Estimating the health impact of air pollution in Scotland, and the resulting benefits of reducing concentrations in city centres

Duncan Lee[a,*], Chris Robertson[b], Colin Ramsay[c], Colin Gillespie[d], Gary Napier[a]

[a]*School of Mathematics and Statistics, University of Glasgow*
[b]*Department of Mathematics and Statistics, University of Strathclyde*
[c]*Health Protection Scotland*
[d]*Scottish Environment Protection Agency*

Abstract

Air pollution continues to be a key health issue in Scotland, despite recent improvements in concentrations. The Scottish Government published the Cleaner Air For Scotland strategy in 2015, and will introduce Low Emission Zones (LEZs) in the four major cities (Aberdeen, Dundee, Edinburgh and Glasgow) by 2020. However, there is no epidemiological evidence quantifying the current health impact of air pollution in Scotland, which this paper addresses. Additionally, we estimate the health benefits of reducing concentrations in city centres where most LEZs are located. We focus on cardio-respiratory disease and total non-accidental mortality outcomes, linking them to concentrations of both particulate (PM_{10} and $PM_{2.5}$) and gaseous (NO_2 and NO_x) pollutants. Our two main findings are that: (i) all pollutants exhibit significant associations with respiratory disease but not cardiovascu-

*Corresponding author - Duncan Lee, School of Mathematics and Statistics, University of Glasgow, Glasgow, G12 8SQ

Email address: Duncan.Lee@glasgow.ac.uk (Duncan Lee)

lar disease; and (ii) reducing concentrations in city centres with low resident populations only provides a small health benefit.

Keywords: Air pollution, Cardio-respiratory disease, Epidemiological modelling

---

1. Introduction

Air pollution is the biggest environmental risk to health across the world, with the World Health Organisation (WHO) estimating that 3 million deaths are attributable to it each year (World Health Organisation, 2016). Pollution concentrations around the world often exceed safe levels, with an estimated 90% of the population living in areas where pollutants exceed WHO guideline levels (also World Health Organisation, 2016). The true impact on health is difficult to measure directly, and estimates vary with wide uncertainty intervals. The United Kingdom (UK) Royal College of Physicians estimated that up to 40,000 deaths in the UK could be attributable to air pollution each year (Royal College of Physicians, 2016).

The focus of this study is Scotland, UK, where pollution concentrations are now comparatively low, although there are 39 declared Air Quality Management Areas (AQMA, http://www.scottishairquality.co.uk/laqm/aqma), which either breach or are likely to breach legal pollution limits set by the European Union (EU, European Parliament, 2008). The majority of these breaches are for nitrogen dioxide (NO_2, 27 areas) and / or coarse particulate matter (PM_{10}, 24 areas), with only one for sulphur dioxide (SO_2). The Scottish Government published the Cleaner Air For Scotland (CAFS) strategy (http://www.gov.scot/Resource/0048/00488493.pdf) in 2015,

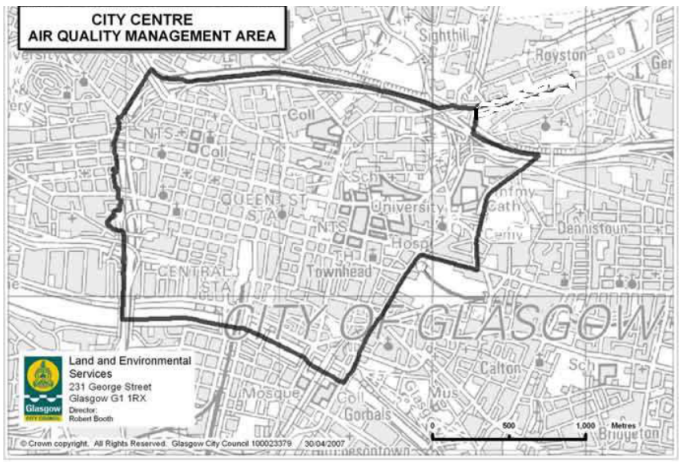

Figure 1: Boundary of Glasgow City Councils Air Quality Management Area, which is the location for the proposed LEZ.

which proposes interventions directed particularly at reducing traffic related pollution. One such intervention is a Low Emission Zone (LEZ), where vehicles that do not meet specified emission standards are banned from, or attract fines for, entering the zone. The first LEZ in Scotland was introduced in the city of Glasgow at the end of 2018 (https://news.gov.scot/news/first-low-emission-zone-for-glasgow), with a phased implementation over 5 years starting with buses that do not meet the EURO 6 emission standard. The other 3 main cities (Aberdeen, Dundee and Edinburgh) are mandated by the Scottish Government to follow suit by the end of 2020. The location for the Glasgow LEZ is the city centre (see Figure 1), bounded by the M8 motorway (west and north), river Clyde (south) and High street (east).

The city centre has been specified for the LEZ because it was identified as the area most likely to exceed EU limit values for NO_2 through the assessment

of air quality data. For example, despite the continual improvements in measured NO_2, the Glasgow Kerbside monitoring station (in the city centre) continually exceeds the EU limit of $40\mu g\ m^{-3}$ for annual mean NO_2, with many of the passive diffusion tube sites within the city centre AQMA also continuing to exceed this limit (see http://www.scottishairquality.co.uk/assets/documents//Glasgow_LAQM_Annual_Progress_Report_2017.pdf). Thus as the Glasgow LEZ was located based on achieving regulatory compliance, the improvement of public health was not the primary driver in deciding the location. Possible public health drivers for an air pollution intervention include the reduction of the overall risk from air pollution, and a reduction in the number of disease cases, the latter being naturally targeted at highly populated and high risk areas.

For the Glasgow LEZ, its beneficial health impact will depend on the size, demographics and underlying health of the population who spend time in the LEZ, as well as on the scale of reduction in pollution concentrations that it achieves. Thus while the city centre has the highest pollution concentrations within the city, it also has a very low resident population and thus may have a limited impact on the majority of Glasgow's population. This preceding argument however does not account for people who travel into the city centre for work or pleasure for large periods of time, which illustrates the complexity of comprehensively evaluating the health impact of an LEZ.

Our aims for this paper are two-fold, with the first being to provide up-to-date policy relevant evidence about the impact of long-term exposure to coarse and fine particulate matter (PM_{10} and $PM_{2.5}$) and oxides of nitrogen (NO_2 and NO_x) on a range of health outcomes to address the gap in the

evidence base about the health impacts of current levels of air pollution concentrations in Scotland. Existing studies include Lee et al. (2009); Lee (2012); Willocks et al. (2012); Dibben and Clemens (2015) and Huang et al. (2015), but are based on relatively old data up to 2011. Our second aim is to use our modelling results to estimate the spatially-varying health benefits of reducing air pollution concentrations in Scotland's cities, specifically in city centres where LEZs are most likely to be located. The data and study region are presented in Section 2, while the proposed statistical methodology is outlined in Section 3. The results of the study are presented in Section 4, while a note of caution about comparing the results here to other studies is presented in Section 5. Finally, the key conclusions are presented in Section 6.

2. Data and study design

The study is based in mainland Scotland for the two-year period 2015-2016, and the study region has been spatially partitioned into $K = 1252$ Intermediate Zones (IZ) that have an average population of around 4,000. The health effects associated with air pollution are estimated from the spatial contrasts in population-level disease incidence and air pollution concentrations across the study region, after adjusting for population demographics and socio-economic deprivation.

