

Predicting daily concentrations of nitrogen dioxide, particulate matter and ozone at fine spatial scale in Great Britain

Weiye Wang^{a,*}, Daniela Fecht^a, Sean Beevers^a, John Gulliver^b

^a MRC Centre for Environment and Health, School of Public Health, Imperial College London, Norfolk Place, W2 1PG, London, UK

^b Centre for Environmental Health and Sustainability, George Davies Centre, University of Leicester, University Road, Leicester, LE1 7RH, UK

ARTICLE INFO

Keywords:

Air pollution
Exposure assessment
Spatio-temporal model
Land use regression
Geographic information system
Epidemiology

ABSTRACT

Short-term exposure studies have often relied on time-series of air pollution measurements from monitoring sites. However, this approach does not capture short-term changes in spatial contrasts in air pollution. To address this, models representing both the spatial and temporal variability in air pollution have emerged in recent years. Here, we modelled daily average concentrations of nitrogen dioxide (NO₂), particulate matter (PM_{2.5} and PM₁₀) and ozone (O₃) on a 25 m grid for Great Britain from 2011 to 2015 using a generalised additive mixed model, with penalised spline smooth functions for covariates. The models included local-scale predictors derived using a Geographic Information System (GIS), daily estimates from a chemical transport model, and daily meteorological characteristics. The models performed well in explaining the variability in daily averaged measured concentrations at 48–85 sites: 63% for NO₂, 77% for PM_{2.5}, 80% for PM₁₀ and 85% for O₃. Outputs of the study include daily air pollution maps that can be applied in epidemiological studies across Great Britain. Daily concentration values can also be predicted for specific locations, such as residential addresses or schools, and aggregated to other exposure time periods (including weeks, months, or pregnancy trimesters) to facilitate the needs of different health analyses.

1. Introduction

Epidemiological studies assessing the health impacts of short-term environmental exposures to air pollution often rely on air pollution measurements obtained from monitoring networks (Mo et al., 2019; Stas et al., 2021; Yang et al., 2020). Assigning concentrations based on measurements from a limited number of monitoring sites to individuals (e.g. cohort participants) in large geographic areas often fails to capture the spatial heterogeneity in ambient air pollution, especially for pervasive sources such as traffic-related air pollutants (Rushworth et al., 2014).

More recently, models representing spatial and temporal air pollution patterns have emerged. Some models, using deterministic techniques, estimate the rate of dispersion and dilution of pollutants between emission source(s) and receiver(s) (e.g. address location) as a function of meteorology, terrain and land use. These ‘dispersion models’ often cannot easily be applied over large geographical areas for epidemiological studies, especially at the national scale, due to high demands in terms of detailed data requirements (such as emission inventories)

and computer processing. Land use regression (LUR) is a statistical approach that has been used extensively to estimate long-term (i.e. annual average) air pollution concentrations, often as an alternative to dispersion modelling. Traditional LUR models combine monitored air pollution concentrations with ‘land use’ information (e.g. land cover, population density, traffic indicators) (Hoek et al., 2008) using multiple regression. LUR has been shown in many locations (Beelen et al., 2013; Hystad et al., 2011; Meng et al., 2016; Muttoo et al., 2018; Rose et al., 2011; van Nunen et al., 2017) to adequately represent the spatial variation in pollution concentrations. Increasingly, LUR includes the ability to add temporal components (e.g. meteorological data, chemical transport models (CTM), emission inventories and satellite-derived data) to allow the prediction of concentrations over shorter periods (e.g. daily). Spatio-temporal LUR models often adopt statistical frameworks to represent the non-linear relationships between time-varying variables and air pollution measurements, and this includes generalised additive mixed methods (GAMM), a relative of the generalised linear model (GLM) (Lindström et al., 2014; Stafoggia et al., 2017; Yanosky et al., 2014; Zhang et al., 2018).

Peer review under responsibility of Turkish National Committee for Air Pollution Research and Control.

* Corresponding author.

E-mail address: weiye.wang14@imperial.ac.uk (W. Wang).

