The Annals of Applied Statistics
2018, Vol. 12, No. 4, 2540–2558
https://doi.org/10.1214/18-AOAS1167

A LOCALLY ADAPTIVE PROCESS-CONVOLUTION MODEL FOR ESTIMATING THE HEALTH IMPACT OF AIR POLLUTION

BY DUNCAN LEE[1]

University of Glasgow

Most epidemiological air pollution studies focus on severe outcomes such as hospitalisations or deaths, but this underestimates the impact of air pollution by ignoring ill health treated in primary care. This paper quantifies the impact of air pollution on the rates of respiratory medication prescribed in primary care in Scotland, which is a proxy measure for the prevalence of less severe respiratory disease. A novel bivariate spatiotemporal process-convolution model is proposed, which: (i) has increased computational efficiency via a tapering function based on nearest neighbourhoods; and (ii) has locally adaptive weights that outperform traditional distance-decay kernels. The results show significant effects of particulate matter on respiratory prescription rates which are consistent with severe endpoint studies.

1. Introduction. In the United Kingdom 40,000 premature deaths are attributable to air pollution each year [Royal College of Physicians (2016)], and the epidemiological literature focuses almost exclusively on severe health outcomes such as hospital admissions [e.g., Huang, Lee and Scott (2018)] or deaths [e.g., Dominici, Samet and Zeger (2000)]. However, this underestimates the impact of air pollution because it ignores ill health treated in primary (nonhospitalised) care. One of the few exceptions is Blangiardo, Finazzi and Cameletti (2016) who linked air pollution concentrations to rates of respiratory medication in general practice (GP) surgeries in England, the latter being a proxy for the prevalence of respiratory disease not requiring hospitalisation.

In Scotland, the focus of this study, primary medical care is largely provided by doctors grouped within GP surgeries who prescribe medication that is generically referred to as a prescription. The aim of this study is to estimate the impact of particulate air pollution on respiratory prescription rates, a surrogate measure of nonhospitalised respiratory ill health. If positive associations are observed, then it provides evidence that the health burden from air pollution is larger than previously thought in Scotland because existing studies have only focused on severe outcomes such as hospitalisation rates [e.g., see Huang, Lee and Scott (2018)]. A secondary aim of this study is to quantify the magnitude of any health inequalities in respiratory prescription rates across Scotland, particularly focusing on variation between

Received October 2017; revised February 2018.

[1]Supported by the UK Medical Research Council (MRC) Grant MR/L022184/1.

Key words and phrases. Air pollution, bivariate spatiotemporal modelling, process-convolution models, respiratory medication rates.

the 14 regional health boards to which health care spending is devolved. The policy context for this study is the *Cleaner Air For Scotland* (CAFS) strategy, which will introduce air pollution reduction strategies, such as low emission zones, by the end of 2018. However, existing epidemiological studies in Scotland use relatively old data up to 2011 [e.g., Dibben and Clemens (2015), Huang, Lee and Scott (2018)], making current epidemiological evidence vital to inform these policies. Additionally, this is the first study in Scotland to quantify the impact of air pollution in a nonhospitalised setting and, unlike Blangiardo, Finazzi and Cameletti (2016), focuses on medications that prevent and relieve the symptoms of respiratory disease such as asthma and chronic obstructive pulmonary disease (COPD).

The respiratory prescription data in this study relate to GP surgeries that have a single geographical coordinate while their patient populations are drawn from the surrounding area. In urban areas the patient populations from spatially close surgeries will overlap, meaning that the data are not of areal unit type. Furthermore, they are not strictly in the geostatistical paradigm either because data for a GP surgery relates to its surrounding patient population and not to the measurement of a random quantity at a single location. Therefore, we propose a novel bivariate spatiotemporal process-convolution [PC, Higdon (1998)] model for the data, which represents spatial correlation as a spatially weighted moving average of a noise process. Our proposed model extends current PC models in two main ways. First, exploratory analysis shows that some pairs of spatially close GP surgeries have similar data values suggesting correlation while other pairs have very different values suggesting no correlation. Furthermore, the spatial correlations do not always decay with increasing distance apart, in that for GP surgeries (i, j, k) all close together if surgery i is geographically closer to surgery j than surgery k, then it often has a data value closer to that from surgery k than surgery j. We model this structure via a novel random weighting scheme using Dirichlet priors, which do not enforce a rigid parametric distance-decay form as kernel functions do.

The second methodological novelty of this paper is computational because PC models are expensive to implement in a Bayesian paradigm via Markov chain Monte Carlo (MCMC) simulation when the matrix of spatial weights is dense; see Section 3 for details. Therefore, we propose a novel computationally efficient tapering approach for making the weight matrix sparse. Computationally efficient modelling can be undertaken via numerous approaches, including covariance tapering [Furrer, Genton and Nychka (2006)], low-rank models [Banerjee et al. (2008)] and nearest-neighbour models [Datta et al. (2016)]. Our approach fuses covariance tapering with nearest-neighbour modelling by applying a tapering function to the weight matrix based on being one of the nearest neighbours. The methodology is presented in Section 3 following the exploratory analysis in Section 2, while the results are presented in Section 4. Finally, the paper concludes with a discussion in Section 5.

2. Motivating study. Scotland contains 966 GP surgeries arranged within 14 regional health boards, although data are only available for $K = 939$ surgeries due to missing data problems (see below). The numbers of surgeries per health board varies between six (Orkney) and 239 (Greater Glasgow and Clyde), and they are unevenly distributed across the country and concentrated mainly in the cities. Further details are provided in Section 1 of the Supplementary Material [Lee (2018)].

2.1. *Prescription data.* Newly released monthly data on respiratory prescribing rates were obtained from the Information Services Division of NHS Scotland between October 2015 and August 2016, and are considered as a proxy for the prevalence of respiratory disease not requiring hospitalisation. The data comprise counts for each GP surgery and month of the total numbers of prescriptions for medication that: (i) prevent (corticosteroids); and (ii) relieve (short-acting β_2 agonists), the symptoms of respiratory conditions. The monthly scale was retained because temporal aggregation would make the data less finely resolved and further away from the individual level. The monthly prescription counts at a surgery will depend on the size and age-sex structure of its patient population, which is accounted for using indirect standardisation. Specifically, the number of male and female patients in age groups: 0–4, 5–14, 15–24, 25–44, 45–64, 65–74, 75–84 and 85+ were obtained for 2015 and 2016 (monthly data were not available) for each surgery, which were then multiplied by national age-sex specific rates of asthma and COPD and then summed to estimate the expected number of patients with respiratory disease for each surgery. These expected numbers were then scaled, separately for preventer and reliever medication, so that the total of the observed and expected counts were equal for each medication. However, 27 surgeries had missing patient size data and thus were removed.

The standardised prescription rate (SPR) equals the observed divided by the expected count of prescriptions for each surgery, month and medication type (reliever or preventer) and is equivalent to a standardised mortality ratio (SMR) when modelling death rates. For interpretation an SPR of 1.2 represents a 20% increased rate compared to the Scotland average. The monthly temporal patterns [shown in Section 1 of the Supplementary Material, Lee (2018)] in the SPR show little month-to-month variation with only December having markedly increased rates. The average (mean over month) spatial patterns in the SPR for preventer and reliever medications are displayed in Figure 1 where the SPR has been classified into three groups: Low (<0.9), Average ($[0.9, 1.1]$) and High (>1.1), to aid the visualisation. The main high rate areas are the city of Glasgow and parts of southern Scotland where as low rates are observed in Grampian and Edinburgh. Finally, the SPRs for the two medication types have a Pearson's correlation coefficient of 0.60, and further exploratory analysis is given in Section 1 of the Supplementary Material [Lee (2018)].