2.1. Disease data

The data are counts of the numbers of disease events from the populations living in each IZ in the two-year study period, and we consider the following 5 outcomes: respiratory hospitalisations and mortalities (ICD-10 codes

J00-J99), cardiovascular hospitalisations and mortalities (ICD-10 codes I00-I99), and total non-accidental mortalities. For the hospitalisation outcomes the data relate to the total numbers of events rather than the number of first events, so that an individual who has multiple hospitalisation events within the two-year period will contribute more than one event to the count data. All of these outcomes have been associated with air pollution in the existing literature (see Schwartz et al., 2001; Brook et al., 2004 and Lee et al., 2009), and cardiovascular and respiratory disease are two of Scotland's leading causes of deaths (see http://www.gov.scot/Topics/Statistics/Browse/Health/TrendMortalityRates). These data are summarised in Table 1, where the figure for 0% represents the minimum number of counts recorded for the health outcome in any of the IZs and 100% of the distribution represents the maximum count recorded among all the IZs.

The area level disease counts depend on the size and age-sex structure of the population at risk within each areal unit (IZ), which is accounted for by computing the expected number of disease events in each IZ using indirect standardisation. Specifically, the population living within each IZ is split into strata based on 5-year age bands and sex, and the number of people in each strata is multiplied by national strata specific disease rates, which are then summed over strata to compute the expected count. Letting (Y_k, e_k) respectively denote the observed and expected numbers of disease events in the kth IZ, an exploratory measure of disease risk is the Standardised Morbidity / Mortality Ratio (SMR), which is computed as $\text{SMR}_k = Y_k/e_k$. An SMR of one corresponds to an average risk area, while an SMR of 1.2 corresponds to a 20% increased risk of disease compared to the Scottish

average.

The spatial distribution of the SMR is summarised in Table 1, which shows that the average SMR is close to 1 in all cases, and that generally the mortality outcomes have a wider range of SMR values than the hospitalisation outcomes due to the mortality outcomes having smaller numbers of incidents and hence being a more unstable ratio. The spatial pattern in the SMR for respiratory hospitalisations is displayed in panel A of Figure 2, which shows that the majority of the IZs are in the heavily populated central belt of Scotland containing the two largest cities Glasgow and Edinburgh. A large number of the high SMRs (dark colours) are in the city of Glasgow, which is known to exhibit some of the worst health in the United Kingdom (Walsh et al., 2017). The SMRs for the remaining disease outcomes exhibit similar spatial patterns, with correlations ranging between 0.48 (between cardiovascular and respiratory mortality) and 0.77 (between cardiovascular and total non-accidental mortality).

2.2. Air pollution data

The network of air pollution monitors and diffusion tubes is relatively sparse in Scotland (see `http://www.scottishairquality.co.uk`), and is not sufficient for the small-area scale of this study. Therefore in common with Haining et al. (2010) and Lee et al. (2009) we utilise modelled concentrations instead, specifically annual averages for 2015 and 2016 from the Pollution Climate Mapping (PCM) model (`https://uk-air.defra.gov.uk/data/pcm-data`) developed for the Department for the Environment, Food and Rural Affairs (DEFRA). This model estimates concentrations on a 1km square grid, which are spatially misaligned with the irregularly shaped Inter-

Table 1: Summary of the spatial distribution of the disease and pollution data across the 1252 Intermediate Zones.

Variable	**Percentiles of the distribution**				
	0%	**25%**	**50%**	**75%**	**100%**
Disease incidents (total counts)					
Cardiovascular hospitalisation	26	101	131	166	354
Cardiovascular mortality	2	16	22	30	90
Respiratory hospitalisation	34	108	148	200	530
Respiratory mortality	0	7	11	15	50
Total non-accidental mortality	7	63	84	109	303
Disease risk (SMR)					
Cardiovascular hospitalisation	0.44	0.83	0.98	1.17	2.16
Cardiovascular mortality	0.19	0.80	0.99	1.20	2.76
Respiratory hospitalisation	0.33	0.75	1.00	1.32	2.48
Respiratory mortality	0.00	0.67	0.96	1.31	3.44
Total non-accidental mortality	0.28	0.82	0.99	1.18	2.27
Air pollutants (in $\mu g\ m^{-3}$)					
NO_2	1.3	5.8	9.8	14.0	38.3
NO_x	1.7	7.6	13.3	19.8	74.7
$PM_{2.5}$	3.2	5.6	6.1	6.5	9.1
PM_{10}	5.5	9.0	10.0	10.8	13.9

mediate Zones that the disease data relate to. Such spatial misalignment is often addressed by simple averaging (see Haining et al., 2010), which is the approach adopted here. Specifically, each 1km grid square has an associated centroid (central point), and the estimated pollution concentration for an IZ is the mean of the grid square concentrations whose centroids lie within the IZ. Any IZ that does not contain a grid square centroid is assigned the pollution concentration from the nearest grid square.

In this study we consider concentrations of nitrogen dioxide (NO_2), nitrogen oxides (NO_x), and coarse (PM_{10}) and fine ($PM_{2.5}$) particulate matter, all of which are measured in $\mu g\ m^{-3}$. These pollutants are chosen because they are the ones responsible for all but one of Scotland's air quality management areas. The spatial distribution of $PM_{2.5}$ is displayed in panel B of Figure 2, which shows it is highest in the cities of Glasgow and Edinburgh as well as around the east and south east coasts, the latter due to transboundary pollution from continental Europe and England respectively.

A summary of the spatial distributions of all 4 pollutants is displayed in Table 1, which shows that the 2-year annual average concentrations are generally low. They also exhibit relatively little variation, with standard deviations of 5.5 (NO_2), 8.8 (NO_x), 0.8 ($PM_{2.5}$) and 1.4 (PM_{10}) respectively. Thus presenting the estimated PM_{10}-disease associations as relative risks for a $10\mu g\ m^{-3}$ increase in concentrations, as is done in existing time series studies (see Dominici et al., 2004), would not be sensible, because $10\mu g\ m^{-3}$ does not represent a plausible increase given the data. Therefore in the results we specify relative risks based on a $5\mu g\ m^{-3}$ increase for NO_2 and NO_x, and a $1\mu g\ m^{-3}$ increase for $PM_{2.5}$ and PM_{10}, although we accept this is, as it has

to be, a somewhat arbitrary choice. Finally, the four pollutants are highly correlated spatially, with correlations of: 0.99 between NO_x and NO_2; 0.98 between PM_{10} and $PM_{2.5}$; and between 0.66 and 0.69 for all other pairs of pollutants.

2.3. Confounder data

One of the main factors affecting cardio-respiratory disease incidence is smoking (Hawthorne and Fry, 1978), and therefore areas with higher smoking prevalences are likely to exhibit higher numbers of disease incidents. However smoking prevalence data are unavailable at the IZ scale, but Kleinschmidt et al. (1995) have shown a strong link between smoking rates and socio-economic deprivation. Therefore we use the Scottish Index of Multiple Deprivation (SIMD, http://www.gov.scot/Topics/Statistics/SIMD) in our models as a proxy for smoking. The SIMD is a composite index consisting of deprivation indicators in the domains of access to services, crime, education, employment, health, housing and income, which are weighted and combined to create the final index.