<https://doi.org/10.1016/j.apr.2022.101506>

Received 24 March 2022; Received in revised form 28 June 2022; Accepted 10 July 2022

Available online 12 July 2022

1309-1042/© 2022 Turkish National Committee for Air Pollution Research and Control. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

There is a wealth of health data from routinely collected health registries and longitudinal cohort studies to facilitate epidemiological studies on both the short- and long-term effects of air pollution. There are some examples of high resolution spatio-temporal models, developed to provide adequate exposure assessment for these types of study (Chen et al., 2018; de Hoogh et al., 2019; Shtein et al., 2019; Silibello et al., 2021; Stafoggia et al., 2017; Yu et al., 2022), but these are limited at the national scale for Great Britain (GB). Here, we developed spatio-temporal models to predict daily average ambient concentrations of nitrogen dioxide (NO₂), particulate matter (PM) with aerodynamic diameter $\leq 2.5 \mu\text{m}$ (PM_{2.5}), particulate matter with aerodynamic diameter $\leq 10 \mu\text{m}$ (PM₁₀) and ozone (O₃). We aimed to develop national-scale models that are suitable for predicting exposure at residential addresses and for use in air pollution epidemiological studies in Great Britain.

2. Material and methods

We developed generalised additive mixed models for NO₂, PM_{2.5}, PM₁₀ and O₃ for five years from January 1, 2011 to December 31, 2015 using air pollution measurements from fixed sites, spatial data on land use, road network, traffic, topography, building heights and population density, and spatio-temporal data on meteorological conditions and air pollution maps from chemical transport models.

2.1. Air pollution measurement data

We obtained verified daily mean concentration measurements for NO₂, PM_{2.5}, PM₁₀ and O₃ from 2011 to 2015 from the Automatic Urban and Rural Monitoring Network (AURN). AURN classifies monitoring sites as background urban, background suburban, background rural, traffic urban, industrial urban and industrial suburban. We did not include the small number ($n < 6$) of industrial sites (including ports and airports) as the focus was on optimising model performance at residential addresses for exposure assessment. Monitoring sites with more than 75% daily mean measurements each month and year were included to ensure sufficient daily data availability. Less stringent criteria of 50% data completeness were applied for PM_{2.5} and PM₁₀ due to fewer overall sites. The spatial distribution of the monitoring sites for NO₂ ($n = 85$), PM_{2.5} ($n = 56$), PM₁₀ ($n = 48$) and O₃ ($n = 57$) is shown in the Supplementary Material (Figure A.1).

2.2. Potential predictor variables

We used ERA5-Land reanalysed hourly climate models from the European Centre for Medium-Range Weather Forecasts (ECMWF) with a spatial resolution of approximately $9 \times 9 \text{ km}$ to obtain dew point temperature, temperature measured at 2-m height, and u-component and v-component of wind measured at 10-m height. Based on this information, we computed daily mean temperature, relative humidity, wind speed and wind direction as potential predictors.

Chemical transport models simulate the formation, advection, deposition and dispersion of air pollutants in 3 dimensions, based on emission inventories (location, strength, size) and meteorological inputs (Akita et al., 2014). CTMs have been increasingly used to predict air pollution concentrations at the national/continental scale. We used hourly CTM estimates extracted from the Monitoring Atmospheric Composition and Climate – Interim Implementation ENSEMBLE (MACC-II ENSEMBLE) model provided by Copernicus Atmosphere Monitoring Service (Marécal et al., 2015). MACC-II ENSEMBLE (referred to as “MACC” in the following context) is the median of seven state-of-the-art numerical air quality models in Europe, namely CHIMERE, EMEP, EURAD-IM, LOTOS-EUROS, MATCH, MOCAGE and SILAM with a spatial resolution of 0.1° .

Spatial predictor variables on land cover, road network, road traffic, topography, building, and population density were used to characterise the local environment and represent emission sources or sinks of air

pollution. Description of the spatial predictors, including data source and resolution, is presented in the Supplementary Material (Table A.1).

We used a combined version of the national Land Cover Map 2007 (LCM2007), derived from LANDSAT and processed by the Centre for Ecology & Hydrology, and CORINE Land Cover 2012, an EU-wide map compiled by the European Environmental Agency. LCM2007 has a high spatial resolution of $25 \times 25 \text{ m}$ but includes only two urban land cover classes, limiting its use in urban areas. CORINE has a lower resolution of $100 \times 100 \text{ m}$ but differentiates eleven urban land cover classes, including urban parks. The combined product replaces the two land cover classes from LCM2007 with those from CORINE (Fecht et al., 2014). Land cover classes were then aggregated into nine main groups: high density urban, low density urban, other built-up, ports, industrial, natural areas, green space, agriculture and watercourses.