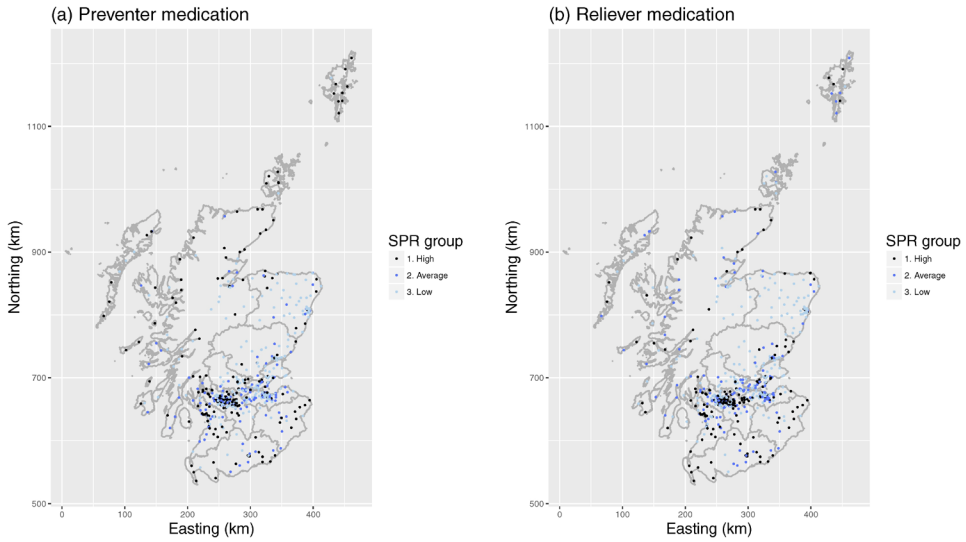

FIG. 1. *The spatial distribution in average SPR for*: (a) *preventer medication*; *and* (b) *reliever medication. The SPR has been categorised as*: *Low* (<0.9), *Average* ([0.9, 1.1]) *and High* (>1.1). *The grey lines denote the health board boundaries*.

2.2. *Air pollution*. Data on concentrations of particulate matter less than 10 (PM_{10}) and 2.5 ($PM_{2.5}$) microns in size (measured in μgm^{-3}) are available from two sources. The first are measured monthly mean concentrations at 76 point-locations for PM_{10} and 17 point-locations for $PM_{2.5}$, which were obtained from http://www.scottishairquality.co.uk. These data contain 6.7% and 7.2% missing values respectively for each pollutant over the 11 months. The locations of these pollution monitors are clustered mainly in Glasgow and Edinburgh with large parts of Scotland having no observations, which makes spatial prediction via Kriging difficult. Therefore, a second source of pollution data we use are modelled annual average concentrations provided by the Department for the Environment, Food and Rural Affairs (DEFRA, https://uk-air.defra.gov.uk/data/pcm-data), which are on a regular 1-kilometer grid and provide complete spatial coverage of Scotland. Further details of the pollution data are provided in Section 2 of the Supplementary Material [Lee (2018)].

2.3. *Other covariates*. The main confounder for respiratory illness is smoking, but as no data are available to quantify smoking rates we use socioeconomic deprivation as a proxy measure due to its strong relationship. Specifically, the percentage of patients from each GP surgery who live in the 15% most deprived data-zones (a small-area geography), as measured by the Scottish Index of Multiple Deprivation (SIMD), is available as is the median property price surrounding each GP surgery. These covariates only vary in space and not in time because monthly level

data are not available. Meteorology also impacts respiratory health, and monthly modelled average temperatures and relative humidities were obtained from the Centre for Environmental Data Analysis (2016, 2017). Additionally, the percentage of patients from each GP surgery who are white was also included (the only ethnicity variable available) to adjust for ethnic differences in respiratory disease rates [Bhopal et al. (2015)]. Finally, we include an indicator variable for December because it exhibits around a 30% (preventer) and a 25% (reliever) higher rate of prescription and is a clear outlier compared to the other months. This December phenomenon is well known amongst GPs and is caused by the public stocking up on medication during the Christmas holidays when surgeries are closed. Thus, this variable is essentially an indicator for Christmas holidays.

2.4. *Exploratory analysis*. Initially, overdispersed quasi-Poisson log-linear models (where $\texttt{variance} = \xi \times \texttt{mean}$) were fitted to the prescription count data separately for each medication type, and the covariates included the expected numbers of prescriptions (an offset), $PM_{2.5}$ concentrations and the confounding factors described above. The residuals exhibited substantial overdispersion ($\hat{\xi} = 8.2$) and spatiotemporal and between medication correlations, the latter of which appear to be separable in space, time and between medications [see Section 3 of the Supplementary Material, Lee (2018)] which motivates the model proposed in the next section. However, the spatial autocorrelation is not universal because there are numerous pairs of spatially close GP surgeries that have very different residual values. To illustrate, we considered the absolute differences in the residuals between each GP surgery and its eight geographically closest neighbours. Table 1 presents the percentages of GP surgeries for which the most similar residual (in terms of absolute difference) was from its closest, second closest, third closest, etc. neighbouring surgery, where for ease of presentation the results are averaged over the 11 months of data.

TABLE 1
Percentage of surgeries whose residual is most similar to that of the kth closest surgery. Only the eight closest surgeries were considered

***k*th closest surgery**	**Preventer medication**	**Reliever medication**
1	17.89%	16.93%
2	14.70%	16.51%
3	13.63%	12.99%
4	11.93%	10.76%
5	12.35%	11.71%
6	9.80%	11.40%
7	9.69%	8.95%
8	10.01%	10.76%

The table shows that the percentages decrease as one moves further away from the surgery in question as expected, but that the percentages of surgeries for which the spatially closest residual was the most similar was only 16%–18%. This suggests that, within a spatially close set of GP surgeries, modelling correlation in terms of a simple distance-decay form is not appropriate as the absolute differences do not simply increase with increasing distance apart. Furthermore, the similarities are not reversible, as in only 42.4% (preventer medication) and 42.9% (reliever medication) of cases does it follow that if surgery k has a residual that is most similar to that from surgery j, then the residual from surgery j is most similar to that from surgery k. This asymmetry suggests that the weights in the process convolution model described in the next section should not be symmetric.

3. Methodology. We propose a novel bivariate locally adaptive spatiotemporal process-convolution model to estimate the effects of air pollution on respiratory prescribing rates with inference in a Bayesian framework using MCMC simulation. A bivariate model is chosen because the data are naturally bivariate given the two different medication types being modelled. Additionally, given the highly parameterised nature of the localised weight model proposed in this section, a bivariate approach provides double the number of data points with which to estimate the localised weights compared to a univariate model. The model is outlined below, whilst details of the MCMC algorithm, reproducibility materials and the model's correlation structures are presented in Sections 4 and 5 of the Supplementary Material [Lee (2018)].