However, as the health domain in this overall index contains similar variables to the disease outcome variables, it cannot be used as a covariate in the models. Therefore in the modelling described in Section 4 we consider the indicators for the 6 individual domains, excluding health, as possible covariates. The crime indicator has a single IZ with a missing value, which is imputed by computing the average value from geographically neighbouring areas (those sharing a common border). Naturally however these six indicators are highly correlated, with the highest correlation being between the income and employment domains (correlation of 0.98), which thus means

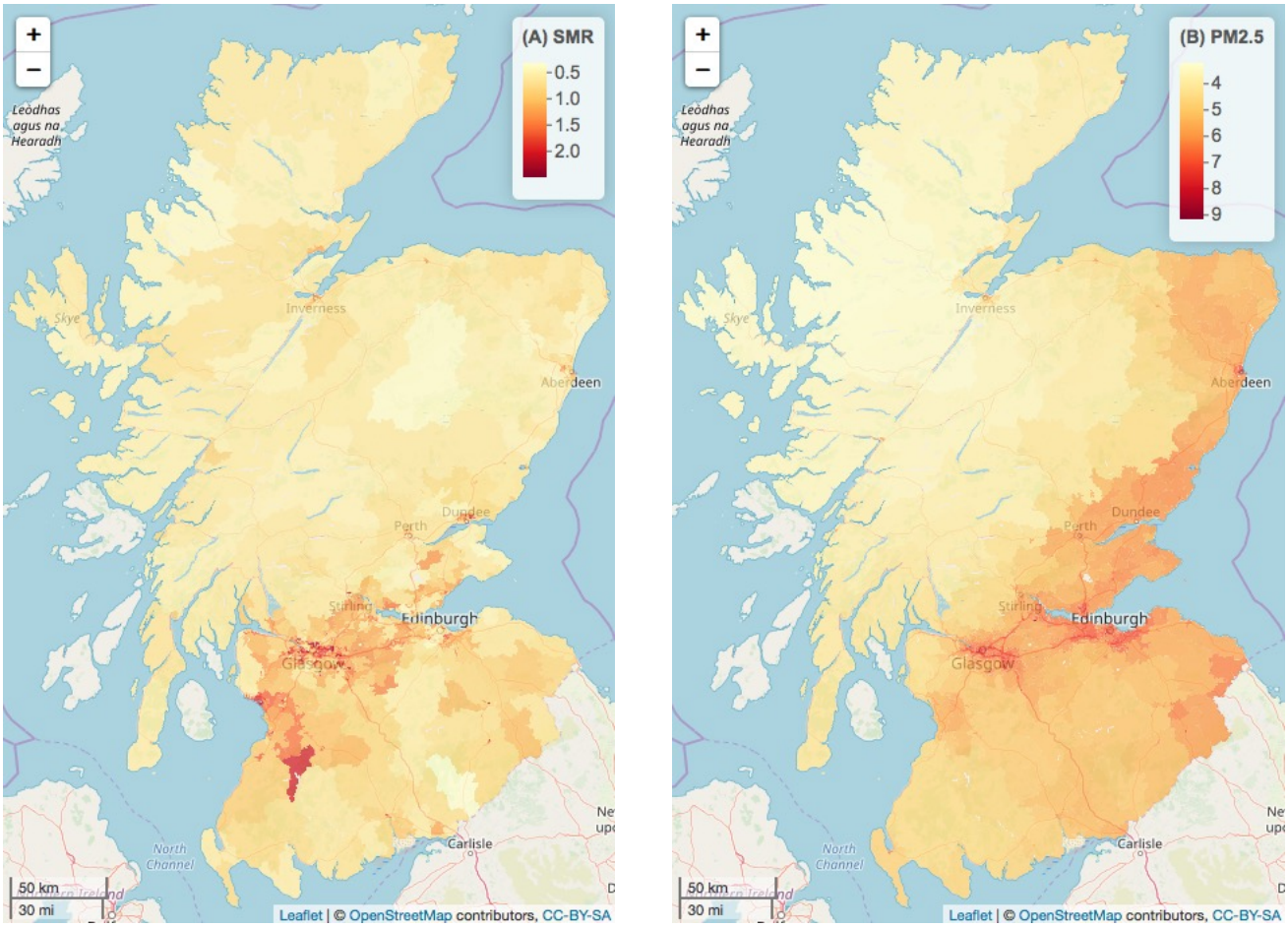

Figure 2: Display of the data. The left panel (A) shows the standardised morbidity ratio for respiratory hospitalisations, while the right panel (B) presents the average concentrations of $PM_{2.5}$.

we do not include them in the same model. Finally, we also have the average number of dwellings per hectare, which is a proxy measure of property density and hence urbanicity.

2.4. Assessment of residual spatial autocorrelation

Here we examine whether the disease outcomes contain residual spatial autocorrelation after covariate adjustment, because this will affect the choice of model that is appropriate for these data. To assess the presence or absence of such correlation, overdispersed quasi-Poisson log-linear models were fitted to each disease outcome separately, where the expected disease counts e_k were included as an offset term. The covariates included in the models were selected from the set described in the previous section, where the selection was based on the significance (at the 5% level) of their association with the disease outcomes and their pairwise correlations. The residuals from these models contained substantial overdispersion, with the estimated overdispersion parameter $\hat{\omega}$ (where $\text{Var}[Y_k] = \omega\mathbb{E}[Y_k]$) ranging between 1.35 and 6.41 across the 5 disease outcomes. The residuals also contained substantial spatial autocorrelation, which was assessed by performing permutation tests based on Moran's I statistic (Moran, 1950). The Moran's I statistics ranged between 0.04 and 0.38 and had p-values less than 0.01 in all cases, which suggests that spatially correlated random effects that also account for overdispersion should be included in the final model.

However, the residual surfaces do not vary smoothly in space, and instead exhibit subregions of spatial smoothness separated by abrupt step changes. This is illustrated in Section 1 of the supplementary material, which displays maps of the residuals from the model applied to the respiratory hospitali-

sations data zoomed in to the cities of Glasgow and Edinburgh. The maps show that while most pairs of spatially neighbouring IZs exhibit similar residual values suggesting spatial autocorrelation, there are numerous examples of large step-changes between spatially neighbouring IZs. This suggests that a globally smooth spatial autocorrelation structure is unlikely to be appropriate for these data, which motivates the use of the locally adaptive spatial smoothing model described in the next section.

3. Methodology

We quantify the impact of air pollution on disease risk using the spatial hierarchical regression model proposed by Lee and Mitchell (2013), because it allows for localised spatial autocorrelation that is present between some pairs of neighbouring areas but absent between other pairs. Inference is undertaken in a Bayesian paradigm using Integrated Nested Laplace Approximations (INLAs, Rue et al., 2009), utilising the `R` package `INLA` (`http://www.r-inla.org`). The overall model is presented in Section 3.1, while the iterative estimation algorithm is presented in Section 3.2. The model is fitted separately for each disease outcome, because this ensures that the cross correlations between the disease outcomes do not affect the estimated pollution-health relationships.

3.1. Overall model

Recall that (Y_k, e_k) respectively denote the observed and expected numbers of disease events in IZ k for $k = 1, \ldots, K$, while x_k denotes the concentration of a single pollutant and $\mathbf{z}_k = (1, z_{k1}, \ldots, z_{kp})$ denotes a vector of p

confounders including an intercept term. The data likelihood model is given by

$$\begin{aligned} Y_k &\sim \text{Poisson}(e_k\theta_k) \quad \text{for } k = 1, \ldots, K \\ \ln(\theta_k) &= \mathbf{z}_k^\top \boldsymbol{\alpha} + x_k\beta + \phi_k, \end{aligned} \tag{1}$$

where θ_k is the risk of disease relative to e_k and can be interpreted on the same scale as the SMR. The regression parameters corresponding to each confounder ($\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$) and the air pollution covariate (β) are assigned independent weakly informative Gaussian prior distributions, with a mean of zero and a variance of 100,000. The remaining term in the linear predictor is a set of random effects $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_K)$, which account for the residual overdispersion and spatial autocorrelation in the disease data not captured by the covariates. The spatial structure of the K IZs is quantified by a non-negative symmetric $K \times K$ neighbourhood matrix $\mathbf{W}$, and here we use the common binary specification where $w_{ki} = 1$ if areas (k, i) share a common border (denoted $k \sim i$) and $w_{ki} = 0$ otherwise (also $w_{kk} = 0\ \forall k$). Then based on $\mathbf{W}$ we model $\boldsymbol{\phi}$ using the conditional autoregressive (CAR) prior proposed by Leroux et al. (2000):

$$\begin{aligned} \phi_k|\boldsymbol{\phi}_{-k}, \mathbf{W}, \tau^2, \rho &\sim \text{N}\left(\frac{\rho\sum_{i=1}^K w_{ki}\phi_i}{\rho\sum_{i=1}^K w_{ki} + 1 - \rho}, \frac{\tau^2}{\rho\sum_{i=1}^K w_{ki} + 1 - \rho}\right) \\ \tau^2 &\sim \text{Inverse-gamma}(1, 0.01) \\ \rho &\sim \text{Uniform}(0, 1), \end{aligned} \tag{2}$$

where $\boldsymbol{\phi}_{-k}$ denotes the vector of random effects except ϕ_k. The prior mean of ϕ_k is a weighted average of the random effects ϕ_i in neighbouring areas

