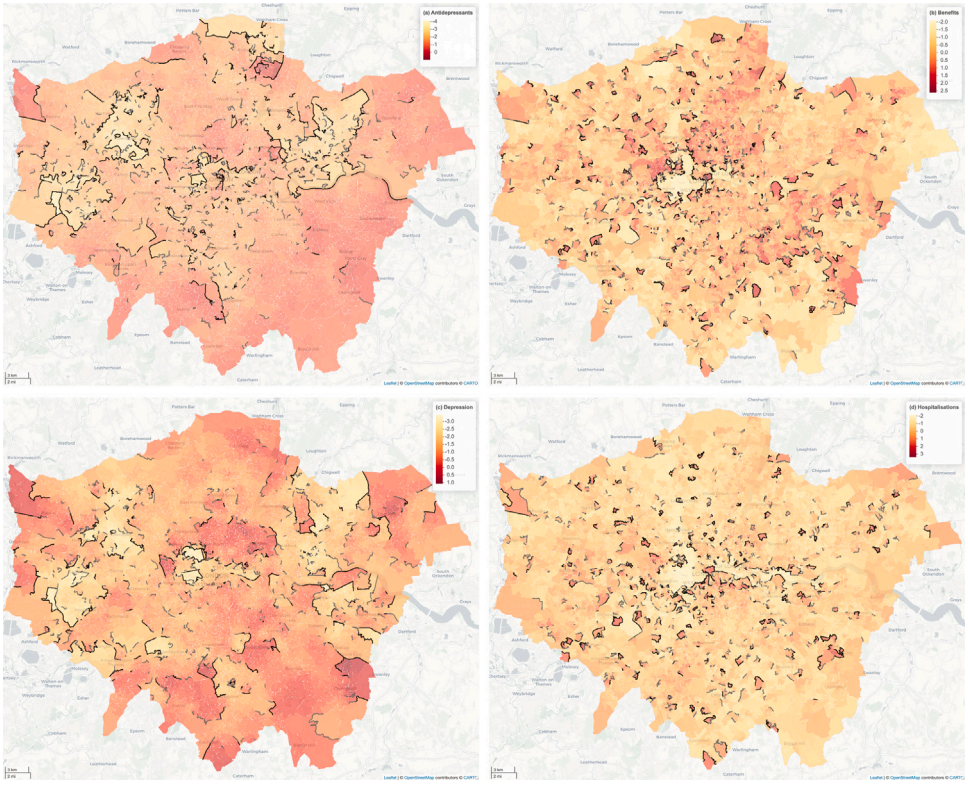


Fig. 3. The fitted values from the RAB model for each of the 4 mental ill health indicators for Greater London. Boundaries are identified as those areal unit borders that have the largest absolute differences in fitted values between neighbouring LSOAs. The boundaries in black are those above the 95th percentile of $\hat{\Delta}_{kj} = |\hat{\mu}(\mathcal{A}_k) - \hat{\mu}(\mathcal{A}_j)|$ across the region, while those in grey are between the 90th and 95th percentiles.

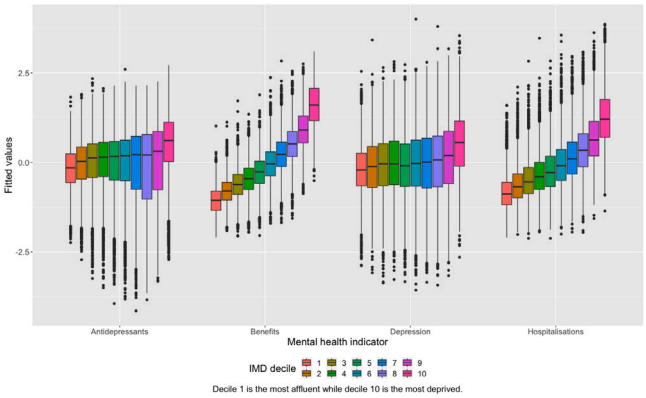


Fig. 4. Boxplots showing the level of socio-economic inequality present in the estimated disease rates for each mental health indicator. Each boxplot represents a decile of the English index of multiple deprivation, where decile 1 is the most affluent and decile 10 is the most deprived.

while the median levels of antidepressant usage and depression diagnoses are almost monotonically increasing. The figure also shows that benefit claims and hospitalisations exhibit the highest levels of socio-economic inequality, while antidepressant usage and depression diagnoses have much lower socio-economic gradients. For example, the absolute difference between the median

fitted values in deciles 10 and 1 are: antidepressants - 0.762; benefits - 2.664; depression - 0.766; and hospitalisations - 2.090; showing a clear split of the indicators into two groups with low and high levels of inequality.