3.1. *Overall model.* Let $\mathbf{s}_k = (s_{k1}, s_{k2})$ for $k = 1, \ldots, K = 939$ denote the geographical coordinates of the kth GP surgery, while $(Y_{ti}(\mathbf{s}_k), e_{ti}(\mathbf{s}_k))$ respectively denote the observed and expected numbers of prescriptions for surgery k in month t $(t = 1, \ldots, N = 11)$ for medication type i, $(i = 1$ is preventer and $i = 2$ is reliever). Additionally, let $\mathbf{x}_t(\mathbf{s}_k) = (x_{t1}(\mathbf{s}_k), \ldots, x_{tp}(\mathbf{s}_k))$ denote a $p \times 1$ vector of covariates including a column of ones for the intercept term, while $\tilde{Z}_t(\mathbf{s}_k)$ denotes the pollution concentrations. Then the first level of the proposed model is given by

$$
\begin{aligned}
Y_{ti}(\mathbf{s}_k) &\sim \text{Poisson}\big(e_{ti}(\mathbf{s}_k) r_{ti}(\mathbf{s}_k)\big), \\
\ln\big(r_{ti}(\mathbf{s}_k)\big) &= \mathbf{x}_t(\mathbf{s}_k)^\top \boldsymbol{\beta}_i + \tilde{Z}_t(\mathbf{s}_k)\beta_i^z + \sum_{j=1}^{K} w_{kj}\theta_{ti}(\mathbf{s}_j), \\
\big(\boldsymbol{\beta}_i, \beta_i^z\big) &\sim \text{N}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta).
\end{aligned} \tag{3.1}
$$

A negative binomial data model was also considered, but after including the process-convolution $\phi_{ti}(\mathbf{s}_k) = \sum_{j=1}^{K} w_{kj}\theta_{ti}(\mathbf{s}_j)$ no overdispersion remained (see Section 4.1). In (3.1) $r_{ti}(\mathbf{s}_k)$ denotes the relative prescribing rate for medication i compared to the scaled expected numbers $e_{ti}(\mathbf{s}_k)$ and has the same interpretation as

the SPR. The log rate is modelled by covariates and a bivariate PC latent process—the latter modelling any unmeasured spatiotemporal and between medication correlation in the data after covariate adjustment. The regression parameters $(\boldsymbol{\beta}_i, \beta_i^z)$ are medication specific and are assigned a weakly informative Gaussian prior with a zero mean ($\boldsymbol{\mu}_\beta = \mathbf{0}$) and a large diagonal variance matrix ($\boldsymbol{\Sigma}_\beta = 1000\mathbf{I}$).

The PC $\phi_{ti}(\mathbf{s}_k)$ is a spatially weighted average of the latent process $\{\theta_{ti}(\mathbf{s}_j)\}$ at the K data locations, and spatial autocorrelation is induced by the spatial weight matrix $\mathbf{W} = (w_{kj})_{K\times K}$. Here w_{kj} is large for geographically close surgeries (k, j) and small for surgeries further apart, and further details are given in Section 3.3. Temporal and between medication correlations are induced by modelling $\boldsymbol{\theta}_t(\mathbf{s}_k) = (\theta_{t1}(\mathbf{s}_k), \theta_{t2}(\mathbf{s}_k))$ as a spatially independent bivariate first order autoregressive process:

$$
\begin{aligned}
&\boldsymbol{\theta}_t(\mathbf{s}_k) \sim \mathrm{N}\big(\gamma\boldsymbol{\theta}_{t-1}(\mathbf{s}_k), \boldsymbol{\Sigma}\big) \qquad \text{for } t = 2, \ldots, N,\\
&\boldsymbol{\theta}_1(\mathbf{s}_k) \sim \mathrm{N}(0, \boldsymbol{\Sigma}),\\
&\ln\left[\frac{1+\gamma}{1-\gamma}\right] \sim \mathrm{N}(0, 1/0.4),\\
&f_J(\boldsymbol{\Sigma}) \propto 1/|\boldsymbol{\Sigma}|^{3/2}.
\end{aligned} \tag{3.2}
$$

Here γ controls the temporal autocorrelation, with $\gamma = 0$ corresponding to temporal independence while $\gamma = 1$ corresponds to a nonstationary temporally dependent random walk. It is assigned a Gaussian prior on the transformed $\ln[(1+\gamma)/(1-\gamma)]$ scale following the suggestion of Riebler, Held and Rue (2012), and the prior precision is 0.4 as this provides a relatively noninformative prior [see Figure 2 in Riebler, Held and Rue (2012)]. The conditional covariance of $\boldsymbol{\theta}_t(\mathbf{s}_j)$ is represented by $\boldsymbol{\Sigma}$, which is assigned a weakly informative Jeffeys prior [for details see Liechty, Liechty and Müller (2004)]. A sensitivity analysis to these choices is presented in Section 6 of the Supplementary Material [Lee (2018)].

3.2. *Pollution model for* $\tilde{Z}_t(\mathbf{s}_k)$. Monthly concentrations of $\mathrm{PM}_{2.5}$ and PM_{10} need to be predicted for each GP surgery using both the measured and modelled pollution data described in Section 2. In the main analysis concentrations are predicted at each surgeries coordinates, while Section 6 of the Supplementary Material [Lee (2018)] presents a sensitivity analysis by instead using the average concentration within 1 km and 2 km circular buffers of each surgery. A preliminary cross validation exercise was undertaken to compare a number of models for predicting pollution concentrations, and details are given in Section 2 of the Supplementary Material [Lee (2018)]. A simple linear model performed best, where the square root of the monitored concentrations was regressed against the modelled concentrations (also square rooted), a monthly factor variable and the site type (a *background* or a *roadside* site) variable. Thus, letting

$\mathbf{Z} = (Z_1(\mathbf{p}_1), \ldots, Z_N(\mathbf{p}_D))_{ND\times 1}$ represent the vector of monthly averaged pollution concentrations at the D monitor locations $(\mathbf{p}_1, \ldots, \mathbf{p}_D)$ for all time periods, the model has the form

$$\begin{aligned} \sqrt{\mathbf{Z}} &\sim \mathrm{N}(\mathbf{B}\boldsymbol{\omega}, \tau^2\mathbf{I}), \\ f(\boldsymbol{\omega}, \tau^2) &\propto \tau^{-2}, \end{aligned} \tag{3.3}$$

where $\mathbf{B}_{ND\times 14}$ denotes the design matrix of covariates. Additionally, $\mathbf{I}$ is the $ND \times ND$ identity matrix, $\boldsymbol{\omega}$ are the corresponding regression parameters, and τ^2 is the error variance. A square root transformation is applied because the pollution data are nonnegative and skewed to the right. The joint prior distribution $f(\boldsymbol{\omega}, \tau^2)$ was suggested by Gelman et al. (2004) to ensure conjugacy with the data likelihood, allowing the posterior predictive distributions to have a closed form (on the square root scale). Let $\tilde{\mathbf{Z}} = (\tilde{Z}_1(\mathbf{s}_1), \ldots, \tilde{Z}_N(\mathbf{s}_K))_{NK\times 1}$ denote the unknown pollution concentrations at the K GP surgery locations for all N months; then their posterior predictive distribution is

$$\sqrt{\tilde{\mathbf{Z}}}|\sqrt{\mathbf{Z}} \sim \text{Multivariate-t}_{NK-14}(\mathbf{m}_{\tilde{Z}} = \tilde{\mathbf{B}}\hat{\boldsymbol{\omega}}, \mathbf{V}_{\tilde{Z}} = s^2[\mathbf{I} + \tilde{\mathbf{B}}\mathbf{V}\tilde{\mathbf{B}}^\top]), \tag{3.4}$$

where $\hat{\boldsymbol{\omega}} = (\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{z}$, $\mathbf{V} = (\mathbf{B}^\top\mathbf{B})^{-1}$ and $s^2 = \frac{1}{NK-14}(\mathbf{z} - \mathbf{B}\hat{\boldsymbol{\omega}})^\top(\mathbf{z} - \mathbf{B}\hat{\boldsymbol{\omega}})$. The degrees of freedom $(NK - 14)$ comes from the choice of covariates outlined above, while $\tilde{\mathbf{B}}$ denotes the covariates at the prediction locations where in all cases the site type is set to *background* rather than *roadside* so as to be more representative of population exposure (i.e., people don't spend their lives next to a main road).