(those for which $w_{ki} = 1$), which thus induces spatial autocorrelation into $\boldsymbol{\phi}$. The strength of this spatial autocorrelation is controlled by ρ, where $\rho = 1$ corresponds to strong spatial autocorrelation and simplifies to the intrinsic CAR model of Besag et al. (1991), while $\rho = 0$ corresponds to independence ($\phi_k \sim \text{N}(0, \tau^2)$). However, model (2) enforces the random effects to exhibit a single global level of spatial smoothness controlled by ρ, which can be seen from its implied partial autocorrelations:

$$\text{Corr}[\phi_k, \phi_i | \boldsymbol{\phi}_{-ki}] = \frac{\rho w_{ki}}{\sqrt{(\rho \sum_{j=1}^{K} w_{kj} + 1 - \rho)(\rho \sum_{j=1}^{K} w_{ij} + 1 - \rho)}}. \quad (3)$$

Thus if ρ is close to one then all pairs of random effects in neighbouring areas where $w_{ki} = 1$ will be partially autocorrelated, whilst if ρ is zero then they will all be independent. However, the exploratory analysis showed that such global spatial smoothness is inappropriate for our data, because some pairs of neighbouring areas have very similar values, suggesting ρ should be close to one, whilst other neighbouring pairs have very different values, suggesting ρ should be close to zero.

Therefore we take the approach of Lee and Mitchell (2013) and estimate each element in the set $\{w_{ki} | k \sim i\}$ as 0 or 1, rather than assuming it is fixed equal to 1. Equation (3) shows that if $w_{ki} = 1$ then (ϕ_k, ϕ_i) will be modelled as partially autocorrelated and hence smoothed over in the modelling, while if $w_{ki} = 0$ then (ϕ_k, ϕ_i) are modelled as conditionally independent and no such spatial smoothing will be enforced. This potentially allows for isolated islands of correlation, where an area is not correlated to any of its neighbours. The major challenge when estimating $\{w_{ki} | k \sim i\}$ is overparameterisation,

because there are $K = 1252$ data points and 3281 neighbourhood elements $\{w_{ki}|k \sim i\}$ to be estimated. Therefore we update $\{w_{ki}|k \sim i\}$ deterministically based on the remaining model parameters $\boldsymbol{\Theta} = (\boldsymbol{\alpha}, \beta, \boldsymbol{\phi}, \tau^2, \rho)$ in an iterative algorithm, rather than assigning each w_{ki} parameter a Bernoulli prior distribution. The algorithm proposed by Lee and Mitchell (2013) and used here is outlined below.

3.2. Iterative estimation algorithm

The algorithm iterates between updating: (i) $\boldsymbol{\Theta}|\mathbf{W}$ and (ii) $\mathbf{W}|\boldsymbol{\Theta}$ until convergence of $\mathbf{W}$ as follows.

Estimation Algorithm

1: Estimate a starting posterior distribution for $\boldsymbol{\Theta}$, by fitting model (1)-(2) based on the assumption that the random effects are independent ($\rho = 0$).

2: Iterate the following two steps for $j = 1, 2, \ldots, j^*$, until one of the two termination conditions for $\mathbf{W}$ outlined in step 3 are met.

- **a:** Estimate $\mathbf{W}^{(j)}$ deterministically based on the current posterior distribution $f(\boldsymbol{\Theta}^{(j-1)}|\mathbf{Y}, \mathbf{W}^{(j-1)})$, by setting $w_{ki}^{(j)} = w_{ik}^{(j)} = 1$ if the marginal 95% posterior credible intervals for $(\phi_k^{(j-1)}, \phi_i^{(j-1)})$ overlap and areas (k, i) share a common border. Otherwise, set $w_{ki}^{(j)} = w_{ik}^{(j)} = 0$.
- **b:** Estimate the posterior distribution $f(\boldsymbol{\Theta}^{(j)}|\mathbf{Y}, \mathbf{W}^{(j)})$ by fitting model (1)-(2) using INLA.

3: After j^* iterations one of the following termination conditions will apply.

Case 1 - The sequence of $\mathbf{W}$ estimates is such that $\mathbf{W}^{(j^*)} = \mathbf{W}^{(j^*+1)}$, which is the estimated hyperparameter matrix $\hat{\mathbf{W}}$.

Case 2 - The sequence of $\mathbf{W}$ estimates forms a cycle of m different states $(\mathbf{W}^{(j^*)}, \mathbf{W}^{(j^*+1)}, \ldots, \mathbf{W}^{(j^*+m-1)}, \mathbf{W}^{(j^*+m)})$, where $\mathbf{W}^{(j^*)} = \mathbf{W}^{(j^*+m)}$. In this case the estimated hyperparameter matrix $\hat{\mathbf{W}}$ is the value from the cycle of m states that has the minimal level of residual spatial autocorrelation, as measured by the absolute value of Moran's I statistic.

When one of the termination conditions has been met $\hat{\mathbf{W}}$ is the estimated spatial structure of the random effects, and $\boldsymbol{\Theta}$ is summarised by the posterior distribution $f(\boldsymbol{\Theta}|\mathbf{Y}, \hat{\mathbf{W}})$.

The algorithm is initialized by assuming the random effects are independent so that initial spatial smoothness constraints are not imposed on the random effects. The update of $\mathbf{W}$ in step 2a assumes that if there is a substantial difference between the current estimates of (ϕ_k, ϕ_i), that is their 95% credible intervals do not overlap, then they should be modelled as conditionally independent, otherwise they are modelled as autocorrelated. In practice, the $\mathbf{W}$ estimates converge to a single value (Case 1) after a small number of iterations in almost all cases, and full details of the algorithm are given in Lee and Mitchell (2013).

4. Results

This section presents the results of the study, including the model building process, pollution-health relative risk estimates, and the impact of air pollution reductions on health.

4.1. Model building

We fit single disease and single pollutant models in this study, resulting in 20 different disease-pollutant combinations. Single disease models ensure that any cross correlations between the disease outcomes do not affect the estimated pollution-health relationships, while single pollutant models are used because of the high collinearity between the four pollutants (pairwise correlations range between 0.66 and 0.99) which hinders reliable joint estimation. To assess the robustness of our results to model choice we fit 2 different spatial autocorrelation models to the data, which are the Poisson log-linear Leroux CAR model ((1) and (2)), and the Poisson log-linear locally adaptive CAR model ((1) and (2) with the estimation of $\mathbf{W}$ as described in Section 3.2).