7. Discussion

This paper has developed the first computationally efficient approach to localised spatial smoothing for areal unit data, which allows for both sub-regions of spatial smoothness and locations exhibiting abrupt step changes. The localised spatial structure in the disease rates is captured by a set of anisotropic spatial basis functions constructed using an ancillary data set, which are included in the mean structure of a ridge regression model. Thus unlike CAR-based random effects models that enforce spatial smoothing through the covariance structure, this approach spatially smooths the disease rates through the mean function, and could be viewed as a localised smoothing analogue of a spatial smoothing spline (Ugarte et al., 2012).

The simulation studies have shown that this approach has a much smaller computational cost than the existing localised smoothing model proposed by Lee et al. (2021), which itself is much faster than other localised smoothing models such as Lee and Mitchell (2012) and Rushworth et al. (2017) because it utilises INLA for fast parameter estimation. Additionally, it generally outperforms both the global and localised spatial smoothing alternatives, in terms of both in-sample and out-of-sample disease rate estimation and the accuracy of boundary identification. Taken together, these two factors suggest that in the data rich world in which we live, using computationally efficient model fitting tools in conjunction with ancillary data appears to be a promising paradigm for modelling big areal unit data with complex autocorrelation structures, rather than using computationally intensive and highly parameterised autocorrelation structures.

However, there is no such thing as a free lunch when modelling localised spatial smoothness, because one has to either use complex autocorrelation models that are computationally expensive to implement or utilise ancillary data as we have done here. The latter allows localised spatial smoothing to be implemented at scale for large spatial data sets, but it is critically dependent on the existence of ancillary data that are correlated with the study data, and will not work without it. Thus, our approach is appropriate for modelling slowly changing non-infectious diseases where previous years of data will exhibit similar spatial patterns and can thus be used as ancillary data. However, our approach is not appropriate for fast-moving infectious diseases such as Covid-19, where the spatial pattern in the study data cannot be easily predicted from the spatial patterns in earlier data.

A main finding from the England study is that all four mental ill health indicators exhibit socio-economic inequalities, with LSOAs that are more deprived having higher levels of mental ill health than LSOAs that are more affluent. These inequalities are strongest for benefit claims and hospitalisations, although they also exist to a lesser extent for antidepressant usage and doctor diagnoses of depression. These results agree with both national (Akhter et al., 2018) and international (Lund et al., 2011) studies, and reducing these inequalities is a priority for the UK government (https://www.gov.uk/government/publications/health-matters-reducing-health-inequalities-in-mental-illness). The other main finding is that at a larger local authority level there is a noticeable north-south divide in the prevalence of mental ill health, with much higher levels in the north of the country compared with the south. This finding is supported by recent UK Government data (https://commonslibrary.parliament.uk/health-inequalities-income-deprivation-and-north-south-divides/) showing that three of the top five areas for depression and serious mental ill health are in the north of the country.

This paper opens up numerous avenues for future research, including both application orientated and methodological directions. From an application perspective we have focused exclusively on the spatial smoothing of noisy disease rates, but the disease mapping literature also focuses on quantifying the impacts of exposures such as air pollution (e.g., Lee et al., 2022). Thus, further work will examine whether the improved estimation performance observed here for disease rates carries over to covariate effects. In a methodological direction ridge regression was used here as the model fitting algorithm, due to its computational efficiency, closed form parameter estimates (in the Gaussian case), and the ease of constructing approximate 95% confidence intervals to accompany the fitted disease rates. However, it is unclear whether alternative machine learning algorithms, such as random forests and gradient boosting machines, would lead to improved disease rate estimation, and a particular challenge in these cases will be the quantification of uncertainty in the smoothed disease rates.

One limitation of this paper is that it only considers a single disease outcome over a purely spatial domain. Therefore, future work will extend the RAB framework to model multiple disease outcomes in space and time, which would require correlations over space, time and between disease outcomes to be modelled. A natural approach to capturing temporal autocorrelation would be to utilise the same set of anisotropic spatial basis functions for each time period, and smooth time specific regression parameters $\boldsymbol{\beta}_t$ via an autoregressive ridge penalty similar to that used in penalised splines (Eilers and Marx, 1996). Between disease correlations could be captured in numerous ways, including by replacing the univariate Gaussian noise model in (4) with a multivariate model with correlated errors. Alternatively, the disease specific regression parameters could be smoothed between diseases via an extended ridge-style penalty as suggested above, or a common set of basis functions could be added to the model that have been computed by averaging the disease specific basis functions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author would like to thank the editor and reviewer whose suggestions have improved the motivation for and content of the paper. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spasta.2023.100796.

References

Akhter, N., Bambra, C., Mattheys, K., Warren, J., Kasim, A., 2018. Inequalities in mental health and well-being in a time of austerity: Follow-up findings from the Stockton-on-Tees cohort study. SSM - Popul. Health 6, 75–84.