Two approaches are considered for combining (3.1) and (3.4), and the results are compared to assess its impact on the estimated pollution-health association. The first assumes that the GP surgery pollution concentrations are fixed and equal to the posterior predictive mean of (3.4), which is the most common approach in the literature and thus allows a comparison with previous work. However, this ignores the predictive uncertainty in (3.4), and thus the second approach feeds the uncertainty in the pollution predictions forward into the disease model (3.1). Specifically, a new sample of $\tilde{\mathbf{Z}}$ is generated from (3.4) at each MCMC iteration when fitting the health model, which allows for the pollution uncertainty. This approach cuts the feedback between the pollution and disease models [see also Blangiardo, Finazzi and Cameletti (2016), Lee et al. (2017)] because it makes no sense to allow the disease data to influence the pollution concentrations as this is first, biologically implausible, and second, it is the relationship in the other direction one is trying to estimate.

3.3. *Weight model for* $\mathbf{W}$. We compare two approaches for modelling the spatial weights $\mathbf{W}$, a kernel-smoothing approach similar to Higdon (1998) and a novel locally adaptive specification that is a methodological contribution of this paper. Another methodological contribution is computational because fitting model

(3.1)–(3.2) using a Bayesian MCMC algorithm is computationally intensive due to the need for repeated evaluation of the data likelihood. For example, the full conditional distribution of $\boldsymbol{\theta}_t(\mathbf{s}_b)$ is proportional to

$$
\begin{aligned}
&f\big(\boldsymbol{\theta}_t(\mathbf{s}_b)|-\big)\\
&\quad\propto \prod_{i=1}^{2}\prod_{k=1}^{K}\text{Poisson}\left(Y_{ti}(\mathbf{s}_k)\,\middle|\, e_{ti}(\mathbf{s}_k)\exp\left\{\mathbf{x}_t(\mathbf{s}_k)^\top\boldsymbol{\beta}_i+\sum_{j=1}^{K}w_{kj}\theta_{ti}(\mathbf{s}_j)\right\}\right)\\
&\qquad\times \mathrm{N}\big(\boldsymbol{\theta}_t(\mathbf{s}_b)|\gamma\boldsymbol{\theta}_{t-1}(\mathbf{s}_b),\boldsymbol{\Sigma}\big)\mathrm{N}\big(\boldsymbol{\theta}_{t+1}(\mathbf{s}_b)|\gamma\boldsymbol{\theta}_t(\mathbf{s}_b),\boldsymbol{\Sigma}\big),
\end{aligned}
$$

where '$-$' denotes the data and all other parameters. Thus updating $\boldsymbol{\theta}_t(\mathbf{s}_b)$ requires evaluating the data likelihood for all $2K$ data points at time t, which when iterated over all $K\times N$ elements $\{\boldsymbol{\theta}_t(\mathbf{s}_b)\}$ for a large number of MCMC iterations is computationally demanding. Therefore, we propose a novel tapering approach based on the nearest neighbour ideas of Datta et al. (2016), which makes $\mathbf{W}$ sparse and hence removes most of the data likelihood terms from the above full conditional resulting in faster inference (see Section 4 for details). Specifically, for the kth GP surgery we only allow the weights $\{w_{k1},\dots,w_{kK}\}$ to be nonzero for the noise process $\{\boldsymbol{\theta}_t(\mathbf{s}_1),\dots,\boldsymbol{\theta}_t(\mathbf{s}_K)\}$ at the m nearest GP surgeries (including at the surgery itself via a nonzero w_{kk}), where the sensitivity of the results to the choice of m is evaluated in the next section. This is achieved by combining the weights described below with the tapering function

$$
I(\mathbf{s}_k,\mathbf{s}_j)=\begin{cases}1 & \text{if location } \mathbf{s}_j \text{ is one of the } m\\ & \text{closest points to location } \mathbf{s}_k,\\ 0 & \text{otherwise.}\end{cases}
\tag{3.5}
$$

We base the tapering function on a fixed number of closest neighbours m rather than on a fixed threshold distance from location $\mathbf{s}_k$ because the data locations are irregularly spaced. Thus, a single threshold distance would not be simultaneously appropriate in a city (where lots of surgeries are close by) and in a rural area (where no surgeries are close by).

3.3.1. *Kernel smoothing weights*. To provide a comparator to the locally adaptive weights proposed below, we combine the kernel smoothing weights proposed by Higdon (1998) with the tapering function (3.5) to give

$$
\begin{aligned}
w_{kj}&=\frac{I(\mathbf{s}_k,\mathbf{s}_j)\frac{1}{\sqrt{2\pi/\alpha_k}}\exp(-\alpha_k\frac{\|\mathbf{s}_k-\mathbf{s}_j\|}{2})}{\sum_{i=1}^{K}I(\mathbf{s}_k,\mathbf{s}_i)\frac{1}{\sqrt{2\pi/\alpha_k}}\exp(-\alpha_k\frac{\|\mathbf{s}_k-\mathbf{s}_i\|}{2})},\\
\alpha_k&=\mathbf{v}_k^\top\boldsymbol{\delta},\\
\delta_r&\sim\text{Uniform}(0,100),\qquad r=1,\dots,q,
\end{aligned}
\tag{3.6}
$$

where $\|\cdot\|$ denotes Euclidean distance. These weights induce spatial smoothness into $\phi_{ti}(\mathbf{s}_k) = \sum_{j=1}^{K} w_{kj}\theta_{ti}(\mathbf{s}_j)$ because the closer two locations $(\mathbf{s}_k, \mathbf{s}_j)$ are the larger the weight w_{kj} is, as $\alpha_k > 0$. The speed at which the weights decay to zero with increasing distance is controlled by the bandwidth α_k with small values corresponding to long-range spatial smoothness, while as $\alpha_k \to \infty$ then $w_{kk} \to 1$ resulting in spatial independence. The bandwidth varies in space via the regression model $\alpha_k = \mathbf{v}_k^\top \boldsymbol{\delta}$ where both the covariates and their regression parameters are constrained to be positive as $\alpha_k > 0$.