Each disease outcome is modelled by the expected numbers of disease events as an offset, one of the four pollutants, and a subset of the confounders outlined in Section 2, the latter including the dwellings per hectare variable and the 6 domain specific indicators of the SIMD. The main challenge with confounder selection is collinearity, because the education, employment and income domains all have high pairwise correlations above 0.86. Fitting models with each of these variables separately shows that the income domain variable describes the most variation in the data, and thus is the one re-

tained with the other two being discarded. The remaining confounders do not exhibit this collinearity problem, and as they are all significantly related to most of the disease outcomes, they are retained in all models for consistency. Therefore, the set of confounders included in each model are the access to services, crime, housing and income domains of the SIMD, as well as the dwellings per hectare variable.

The overall fits to the data of each model are presented in Table 2, which displays their Watanabe Akaike Information Criterion (WAIC, Watanabe, 2010) value and the estimated effective number of independent parameters ($p.w$). The results presented relate to when $PM_{2.5}$ was the pollutant included in the model, but the results for the other pollutants are almost identical and are not shown for brevity. The locally adaptive CAR model fits the two hospitalisation outcomes and total non-accidental mortality outcome better than the Leroux CAR model, with reductions in WAIC of 135 (cardiovascular), 209 (respiratory) and 22 (mortality) respectively. These improvements in model fit are achieved despite the locally adaptive model having a smaller effective number of independent parameters than the Leroux model. This phenomenon occurs because the random effects from the Leroux CAR model are globally spatially smooth, which hence forces smoothness between residual risks in geographically neighbouring IZs, even if those residual risks are very different. This inflates the random effects variance τ^2 because the residual risks are not spatially smooth, which results in a greater number of effective parameters. In contrast, the locally adaptive model does not smooth residual risks in geographically adjacent IZs where those residual risks are very different, because it sets the corresponding $w_{ki} = w_{ik}$ elements equal to zero

Table 2: Watanabe-Akaike Information Criterion (WAIC) and the effective number of independent parameters from the Leroux and locally adaptive CAR models. For the latter the number of $\{w_{ki}\}$ elements estimated as zero is also presented.

Disease outcome	**Model**	**WAIC**	**p.w**	**Number of $\{w_{ki}\}$ set to zero**
Cardiovascular hospitalisations	Leroux	10,506	564	-
	Adaptive	10,371	495	115 (3.5%)
Cardiovascular mortality	Leroux	7,753	276	-
	Adaptive	7,753	275	2 (0.1%)
Respiratory hospitalisations	Leroux	10,706	616	-
	Adaptive	10,497	522	386 (11.8%)
Respiratory mortality	Leroux	6,874	244	-
	Adaptive	6,871	251	0 (0%)
Total non-accidental mortality	Leroux	9,782	528	-
	Adaptive	9,760	489	150 (4.6%)

and thus does not assume any partial autocorrelations between the random effects in those IZs. To illustrate the locations of these step changes in the random effects surface, the locations of the borders for which $w_{ki} = w_{ik} = 0$ are displayed for Edinburgh and Glasgow in Section 5 of the supplementary material.

The largest number of $\{w_{ki}\}$ elements estimated as zero is 386 (11.8%) for respiratory hospitalisations, while 115 (3.5%) and 150 (4.6%) were set to zero for cardiovascular hospitalisations and total non-accidental mortality respectively. In contrast, the two CAR models exhibit the same overall fit for the other two disease outcomes, which occurs because the locally adaptive model hardly estimates any $w_{ki} = 0$ and hence it simplifies to the Leroux CAR model.

4.2. Pollution-health effects

The effects of each pollutant on each disease outcome estimated from the locally adaptive CAR model are presented in Table 3, while the corresponding effects for the non-pollutant covariates are presented in Section 3 of the supplementary material accompanying this paper. For completeness, the pollution-disease effects estimated from the model with the Leroux CAR prior are displayed in Section 4 of the supplementary material, and show little change to those presented here, suggesting our results are robust to the choice of spatial autocorrelation model. Table 3 displays relative risks and 95% credible intervals for a $5\mu g\ m^{-3}$ increase in NO_2 and NO_x and a $1\mu g\ m^{-3}$ increase in $PM_{2.5}$ and PM_{10}, because as discussed in Section 2.2, these are realistic increases for each of the pollutants.

Table 3 shows that in this study air pollution only has a significant association with respiratory disease in our data. This is shown prominently for respiratory hospitalisations, where all four pollutants exhibit significant associations. For respiratory mortality the estimated associations are largely similar in size, and the lack of significance at the traditional 5% level (except for $PM_{2.5}$) is because of the much wider credible intervals for this disease outcome, resulting from the much lower numbers of disease counts (less data) compared with respiratory hospitalisations. The effect sizes for respiratory hospitalisations range between a 1.4% and a 5.8% increased risk for the given pollutant increases, although given the differing levels of spatial variation in the pollutants these risks are not directly comparable. Cardiovascular disease and total non-accidental mortality appear to have no relationship with any of the four pollutants, because all 12 of the 95% credible intervals contain

Table 3: Estimated relative risks and 95% credible intervals for the pollution-disease effects from the model with the locally adaptive CAR prior. The results for NO_2 and NO_x relate to a $5\mu g\ m^{-3}$ increase whilst those for $PM_{2.5}$ and PM_{10} relate to a $1\mu g\ m^{-3}$ increase. The significant associations are shown in bold.

Disease outcome	**Pollutant**			
	NO_2	**NO_x**	**$PM_{2.5}$**	**PM_{10}**
Cardiovascular hospitalisations	1.012	1.006	1.018	1.006
	(0.994, 1.030)	(0.995, 1.016)	(0.997, 1.040)	(0.995, 1.017)
Cardiovascular mortality	0.988	0.993	0.995	0.997
	(0.970, 1.006)	(0.982, 1.005)	(0.994, 1.016)	(0.987, 1.008)
Respiratory hospitalisations	**1.028**	**1.014**	**1.058**	**1.023**
	(1.008, 1.048)	**(1.002, 1.025)**	**(1.034, 1.083)**	**(1.011, 1.035)**
Respiratory mortality	1.032	1.017	**1.045**	1.014
	(0.997, 1.067)	(0.996, 1.038)	**(1.002, 1.090)**	(0.992, 1.035)
Total non-accidental mortality	1.003	1.001	1.012	1.005
	(0.986, 1.020)	(0.990, 1.011)	(0.992, 1.033)	(0.995, 1.016)

the null risk of one, and the estimated risks are mostly very close to one.

4.3. Estimating the health impact of pollution reductions

We now use the modelling results to quantify the health impact of reducing air pollution concentrations in each IZ within the four main Scottish cities, Aberdeen, Dundee, Edinburgh and Glasgow, which will illustrate the potential health impact of the planned LEZs. We do this by computing the expected reduction in the numbers of disease cases (hospital admissions or mortalities) in each IZ over 2015-2016 if average concentrations over that two-year period had reduced by $\omega\mu g\ m^{-3}$. We undertake this analysis for each pollutant and disease outcome separately because we have implemented

single pollutant and single disease models, and note that these estimated reductions should not be summed over pollutants or diseases as they are not independent. From equation (1) the estimated reduction in the expected number of disease events for the kth IZ, $\mathbb{E}[Y_k]$, if pollutant x_k reduced by $\omega\mu g\ m^{-3}$ is given by:

$$\begin{aligned}
\textbf{Reduction}_k &= \mathbb{E}[Y_k|x_k] - \mathbb{E}[Y_k|x_k - \omega] \qquad (4)\\
&= e_k \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}} + x_k\hat{\beta} + \hat{\phi}_k) - e_k \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}} + (x_k - \omega)\hat{\beta} + \hat{\phi}_k)\\
&= e_k \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}} + x_k\hat{\beta} + \hat{\phi}_k)[1 - \exp(-\omega\hat{\beta})]\\
&= e_k\hat{\theta}_k[1 - \exp(-\omega\hat{\beta})].
\end{aligned}$$

This reduction depends on the estimated air pollution and health effect $\hat{\beta}$, the pollution reduction $\omega\mu g\ m^{-3}$, the underlying size and demographics of the population at risk via e_k, and the estimated level of disease risk via $\hat{\theta}_k = \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}} + x_k\hat{\beta} + \hat{\phi}_k)$. To understand the range of reductions that might be observed, Table 4 displays the estimated total reductions across the four cities resulting from the following pollutant reductions: NO_2 / NO_x - $2\mu g\ m^{-3}$, $5\mu g\ m^{-3}$ and $10\mu g\ m^{-3}$; and $PM_{2.5}$ / PM_{10} - $0.5\mu g\ m^{-3}$, $1\mu g\ m^{-3}$ and $3\mu g\ m^{-3}$. Here each city is defined by its local authority region, and the pollution-disease combinations listed in the table relate to the significant associations from Table 3.