Berchuck, S., Mwanza, J., Warren, J., 2019. Diagnosing glaucoma progression with visual field data using a spatiotemporal boundary detection method. J. Amer. Statist. Assoc. 114, 1063–1074.

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration with two applications in spatial statistics. Ann. Inst. Stat. Math. 43, 1–59.

Brewer, M., Nolan, A., 2007. Variable smoothing in Bayesian intrinsic autoregressions. Environmetrics 18, 841–857.

Charras-Garrido, M, Abrial, D, Goer, J, Dachian, S, Peyrard, N, 2011. Classification method for disease risk mapping based on discrete hidden Markov random fields. Biostatistics 13, 241–255.

Djeudeu, D, Engel, M, Jöckel, K.-H, Moebus, S, Ickstadt, K, 2020. Spatio-temporal analysis of the risk of depression at district-level and association with greenness based on the Heinz Nixdorf Recall Study. Spat. Spatio-Temp. Epidemiol. 33, 100340.

Eilers, P., Marx, B., 1996. Flexible smoothing with B-splines and penalties. Statist. Sci. 11, 89–121.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33, 1–22.

Gao, L., Banerjee, S., Ritz, B., 2022. Spatial difference boundary detection for multiple outcomes using Bayesian disease mapping. Biostatistics 24, 922–944.

Guan, Y., Haran, M., 2018. A computationally efficient projection-based approach for spatial generalized linear mixed models. J. Comput. Graph. Statist. 27, 701–714.

Hoerl, A., Kennard, R., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12, 55–67.

Jack, E., Lee, D., Dean, N., 2019. Estimating the changing nature of Scotland's health inequalities by using a multivariate spatiotemporal model. J. R. Stat. Soc. Ser. A Stat. Soc. 182, 1061–1080.

Lee, D., 2013. CARBayes: An R package for Bayesian spatial modelling with conditional autoregressive priors. J. Stat. Softw. 55, 13.

Lee, D., Meeks, K., Pettersson, W., 2021. Improved inference for areal unit count data using graph-based optimisation. Stat. Comput. 31, 51.

Lee, D., Mitchell, R., 2012. Boundary detection in disease mapping studies. Biostatistics 13, 415–426.

Lee, D., Mitchell, R., 2013. Locally adaptive spatial smoothing using conditional auto-regressive models. J. R. Stat. Soc. Ser. C 62, 593–608.

Lee, D., Robertson, C., McRae, C., Baker, J., 2022. Quantifying the impact of air pollution on Covid-19 hospitalisation and death rates in Scotland. Spat. Spatio-Temp. Epidemiol. 42, 100523.

Leroux, B., Lei, X., Breslow, N., 2000. Estimation of Disease Rates in Small Areas: A New Mixed Model for Spatial Dependence. Springer-Verlag, New York, pp. 135–178.

Lu, H., Carlin, B., 2005. Bayesian areal wombling for geographical boundary analysis. Geograph. Anal. 37, 265–285.

Lund, C., De Silva, M., Plagerson, S., Cooper, S., Chisholm, D, Das, J, Knapp, M, Patel, V, 2011. Poverty and mental disorders: breaking the cycle in low-income and middle-income countries. Lancet 378, 1502–1514.

Ma, H., Carlin, B., Banerjee, S., 2010. Hierarchical and joint site-edge methods for medicare Hospice Service Region boundary analysis. Biometrics 66, 355–364.

MacNab, Y., 2022. Bayesian disease mapping: Past, present, and future. Spat. Stat. 50, 100593.

Moran, P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17–23.

Orozco-Acosta, E., Adin, A., Ugarte, M.D., 2021. Scalable Bayesian modelling for smoothing disease risks in large spatial data sets using INLA. Spat. Stat. 41, 100496.

Public Health England, 2021. Place-based approaches for reducing health inequalities: main report. https://www.gov.uk/government/publications/health-inequalities-place-based-approaches-to-reduce-inequalities/place-based-approaches-for-reducing-health-inequalities-main-report.

Riebler, A., Sørbye, S., Simpson, D., Rue, H., 2016. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. Stat. Methods Med. Res. 25, 1145–1165.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). J. R. Stat. Soc. Ser. B 71, 319–392.

Rushworth, A., Lee, D., Sarran, C., 2017. An adaptive spatiotemporal smoothing model for estimating trends and step changes in disease risk. J. R. Stat. Soc. Ser. C. Appl. Stat. 66, 141–157.

Ugarte, M., Goicoa, T., Etxeberria, J., Militino, A., 2012. A P-spline ANOVA type model in space-time disease mapping. Stoch. Environ. Res. Risk Assess. 26, 835–845.

Wakefield, J., Kim, A., 2013. A Bayesian model for cluster detection. Biostatistics 14, 752–765.

Womble, W., 1951. Differential systematics. Science 114, 315–322.