3.3.2. *Locally adaptive weights.* The weights in (3.6) vary by GP surgery via a spatially varying bandwidth α_k but are still constrained to decay monotonically as the distance between two surgeries increases. However, Section 2.4 and Table 1 shows this is unrealistic for the prescription data, because within a group of spatially close GP surgeries the absolute differences in the residuals from a covariate only model do not increase with increasing distance apart. Specifically, the smallest absolute difference is between the geographically closest surgeries in only 16%–18% of cases, suggesting that within a set of spatially close GP surgeries, weights that monotonically decrease with increasing distance apart [as used in Meyer and Held (2014)] would not be appropriate. Furthermore, there is an asymmetry in these spatial similarities, in that only 43% of the time does it hold that if the residual in surgery k is most similar to that in surgery j, then the residual in surgery j is most similar to that in surgery k. These two observations suggest that for the m nearest GP surgeries the weights should not be constrained to decay monotonically with increasing distance or be symmetric (i.e., $w_{kj} \neq w_{jk}$). This motives the following locally adaptive weight model for $\mathbf{W}$:

$$w_{kj} = \begin{cases} \psi_{kr} & \text{if } \mathbf{s}_j \text{ is the } r\text{th closest point to } \mathbf{s}_k \text{ for } r = 1, \ldots, m, \\ 0 & \text{otherwise,} \end{cases} \tag{3.7}$$

$$\boldsymbol{\psi}_k = (\psi_{k1}, \ldots, \psi_{km}) \sim \text{Dirichlet}(\alpha_1 = 1, \ldots, \alpha_m = 1),$$

where $\boldsymbol{\psi}_k$ represents the nonzero weights from the m nearest surgeries to surgery k. A Dirichlet prior is assigned to $\boldsymbol{\psi}_k$ so that they sum to one, and they do not vary over time for identifiability reasons because the number of weights $K \times m$ is larger than the number of spatial data points K. Spatial autocorrelation is enforced by this model due to the taper function because only the m geographically closest $\theta_{ti}(\mathbf{s}_j)$ values are given nonzero weight in the PC $\phi_{ti}(\mathbf{s}_k)$. However, the use of random weights means that spatially close elements $(\phi_{ti}(\mathbf{s}_k), \phi_{ti}(\mathbf{s}_i))$ can be highly autocorrelated if the weights (w_{ki}, w_{ik}) are high, but can also be close to independent if close to zero weights are estimated, which thus allows for the localised structure observed in the data.

4. Results. The locally adaptive model (3.7) and the kernel smooth model (3.6) with constant ($\alpha_k = \alpha$ for a baseline comparison) and spatially adaptive bandwidths were run with $m = 4, 8, 16$ which assesses model sensitivity to the amount of tapering. For the latter α_k varies by how urban a surgery is (classes *urban*, *small-town* or *rural*) because in cities GP surgeries are close together allowing short range correlations, while longer range correlations are necessary in rural areas. Each model was run separately including $PM_{2.5}$ or PM_{10} (correlation 0.85) as well as deprivation based on SIMD and property price, temperature, relative humidity, the percentage of patients who are white and an indicator variable for December (Christmas holidays). Inference is based on three Markov chains burnt in for 100,000 iterations, and convergence was checked using trace plots and the Geweke statistic. A further 100,000 samples were then generated and thinned by 10 to reduce their autocorrelation resulting in 30,000 samples for inference. The main results are presented below, while a sensitivity analysis is presented in Section 6 of the Supplementary Material [Lee (2018)].

4.1. *Overall model fit.* Table 2 summarises the overall fit of each model by the Watanabe–Akaike Information Criteria [WAIC, Watanabe (2010)] and the Log Marginal Predictive Likelihood [LMPL, Congdon (2005)], and both show little sensitivity to changing the tapering parameter m with a maximum change of 0.6%. However, increasing m increases the computational burden greatly, for example,

TABLE 2
Summary of the overall fit of the models by WAIC and LMPL (including each pollutant), and its relative computing time. For the latter the single α model with $m = 4$ is the baseline, with a relative time of 1

	m	Single α,	Varying α_k	Localised
WAIC (p.w)	4	171,000 (8187)	170,990 (8180)	147,624 (5115)
($PM_{2.5}$)	8	170,917 (8105)	170,907 (8099)	146,972 (4867)
	16	170,912 (8102)	170,903 (8097)	146,724 (4771)
WAIC (p.w)	4	170,942 (8158)	170,934 (8154)	147,641 (5117)
(PM_{10})	8	170,863 (8078)	170,857 (8081)	146,966 (4865)
	16	170,867 (8081)	170,855 (8077)	146,737 (4773)
LMPL	4	−77,735	−77,734	−74,688
($PM_{2.5}$)	8	−77,738	−77,739	−74,282
	16	−77,740	−77,738	−74,144
LMPL	4	−77,732	−77,730	−74,699
(PM_{10})	8	−77,735	−77,735	−74,284
	16	−77,738	−77,734	−74,153
Relative time	4	1	1.07	1.21
	8	1.69	1.77	1.89
	16	3.07	3.12	3.30

by a factor of 3 from $m = 4$ to $m = 16$. In contrast the locally adaptive model is only around 10% to 20% more computationally demanding than the kernel model. WAIC and LMPL evidence that the locally adaptive model fits the data better than either of the kernel weight models, with percentage improvements of around 16% (WAIC) and 5% (LMPL) respectively. In contrast moving from a constant to a spatially-varying bandwidth in the kernel models results in little improvement in model fit. Additionally, unmeasured overdispersion, quantified by the scaled sum of the Pearson residuals, is present for the kernel models, with estimated overdispersion parameters around 2.4 that are consistent across m. In contrast, for the locally adaptive model there is slight underdispersion (values around 0.7) consistent across m, which again suggests that this model is better able to capture the unmeasured structure in the data.

4.2. *Covariate effects*. Table 3 displays estimated relative rates (posterior medians) and 95% credible intervals for the increase in each covariate given in brackets in the first column of the table. With the exception of PM_{10} the results relate to when $PM_{2.5}$ was the pollutant included in the model. The estimated relative rates show little sensitivity to the choice of m.

4.2.1. *Pollution effects*. The pollution relative rates correspond to when the concentrations are fixed at their posterior predictive means from (3.4) because this makes the results comparable with most of the existing literature. The main finding is that both $PM_{2.5}$ and PM_{10} are associated with increased rates of prescription for preventer and reliever respiratory medication, with between a 1.8% and a 2.8% increased prescription rate for a 2 μgm^{-3} increase in each pollutant. The estimated effect sizes are slightly larger for $PM_{2.5}$ compared to PM_{10}, which may be because smaller particles can travel further into the lungs or because the magnitude of the $PM_{2.5}$ concentrations are smaller than those of PM_{10}, and as both pollutants are highly correlated their estimated regression parameters will thus be larger. The estimated effects are slightly larger for reliever medication compared with preventer medication with, for example, 2.8% and 2.0% increased rates being estimated for $PM_{2.5}$ from the locally adaptive model when $m = 8$. The posterior median estimates are similar between the 3 weight models, with the main difference being a widening of the credible intervals for the locally adaptive model. This widening corresponds to the locally adaptive process-convolution accounting for more variation in the data which is illustrated by the comparison of the overdispersion parameters in the previous section.