The table shows that the estimated reductions in disease cases scales with the chosen pollution reductions, as for example increasing the reduction of $PM_{2.5}$ from $0.5\mu g\ m^{-3}$ to $1\mu g\ m^{-3}$ in Edinburgh results in around 400 and 800 fewer respiratory hospitalisations respectively. The biggest reductions

are in Glasgow because it has the largest population in Scotland, with an estimated reduction of 1,576 fewer admissions to hospital over the two year study period (an average of 788 per year) due to respiratory disease if $PM_{2.5}$ reduced by $1\mu g\ m^{-3}$. In contrast, Dundee, the smallest of the four cities, had an estimated reduction in admissions of 352 over the two-year period (on average 176 per year) for the same $1\mu g\ m^{-3}$ decrease in concentrations. Finally, as mortalities are much rarer than hospital admissions, the estimated reductions in respiratory mortalities are much smaller than the corresponding reductions for respiratory hospitalisations.

Equation (4) shows that the health impact of a fixed $\omega\mu g\ m^{-3}$ reduction in a pollutant will vary by IZ, and thus where those reductions are highest would be where pollution reduction policies, such as an LEZ, would have the largest public health benefit. The left column of Figure 3 illustrates this, by displaying, for Edinburgh (top left) and Glasgow (top right), the estimated reductions in respiratory hospitalisations in each IZ between 2015-2016 that would have occurred if NO_2 concentrations had been reduced by $5\mu g\ m^{-3}$. The right column of the figure presents the estimated NO_2 concentrations for the two cities, allowing us to spatially compare the locations with the highest concentrations and the highest health impacts of reducing concentrations. We have chosen to display the results for NO_2 because traffic related interventions, such as the Glasgow LEZ, are designed to reduce this pollutant more than particulates. However, the same general pattern is observed for all 4 pollutants.

The figure shows the same fundamental message for both cities, namely that reducing NO_2 concentrations in a city centre where concentrations are

Table 4: Estimated reductions in the expected numbers of disease events in 2015-2016 in Aberdeen, Dundee, Edinburgh and Glasgow if a pollutant decreased by $\omega \mu g\ m^{-3}$. The values in brackets relate to the region containing the Glasgow LEZ.

Disease / Pollutant	**City**			
	Aberdeen	**Dundee**	**Edinburgh**	**Glasgow** (LEZ)
Respiratory hospitalisations				
NO_2 - $\omega = 2$	71	71	161	316 (5)
NO_2 - $\omega = 5$	176	175	398	784 (13)
NO_2 - $\omega = 10$	347	345	785	1547 (26)
NO_x - $\omega = 2$	35	35	80	158 (3)
NO_x - $\omega = 5$	88	88	199	393 (7)
NO_x - $\omega = 10$	175	175	395	781 (14)
$PM_{2.5}$ - $\omega = 0.5$	179	178	405	799 (14)
$PM_{2.5}$ - $\omega = 1$	354	352	798	1,576 (27)
$PM_{2.5}$ - $\omega = 3$	1005	999	2265	4474 (77)
PM_{10} - $\omega = 0.5$	73	73	165	325 (6)
PM_{10} - $\omega = 1$	144	143	325	641 (11)
PM_{10} - $\omega = 3$	409	406	923	1820 (31)
Respiratory mortality				
$PM_{2.5}$ - $\omega = 0.5$	12	10	22	39 (1)
$PM_{2.5}$ - $\omega = 1$	23	19	44	78 (1)
$PM_{2.5}$ - $\omega = 3$	66	55	126	224 (3)

highest and hence where LEZs are typically located, will likely have a relatively low public health impact in terms of the number of cases. This is because city centres have comparatively low resident populations at risk due to being mainly commercial centres, resulting in smaller estimated reductions in the numbers of disease events. The same observation is true for Aberdeen and Dundee, and the results are displayed in Section 4 of the supplementary material.

The location for the Glasgow LEZ is highlighted by the blue line in Figure 3, and the bottom right panel shows it has some of the highest NO_2 concentrations in the city. However, the bottom left panel shows that the health impact from reducing the concentrations in the LEZ will likely be small, as the three IZs that make up the LEZ combined have a total reduction of 13 hospital admissions over 2015-2016 (on average between 6 and 7 a year), less than 2% of the estimated reduction for the whole city. The corresponding reductions in estimated disease events for the LEZ for the other pollutants, health outcomes, and sizes of pollutant reduction are shown in brackets in Table 4, and again show very small numbers compared with the city of Glasgow as a whole.

5. Comparability of epidemiological air pollution studies

Large numbers of epidemiological air pollution studies have been conducted across the world, which has led researchers and policy-makers to directly compare the results from multiple studies. However this is problematic, because the estimated effect sizes will depend on both the strength of the pollution-health association and the amount of variation in pollution

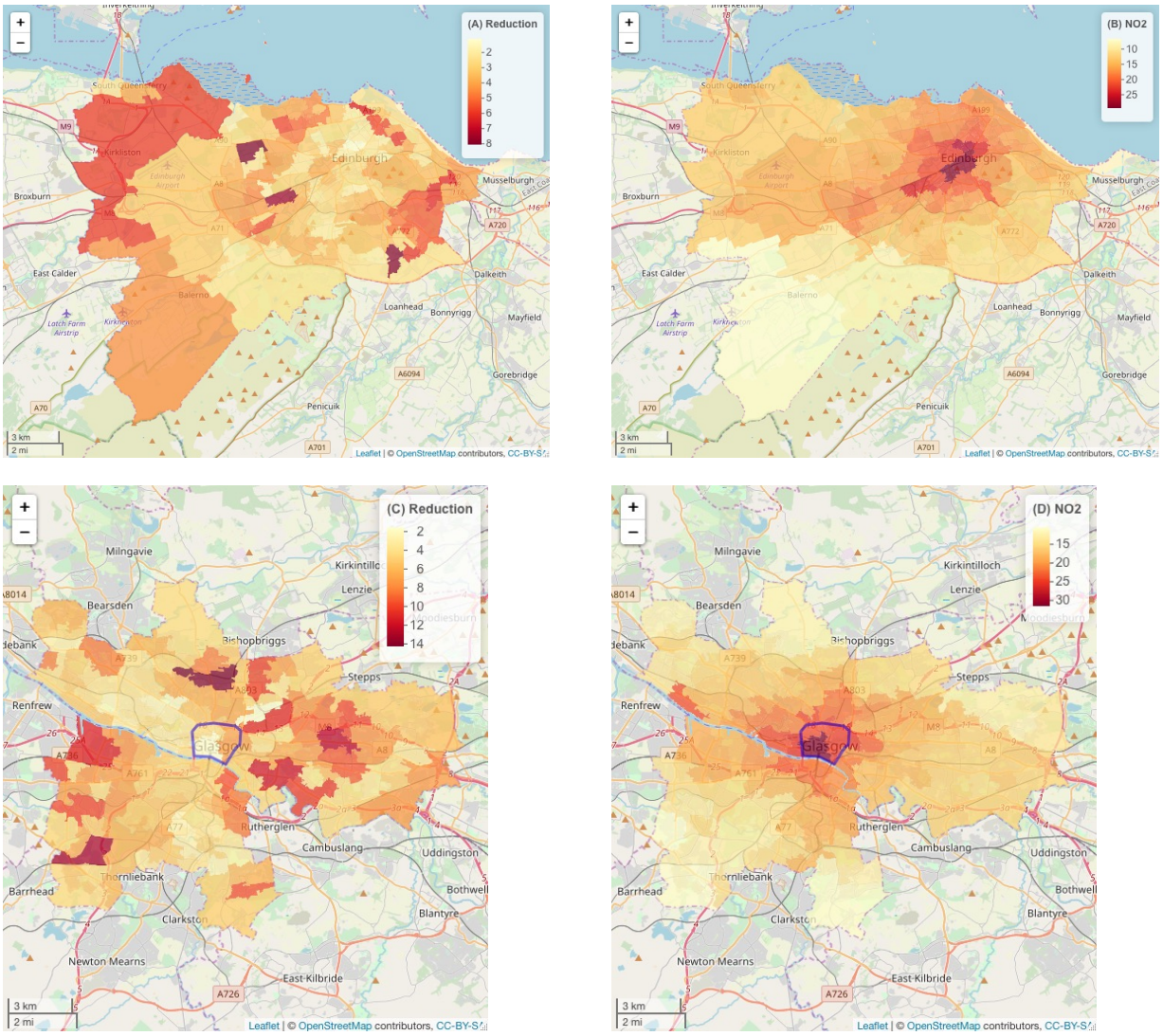

Figure 3: Maps of the estimated reductions in respiratory hospitalisations in each IZ due to a $5\mu g\ m^{-3}$ reduction in NO_2 concentrations (left), and the average NO_2 concentrations (right). The top row refers to Edinburgh and the bottom row refers to Glasgow. The blue line denotes the boundary of the proposed Glasgow LEZ.

concentrations across the study region. To see this note that from equation (1) the risk for area k is given by

$$\hat{\theta}_k = \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}} + x_k\hat{\beta} + \hat{\phi}_k), \tag{5}$$

where the exposure x_k has mean $\bar{x} = \frac{1}{K}\sum_{k=1}^{K} x_k$ and variance $\sigma_x^2 = \frac{1}{n-1}\sum_{k=1}^{K}(x_k - \bar{x})^2$. Now consider a linearly scaled exposure $v_k = (1+\psi)x_k - \psi\bar{x}$, where it is straightforward to show that they have the same mean (i.e. $\bar{v} = \bar{x}$) and the variances are related by $\sigma_v^2 = (\psi+1)^2\sigma_x^2$. Then replacing x_k by v_k in equation (1) yields:

$$\begin{aligned}
\hat{\theta}_k &= \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}}^* + v_k\hat{\beta}^* + \hat{\phi}_k^*) \\
&= \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}}^* + [(1+\psi)x_k - \psi\bar{x}]\hat{\beta}^* + \hat{\phi}_k^*) \\
&= \exp(\mathbf{z}_k^\top \hat{\boldsymbol{\alpha}}^* + x_k(1+\psi)\hat{\beta}^* - \psi\bar{x}\hat{\beta}^* + \hat{\phi}_k^*).
\end{aligned} \tag{6}$$

Comparing (5) and (6) shows that the coefficients for the scaled and unscaled exposures (v_k, x_k) are related by $\hat{\beta}^* = \frac{\hat{\beta}}{1+\psi}$. Therefore, comparing estimated effect sizes between studies with different levels of exposure variation is not appropriate, because the level of exposure variation affects the estimated regression coefficient. This explains the large estimated effect sizes for $PM_{2.5}$ on respiratory disease outcomes observed in this study, because the level of variation in the 2-year average $PM_{2.5}$ concentrations across Scotland is very low (the standard deviation is only $0.81\mu g\ m^{-3}$).

6. Discussion

This paper has presented a new study of the health impact of long-term exposure to air pollution in Scotland using a spatial small-area design, and has used the results to quantify the likely health impact of air pollution reduction interventions such as Low Emission Zones. Our first main finding is that the four pollutants considered here exhibit associations with respiratory disease (hospitalisations and mortality), even though for mortality three of the relative risks are not significant at the 5% level as a result of small numbers of deaths leading to wide credible intervals. In contrast, no significant associations were observed for cardiovascular disease or total non-accidental mortality, with all relative risks being non-significant and close to one in magnitude (ranging between 0.988 and 1.018). No significant associations between cardiovascular disease and air pollution were also found by Willocks et al. (2012) in Scotland and Carey et al. (2013) and Dehbi et al. (2017) in Great Britain using time series and cohort methodologies respectively, which means our findings are consistent with previous studies on British populations.

Our second main finding is that focusing an air pollution reduction intervention, such as an LEZ, on a city centre where concentrations are highest is likely to have a relatively small positive health impact at the national level, because these areas are largely commercial and hence have small resident populations. Even though these areas will routinely see large numbers of people visiting for both shopping and working, their time spent in the area, especially outdoors, will likely be relatively short. The evidence presented here therefore suggests that the LEZ planned for Glasgow may have a rela-

tively small positive net health benefit. We note however that our study has not evaluated the effect of the Glasgow LEZ directly, because the pollution reductions from the LEZ are not known as it will not be fully operational until the end of 2022. However, other studies have directly evaluated the impact of LEZs across Europe, including studies in Amsterdam (Panteliadis et al., 2014) and Munich (Fensterer et al., 2014) where the LEZ appeared to reduce concentrations, and in the UK (London and Birmingham, Jones et al., 2012; Wood et al., 2015) where it did not appear to reduce concentrations. A thorough review of LEZs is beyond the scope of this paper, and the reader is referred to Holman et al. (2015) and AIRUSE (2016).

The choice of where one should locate an air pollution intervention, such as an LEZ, depends on the ultimate goal. If the main aim is to reduce the number of preventable disease cases attributable to air pollution, then Figure 3 suggests that an intervention should be targeted at areas that have both relatively high pollutant concentrations and a relatively large and more vulnerable population. In contrast, if compliance with air quality limits is the key requirement, such as reducing pollution concentrations below European Union limits (European Parliament, 2008), then interventions need only be targeted at areas with the highest concentrations that are not in compliance. Finally, if the aim is an overall reduction in the risk of air pollution, then more geographically wide-reaching air pollution reduction policies are needed.

The main limitation of our work (shared by all other epidemiological air pollution studies) in respect of attempting to predict the potential health impacts of LEZs is the use of ambient residential concentrations as a proxy for personal exposures, which ignores peoples movements such as daily commut-

ing patterns. In future work we will combine the methodology developed here with population movement models, to identify the possible health impacts of reducing air pollution in city centres on personal exposures. Additionally, we will consider the impact of an LEZ on air pollution concentrations in the rest of the city, which are likely to occur because an LEZ will require cleaner buses that will service routes that travel out-with the LEZ area.

A second limitation with this study is that the pollution data are assumed to be true and measured without error, where as in fact they come from the atmospheric PCM model and thus are subject to error and uncertainty. Numerous solutions have been proposed to allow for pollution uncertainty in disease models, and a recent example using fusion modelling (Berrocal et al., 2010) is provided by Blangiardo et al. (2016). A further limitation is the ecological nature of this small-area study, which in common with time series studies (e.g. Dominici et al., 2004), uses population-level disease summaries rather than individual-level data. This means that only group level associations rather than individual-level cause and effect can be estimated, which provides a weaker evidence base. However, individual-level disease data are not available for confidentiality purposes, and population-level small-area studies are commonplace and are critiqued by Wakefield (2007).