The locally adaptive model with $m = 8$ was then fitted to the data allowing for pollution uncertainty, and the estimated relative rates and 95% credible intervals are as follows: Preventer medication: $PM_{2.5}$—1.005 (0.997, 1.012); PM_{10}—1.002 (0.998, 1.006); Reliever medication: $PM_{2.5}$—1.008 (1.003, 1.014); PM_{10}—1.003 (1.001, 1.006). In all cases the relative rates from assuming pollution was fixed have been greatly attenuated to the null rate of 1, which is not surprising because

TABLE 3

Estimated relative rates and 95% *credible intervals for the increase in each covariate given in brackets in column* 1. *All results except for those of* PM_{10} *relate to when* $PM_{2.5}$ *was the pollutant included in the model*

	m	Single α,	Varying α_k	Localised
Preventer medication				
$PM_{2.5}$	4	1.022 (1.015, 1.029)	1.022 (1.015, 1.030)	1.021 (1.011, 1.030)
($2\ \mu gm^{-3}$)	8	1.022 (1.015, 1.029)	1.022 (1.015, 1.029)	1.020 (1.009, 1.031)
	16	1.022 (1.015, 1.029)	1.022 (1.015, 1.030)	1.018 (1.004, 1.033)
PM_{10}	4	1.019 (1.014, 1.024)	1.019 (1.014, 1.024)	1.019 (1.012, 1.026)
($2\ \mu gm^{-3}$)	8	1.019 (1.014, 1.024)	1.019 (1.014, 1.024)	1.019 (1.011, 1.027)
	16	1.019 (1.014, 1.024)	1.019 (1.014, 1.024)	1.018 (1.008, 1.028)
Price	4	0.979 (0.966, 0.992)	0.979 (0.966, 0.991)	0.965 (0.950, 0.981)
(*£57,461*)	8	0.979 (0.967, 0.992)	0.980 (0.966, 0.993)	0.985 (0.967, 1.002)
	16	0.980 (0.967, 0.993)	0.979 (0.966, 0.992)	0.988 (0.971, 1.006)
SIMD	4	1.162 (1.148, 1.176)	1.162 (1.148, 1.177)	1.106 (1.086, 1.126)
(*19.6%*)	8	1.163 (1.149, 1.178)	1.164 (1.149, 1.177)	1.126 (1.103, 1.148)
	16	1.164 (1.150, 1.178)	1.164 (1.150, 1.177)	1.131 (1.109, 1.154)
White	4	1.038 (1.025, 1.051)	1.039 (1.026, 1.053)	1.021 (1.005, 1.038)
(*6.6%*)	8	1.038 (1.025, 1.051)	1.038 (1.025, 1.052)	1.019 (1.000, 1.038)
	16	1.038 (1.025, 1.051)	1.038 (1.025, 1.051)	1.022 (1.002, 1.042)
Humidity	4	1.014 (1.008, 1.019)	1.014 (1.008, 1.019)	1.013 (1.006, 1.020)
(*2.35%*)	8	1.014 (1.009, 1.019)	1.014 (1.009, 1.019)	1.013 (1.004, 1.022)
	16	1.014 (1.009, 1.019)	1.014 (1.009, 1.019)	1.012 (1.001, 1.023)
Temperature	4	1.004 (0.996, 1.011)	1.004 (0.996, 1.011)	1.007 (0.996, 1.019)
(*4.05°C*)	8	1.003 (0.996, 1.011)	1.004 (0.996, 1.011)	1.009 (0.995, 1.024)
	16	1.004 (0.996, 1.011)	1.003 (0.996, 1.011)	1.014 (0.993, 1.035)
December	4	1.254 (1.237, 1.271)	1.254 (1.237, 1.271)	1.259 (1.238, 1.281)
	8	1.253 (1.236, 1.270)	1.253 (1.237, 1.270)	1.263 (1.237, 1.290)
	16	1.254 (1.237, 1.271)	1.253 (1.237, 1.270)	1.264 (1.231, 1.297)

allowing for measurement/prediction error in the pollutant makes an association harder to identify as the covariate value itself is uncertain. Similar attenuation was observed by Lee et al. (2017) and is also observed in classical measurement error models where covariates contain uncertainty.

4.2.2. *Other covariate effects*. Socioeconomic deprivation has a large impact on respiratory prescription rates with both the property price and SIMD covariates showing largely significant effects that point to higher rates of medication for less affluent communities. These effects are more pronounced for reliever medication with, for example, the relative rates for SIMD (locally adaptive model when $m = 8$) being 1.126 for preventer medication compared to 1.215 for reliever medication. The effect of the December Christmas holidays is more consistent between

TABLE 3
(*Continued*)

	m	Single α,	Varying α_k	Localised
Reliever medication				
$PM_{2.5}$	4	1.026 (1.021, 1.031)	1.026 (1.021, 1.031)	1.027 (1.021, 1.035)
($2\ \mu gm^{-3}$)	8	1.026 (1.021, 1.031)	1.026 (1.021, 1.031)	1.028 (1.020, 1.036)
	16	1.026 (1.021, 1.031)	1.026 (1.021, 1.031)	1.026 (1.015, 1.037)
PM_{10}	4	1.022 (1.019, 1.025)	1.022 (1.019, 1.025)	1.023 (1.018, 1.027)
($2\ \mu gm^{-3}$)	8	1.022 (1.019, 1.025)	1.022 (1.019, 1.025)	1.023 (1.017, 1.029)
	16	1.022 (1.019, 1.025)	1.022 (1.019, 1.025)	1.021 (1.014, 1.029)
Price	4	0.925 (0.915, 0.934)	0.925 (0.915, 0.934)	0.916 (0.904, 0.928)
(*£57,461*)	8	0.925 (0.915, 0.935)	0.925 (0.916, 0.934)	0.921 (0.909, 0.933)
	16	0.925 (0.916, 0.934)	0.924 (0.915, 0.934)	0.921 (0.910, 0.934)
SIMD	4	1.248 (1.238, 1.258)	1.248 (1.238, 1.259)	1.201 (1.186, 1.216)
(*19.6%*)	8	1.248 (1.237, 1.258)	1.248 (1.238, 1.258)	1.215 (1.198, 1.23)
	16	1.248 (1.238, 1.259)	1.248 (1.238, 1.257)	1.218 (1.199, 1.238)
White	4	1.057 (1.046, 1.067)	1.058 (1.047, 1.068)	1.036 (1.024, 1.048)
(*6.6%*)	8	1.055 (1.044, 1.065)	1.055 (1.045, 1.065)	1.020 (1.005, 1.035)
	16	1.054 (1.045, 1.065)	1.055 (1.044, 1.064)	1.015 (1.000, 1.030)
Humidity	4	1.010 (1.006, 1.013)	1.009 (1.006, 1.013)	1.011 (1.006, 1.016)
(*2.35%*)	8	1.010 (1.006, 1.013)	1.010 (1.006, 1.013)	1.011 (1.005, 1.017)
	16	1.010 (1.006, 1.013)	1.010 (1.006, 1.013)	1.011 (1.002, 1.019)
Temperature	4	1.008 (1.003, 1.014)	1.008 (1.003, 1.014)	1.009 (1.000, 1.018)
(*4.05°C*)	8	1.008 (1.002, 1.014)	1.008 (1.002, 1.014)	1.010 (0.999, 1.021)
	16	1.008 (1.002, 1.014)	1.008 (1.003, 1.014)	1.014 (0.998, 1.030)
December	4	1.234 (1.224, 1.245)	1.234 (1.224, 1.244)	1.240 (1.226, 1.256)
	8	1.233 (1.223, 1.244)	1.234 (1.223, 1.244)	1.245 (1.227, 1.264)
	16	1.234 (1.223, 1.244)	1.233 (1.222, 1.243)	1.245 (1.219, 1.271)

the two medication types with increased rates between 23%–27% compared to the other months. Increased relative humidity is associated with a small increased risk of around 1% for both medication types while the uncertainty intervals for temperature generally include the null risk of one. Finally, increasing the percentage of a surgeries patients who are white leads to higher prescription rates by around 1%–5%.