Competing interests

All authors declare that they have no competing interest.

Authors contributions

Colin R framed the Secure Challenge research question; DL, GN, and Chris R undertook the statistical analysis; DL wrote the paper; Chris R and Colin R provided the data, while Colin R and CG provided the policy insight on the Glasgow LEZ that helped to shape this work. All authors provided suggestions and edits to improve the paper, and have read and approved the final manuscript.

Acknowledgements

The disease data used in this paper were provided by Health Protection Scotland, while the shapefiles were provided by the Scottish Government. The work was funded by a pilot project grant from the United Kingdom Engineering and Physical Sciences Research Council funded SECURE network grant number EP/M008347/1.

Bibliography

References

AIRUSE, 2016. Low emissions zones in northern and central europe. URL http://airuse.eu/wp-content/uploads/2013/11/R16_AIRUSE-Low-Emission-Zones-CNE.pdf

Berrocal, V. J., Gelfand, A. E., Holland, D. M., 2010. A spatio-temporal downscaler for output from numerical models. Journal of Agricultural, Biological, and Environmental Statistics 15, 176–197.

585 Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration with two
586 applications in spatial statistics. Annals of the Institute of Statistics and
587 Mathematics 43, 1–59.

588 Blangiardo, M., Finazzi, F., Cameletti, M., 2016. Two-stage bayesian model
589 to evaluate the effect of air pollution on chronic respiratory diseases using
590 drug prescriptions. Spatial and Spatio-temporal Epidemiology 18, 1 – 12.

591 Brook, R., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M.,
592 Luepker, R., Mittleman, M., Samet, J., Smith, S., Tager, I., 2004. Air
593 pollution and cardiovascular disease: A statement for healthcare profes-
594 sionals from the expert panel on population and prevention science of the
595 american heart association. Circulaton 109, 2655–2671.

596 Carey, I. M., Atkinson, R. W., Kent, A. J., van Staa, T., Cook, D. G.,
597 Anderson, H. R., 2013. Mortality associations with long-term exposure to
598 outdoor air pollution in a national english cohort. American Journal of
599 Respiratory and Critical Care Medicine 187, 1226–1233.

600 Dehbi, H.-M., Blangiardo, M., Gulliver, J., Fecht, D., de Hoogh, K., Al-
601 Kanaani, Z., Tillin, T., Hardy, R., Chaturvedi, N., Hansell, A. L., 2017.
602 Air pollution and cardiovascular mortality with over 25years follow-up: A
603 combined analysis of two british cohorts. Environment International 99,
604 275 – 281.

605 Dibben, C., Clemens, T., 2015. Place of work and residential exposure to
606 ambient air pollution and birth outcomes in Scotland, using geographically

fine pollution climate mapping estimates. Environmental Research 140, 535–541.

Dominici, F., McDermott, A., Hastie, T., 2004. Improved semiparametric time series models of air pollution and mortality. Journal of the American Statistical Association 99, 938–948.

European Parliament, 2008. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official Journal of the European Union.

Fensterer, V., Kchenhoff, H., Maier, V., Wichmann, H.-E., Breitner, S., Peters, A., Gu, J., Cyrys, J., 2014. Evaluation of the impact of low emission zone and heavy traffic ban in munich (germany) on the reduction of pm10 in ambient air. International Journal of Environmental Research and Public Health 11, 5094–5112.

Haining, R., Li, G., Maheswaran, R., Blangiardo, M., Law, J., Best, N., Richardson, S., 2010. Inference from ecological models: estimating the relative risk of stroke from air pollution exposure using small area data. Spatial Spatio-temporal Epidemiology 1, 123–131.

Hawthorne, V. M., Fry, J. S., 1978. Smoking and health: the association between smoking behaviour, total mortality, and cardiorespiratory disease in west central scotland. Journal of Epidemiology & Community Health 32, 260–266.

Holman, C., Harrison, R., Querol, X., 2015. Review of the efficacy of low

emission zones to improve urban air quality in european cities. Atmospheric Environment 111, 161 – 169.

Huang, G., Lee, D., Scott, E., 2015. An integrated bayesian model for estimating the long-term health effects of air pollution by fusing modelled and measured pollution data: A case study of nitrogen dioxide concentrations in scotland. Spatial and Spatio-temporal Epidemiology 14-15, 63–74.

Jones, A. M., Harrison, R. M., Barratt, B., Fuller, G., 2012. A large reduction in airborne particle number concentrations at the time of the introduction of sulphur free diesel and the london low emission zone. Atmospheric Environment 50, 129 – 138.

Kleinschmidt, I., Hills, M., Elliott, P., 1995. Smoking behaviour can be predicted by neighbourhood deprivation measures. Journal of Epidemiology and Community Health 49, S71–S77.

Lee, D., 2012. Using spline models to estimate the varying health risks from air pollution across scotland. Statistics in Medicine 31, 3366–3378.

Lee, D., Ferguson, C., Mitchell, R., 2009. Air pollution and health in Scotland: a multicity study. Biostatistics 10, 409–423.

Lee, D., Mitchell, R., 2013. Locally adaptive spatial smoothing using conditional autoregressive models. Journal of the Royal Statistical Society Stries C 62, 593–608.

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, Ch. Statistical Models in Epidemiology, the Environment and Clinical Trials, Halloran, M and Berry, D (eds), pp. 135–178.

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

Panteliadis, P., Strak, M., Hoek, G., Weijers, E., van der Zee, S., Dijkema, M., 2014. Implementation of a low emission zone and evaluation of effects on air quality by long-term monitoring. Atmospheric Environment 86, 113 – 119.

Royal College of Physicians, 2016. Every breath we take: The lifelong impact of air pollution. https://www.rcplondon.ac.uk/projects/outputs/every-breath-we-take-lifelong-impact-air-pollution.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion). Journal of the Royal Statistical Society Series B 71, 319–392.

Schwartz, J., Ballester, F., Saez, M., Perez-Hoyos, P., Bellido, J., Cambra, K., Arribas, F., Canada, A., Jose Perez-Boillos, M., Sunyer, J., 2001. The Concentration-Response Relation between Air Pollution and Daily Deaths. Environmental Health Perpespectives 109, 1001–1006.

Wakefield, J., 2007. Disease mapping and spatial regression with count data. Biostatistics 8, 158–183.

Walsh, D., McCartney, G., Collins, C., Taulbut, M., Batty, G., 2017. History,

politics and vulnerability: explaining excess mortality in Scotland and Glasgow. Public Health 151, 1–12.

Watanabe, S., 2010. Asymptotic equivalence of the Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research 11, 3571–3594.

Willocks, L., Bhaskar, A., Ramsay, C., Lee, D., Brewster, D., Fischbacher, C., Chalmers, J., Morris, G., Scott, M., 2012. Cardiovascular disease and air pollution in scotland: no association or insufficient data and study design? BMC Public Health 12, 227.

Wood, H. E., Marlin, N., Mudway, I. S., Bremner, S. A., Cross, L., Dundas, I., Grieve, A., Grigg, J., Jamaludin, J. B., Kelly, F. J., Lee, T., Sheikh, A., Walton, R., Griffiths, C. J., 2015. Effects of air pollution and the introduction of the london low emission zone on the prevalence of respiratory and allergic symptoms in schoolchildren in east london: A sequential cross-sectional study. PLOS ONE 10, 1–12.
URL https://doi.org/10.1371/journal.pone.0109121

World Health Organisation, 2016. Ambient air pollution: A global assessment of exposure and burden of disease. http://apps.who.int/iris/bitstream/10665/250141/1/9789241511353-eng.pdf.