4.3. *Health board inequalities*. A secondary aim of this study is to estimate the inequalities in prescription rates across the 14 regional health boards which is quantified by averaging the process convolution $\phi_{ti}(\mathbf{s}_k)$ from the locally adaptive model (with $m = 8$) over all surgeries and months within each health board. The estimated health board level relative rates for preventer (a) and reliever (b) medications are displayed in Figure 2 where the average rate across Scotland equals

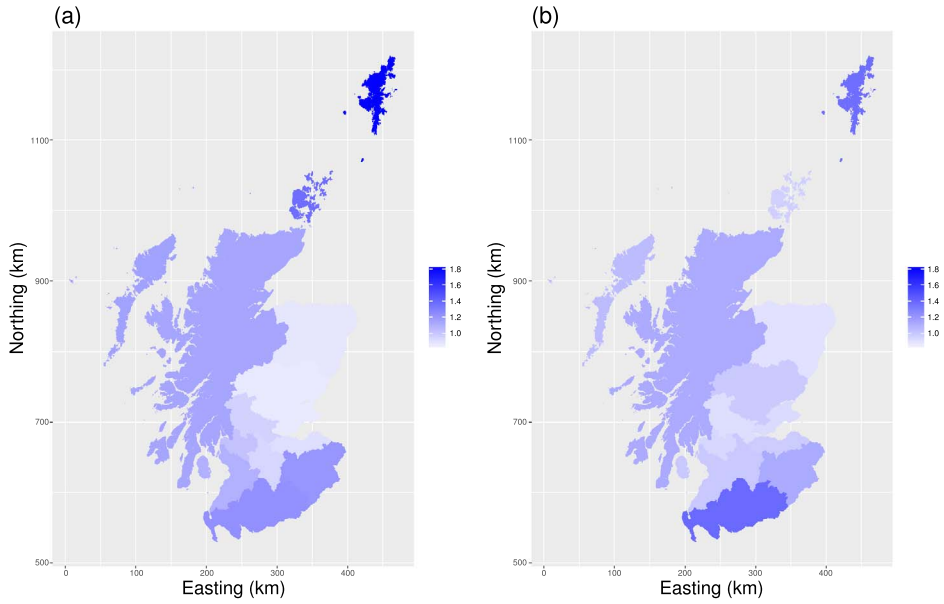

FIG. 2. *The average* (*over time*) *health board level relative rates for*: (a) *preventer*, (b) *and reliever medication after adjusting for the other covariates*.

one. The figure shows that both medication types show similar spatial patterns (Pearson's correlation coefficient is 0.74) with Dumfries and Galloway (far south) and Shetland (far north) being high for both medication types while Grampian in the east is consistently low. The range of estimated relative rates is relatively large with the highest being an 81% increased rate while the lowest is a 25% reduced rate compared to the Scottish average. The other interesting finding is that the highest rates are typically found in the rural health boards without large cities as the health boards containing the four largest cities Glasgow (Greater Glasgow and Clyde), Edinburgh (Lothian), Aberdeen (Grampian) and Dundee (Tayside) all have average or below average relative rates.

4.4. *Correlation structure of the process convolution*. The posterior median between medication correlation from the process convolution $\phi_{ti}(\mathbf{s}_k)$ is relatively strong at 0.71 (from $\frac{\hat{\Sigma}_{12}}{\sqrt{\hat{\Sigma}_{11}\hat{\Sigma}_{22}}}$), while the temporal autocorrelation is very strong with $\hat{\gamma} = 0.98$. The estimated relationship between the spatial autocorrelations and distance apart are presented in Figure 3 for both the kernel (a) and locally adaptive (b) weight models where the figure has been truncated at 10 kilometers to improve the presentation. In each panel of the figure the colours denote the autocorrelations from the other model: (a) locally adaptive weight model and (b) kernel weight model to allow a comparison between the two weight models.

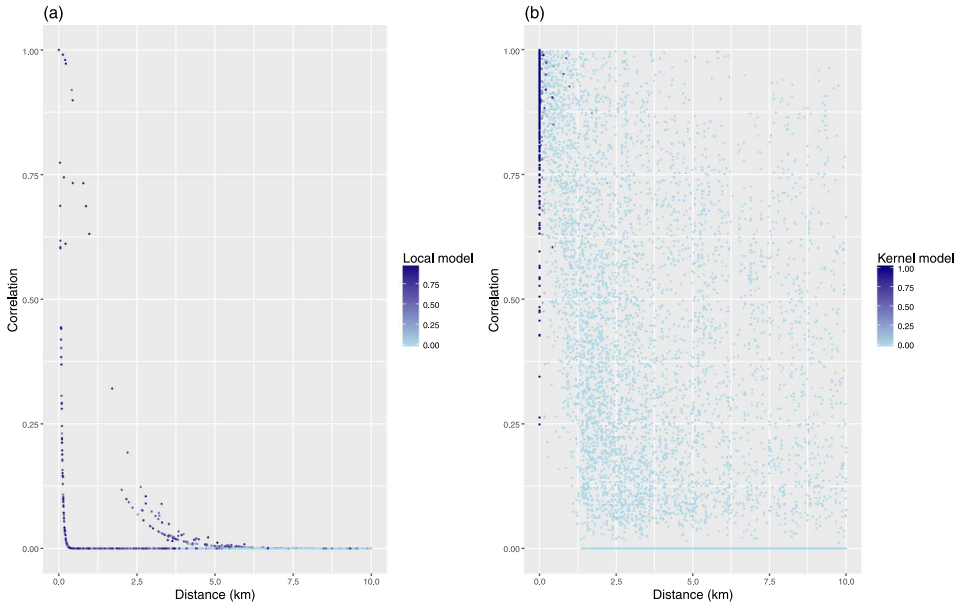

FIG. 3. *The points denote the relationship between distance and correlation for the*: (a) *adaptive kernel weight model and* (b) *locally adaptive weight model. The colours of the points denote the correlations for the same points from the other model*: (a) *locally adaptive weight model and* (b) *kernel weight model*.

Panel (a) shows that the kernel model exhibits different distance-decay relationships for surgeries in urban, small town and rural environments. The shortest range correlations come in urban environments, which is likely to be because surgeries are much closer together, as well as the fact that even short distances between surgeries will likely mean they represent different local communities. The kernel model also produces very few high autocorrelations which is likely because the rigid distance-decay kernels are not realistic, that is, two pairs of surgeries the same distance apart will have very different correlations. In contrast the locally adaptive model in panel (b) shows many more stronger spatial autocorrelations due to its flexible nature which is likely to be the reason for the improved model fit. The shading of the points in panel (a) shows that for the locally adaptive model some pairs of nearby surgeries have very low autocorrelations while other pairs further apart have much higher autocorrelations, a facet that was observed in the exploratory data analysis.

5. Discussion. This paper has proposed a novel bivariate locally adaptive spatiotemporal process-convolution model for estimating the effects of air pollution on respiratory prescribing rates in primary care and has provided data and code to make the research reproducible. It is the first paper to jointly model data on medications that relieve and prevent the symptoms of respiratory disease and is the first

study of its kind in Scotland. We find that both $PM_{2.5}$ and PM_{10} exhibit significant effects on the rates of prescribing for respiratory medication with estimated relative rates from the locally adaptive model ranging between 1.018 and 1.028 for a 2 μgm^{-3} increase in concentrations. For $PM_{2.5}$ this equates to an average of 634 (preventer medication) and 3715 (reliever medication) fewer respiratory prescriptions per month across Scotland if the concentrations fell by 2 μgm^{-3}. These positive effects occur despite Scotland having low particulate matter concentrations with Loader et al. (2016) reporting that in 2015 none of the 76 sites that monitored PM_{10} exceeded the UK Air Quality Strategy limit of an annual mean of 40 μgm^{-3}. What is especially interesting is that the size of the estimated effects are generally similar to those reported for more severe respiratory outcomes using a range of study designs [e.g., Dominici, Samet and Zeger (2000) and Huang, Lee and Scott (2018)], although as a caveat these studies are not directly comparable because of the differing time periods and spatial-temporal scales of the data sets used. Thus this study has provided evidence that current estimates of the health burden attributable to air pollution is likely to be an underestimate because it ignores the effects on ill health treated in primary care.

The methodological novelty of this paper is two-fold, including the use of a tapering function to make the PC more computationally efficient to fit as well as the use of locally adaptive weights. The proposed tapering approach combines covariance tapering with nearest neighbour methods, and Section 4 shows that it reduces the computational burden without having a noticeable impact on the results. The use of locally adaptive weights is also novel, and we have illustrated that for these data it leads to a much better model fit. In this study a pair of GP surgeries close together may not necessarily have correlated prescription rates, a point that has also been observed in the *Wombling* literature for areal unit data [see Lu et al. (2007)]. We note though that data points close together are on average likely to be more similar than those further apart, but that this phenomenon does not hold universally.

The methodology and application presented here suggest two main avenues for methodological extensions. The first is to relax the separability in space, time and between medications of the correlations assumed by the PC which although appropriate for the data analysed here is unlikely to hold in all cases. One possible example would be when jointly modelling medications for different types of diseases with different aetiologies which would likely make the separability assumption less tenable. The second methodological extension would be a joint model for data on different severities of respiratory disease, such as GP prescriptions and hospitalisations, which would allow a direct comparison of their relative associations with air pollution. However, hospitalisation data in Scotland are available at nonoverlapping areal units [see Huang, Lee and Scott (2018)], so the spatial scales (fixed areal units vs. GP practice populations with unknown spatial support) of the two data sets are completely different. Thus to make the results comparable, such a joint model would require novel spatial misalignment methodology to be developed.

Acknowledgment. The data were provided by the Information Services Division of NHS Scotland.

SUPPLEMENTARY MATERIAL

Supplement A: Additional results and data analysis (DOI: 10.1214/18-AOAS1167SUPPA; .pdf). Section 1 contains additional data summaries, while Section 2 presents predictive analysis for the pollution data. Section 3 presents exploratory analysis of the prescription data, while Section 4 presents the reproducibility materials. Section 5 provides theoretical results, while Section 6 presents sensitivity analyses.

Supplement B: Supplementary data and code (DOI: 10.1214/18-AOAS1167SUPPB; .zip). Code and data for applying the model proposed in Section 3 to the GP data.

REFERENCES

BANERJEE, S., GELFAND, A. E., FINLEY, A. O. and SANG, H. (2008). Gaussian predictive process models for large spatial data sets. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 825–848. MR2523906

BHOPAL, R., STEINER, M., CEZARD, G., BANSAL, N., FISCHBACHER, C., SIMPSON, C., DOUGLAS, A. and SHEIKH, A. (2015). Risk of respiratory hospitalization and death, readmission and subsequent mortality: Scottish health and ethnicity linkage study. *Eur. J. Public Health* **25** 769–774.

BLANGIARDO, M., FINAZZI, F. and CAMELETTI, M. (2016). Two-stage Bayesian model to evaluate the effect of air pollution on chronic respiratory diseases using drug prescriptions. *Spatial Spatio-temporal Epidemiol.* **18** 1–12.

CENTRE FOR ENVIRONMENTAL DATA ANALYSIS (2016). HadISDH: Gridded global monthly land surface humidity data version 3.0.0.2016p.

CENTRE FOR ENVIRONMENTAL DATA ANALYSIS (2017). Climatic Research Unit (CRU) Time-Series (TS) version 4.01 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901–Dec. 2016).

CONGDON, P. (2005). *Bayesian Models for Categorical Data*. Wiley, Chichester. MR2191351

DATTA, A., BANERJEE, S., FINLEY, A. O. and GELFAND, A. E. (2016). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *J. Amer. Statist. Assoc.* **111** 800–812. MR3538706

DIBBEN, C. and CLEMENS, T. (2015). Place of work and residential exposure to ambient air pollution and birth outcomes in Scotland, using geographically fine pollution climate mapping estimates. *Environ. Res.* **140** 535–541.

DOMINICI, F., SAMET, J. and ZEGER, S. (2000). Combining evidence on air pollution and daily mortality from the 20 largest US cities: A hierarchical modelling strategy. *J. Roy. Statist. Soc. Ser. A* **163** 263–302.

FURRER, R., GENTON, M. G. and NYCHKA, D. (2006). Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Statist.* **15** 502–523. MR2291261

GELMAN, A., CARLIN, J. B., STERN, H. S. and RUBIN, D. B. (2004). *Bayesian Data Analysis*, 2nd ed. Chapman & Hall/CRC, Boca Raton, FL. MR2027492

HIGDON, D. (1998). A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environ. Ecol. Stat.* **5** 173–190.

HUANG, G., LEE, D. and SCOTT, E. (2018). Multivariate space-time modelling of multiple air pollutants and their health effects accounting for exposure uncertainty. *Stat. Med.*. **37**. 1134–1148.

LEE, D. (2018). Supplement to "A locally adaptive process-convolution model for estimating the health impact of air pollution." DOI:10.1214/18-AOAS1167SUPPA, DOI:10.1214/18-AOAS1167SUPPB.

LEE, D., MUKHOPADHYAY, S., RUSHWORTH, A. and SAHU, S. (2017). A rigorous statistical framework for spatio-temporal pollution prediction and estimation of its long-term impact on health. *Biostatistics* **18** 370–385.

LIECHTY, J. C., LIECHTY, M. W. and MÜLLER, P. (2004). Bayesian correlation estimation. *Biometrika* **91** 1–14. MR2050456

LOADER, A., HECTOR, D., SYKES, D., ROSE, R., TELFER, S. and GRAY, S. (2016). Air Pollution in Scotland 2015. Technical report, Ricardo Energy & Environment.

LU, H., REILLY, C. S., BANERJEE, S. and CARLIN, B. P. (2007). Bayesian areal wombling via adjacency modeling. *Environ. Ecol. Stat.* **14** 433–452. MR2405556

MEYER, S. and HELD, L. (2014). Power-law models for infectious disease spread. *Ann. Appl. Stat.* **8** 1612–1639. MR3271346

RIEBLER, A., HELD, L. and RUE, H. (2012). Estimation and extrapolation of time trends in registry data—Borrowing strength from related populations. *Ann. Appl. Stat.* **6** 304–333. MR2951539

ROYAL COLLEGE OF PHYSICIANS (2016). Every breath we take: The lifelong impact of air pollution. Available at https://www.rcplondon.ac.uk/projects/outputs/every-breath-we-take-lifelong-impact-air-pollution.

WATANABE, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* **11** 3571–3594. MR2756194

SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF GLASGOW
GLASGOW G12 8SQ
UNITED KINGDOM
E-MAIL: Duncan.Lee@glasgow.ac.uk