Road length by road type was extracted from Ordnance Survey (OS)’s Meridian 2 (Ordnance Survey, 2013). Meridian 2 displays road network at a 1:10,000 scale and differentiates between motorway, A-road, B-road, and minor road. Motorway and A-road were grouped to derive a separate class named major road.

Information on road traffic intensity was obtained from Morley and Gulliver (2016). This includes modelled annual average daily traffic (AADT) on all roads in Great Britain, including minor roads. The AADT provided by the Department of Transport contains traffic count data for major roads, but for minor roads, count data on AADT is sparse. The data on traffic flows on minor roads was initially applied in a noise model and significantly improved noise prediction (Morley and Gulliver, 2016). The data includes information on traffic counts for different types of vehicles, including light motor vehicles, medium heavy vehicles, heavy vehicles, powered two-wheelers and all types.

Information on altitude was obtained from the Ordnance Survey Terrain 50 database with a resolution of 50 m. Altitude was extracted from the locations of monitoring sites and included for model development.

Information on buildings is not commonly included as a potential predictor variable in national-scale LUR models. However, studies using building density, building volume and/or aspect ratio (average building height divided by road width) have shown improvement to LUR model performance (Eeftens et al., 2016; Tang et al., 2013). We included building volume as a potential predictor to represent the ventilation in street canyons. We obtained highly detailed footprints and heights of buildings from the OS MasterMap Topography Layer – Building Height Attribute, then calculated building volume (unit: m^3) by multiplying the two attributes.

Information on population and households were obtained as counts for each residential postcode centroid (on average 15 households per postcode, with an average household size of 2.36 residents) from the 2011 Census, provided by the Office for National Statistics (ONS). The population data were used as a proxy for domestic and residential pollution.

2.3. Variable extraction

We quantified spatial predictor variables around monitoring sites by creating circular distance buffers (ranging from 25 m to 5000 m) and linear distance using Geographic Information System (GIS) software ArcGIS v10.4 (ESRI). A complete list of spatial predictor variables is presented in the Supplementary Material (Table A.2). For inverse-distance variables, values for monitoring sites at very close proximity to roads (e.g. less than 3-m) would result in unrealistic concentration values. Therefore, a minimum distance of 3-m (minimum distance to the centre of the road from the roadside) was set for this type of variable.

2.4. Model development

2.4.1. Annual average models

To test the spatio-temporal models’ ability to predict long-term

(annual) air pollution concentrations in comparison with traditional long-term exposure models, we developed separate annual LUR models. The modelling approach followed a standard approach for developing LUR models as described in Eeftens et al. (2012). Potential predictor variable with the highest correlation with all monitored concentrations was first offered to the model, followed by the next ranking variable using a supervised stepwise method. A predictor variable was maintained in the final model if i) the increment of adjusted R^2 was greater than 1%; ii) the coefficient conformed to the pre-determined direction of effect; iii) the p -value was no greater than 0.05. If the resulting model included multiple variables from the same category but different buffers, the variable of less significance was removed. This was to avoid variable overlapping and make the models more intuitively interpretable (Beelen et al., 2013). We included annual CTM estimates (i.e. MACC) in the model as a fixed variable, and other variables were freely selected using the LUR modelling approach. The best combination of predictor variables has the highest R^2 . The output model contains a 'best' set of predictors and associated coefficients.

Each spatial predictor variable was assigned a pre-defined direction of effect. The direction of effect was based on whether the variable acts as a 'source' (+) or 'sink' (-) and was used to guide the variable selection. For NO_2 , $\text{PM}_{2.5}$ and PM_{10} , traffic, population and the built environment (e.g. roads, buildings, industrial lands) are 'sources' that are expected to increase air pollution concentrations. Vegetation, agriculture and watercourse are 'sinks' that are expected to decrease pollution concentrations. The direction of effect was opposite for O_3 . This is because vegetation increases O_3 levels as it emits highly reactive hydrocarbons, especially in growing (warm) seasons, and the presence of nitrogen oxides reacts with the hydrocarbons, which leads to the increase of O_3 concentrations (Kerckhoffs et al., 2015). Using a pre-defined direction of effects improves the application of models and limit the risk of overfitting.

2.4.2. Daily average models

To account for the time-varying spatial variability and reduce computational burden, a two-stage modelling approach was adopted, as described in Yanosky et al. (2014). The modelling approach was developed to represent a trade-off between model complexity and computational efficiency. The approach allows for the inclusion of time-varying smooth spatial terms and smooth terms of spatial predictor variables.

For each pollutant, we developed a daily SpatioTemporal-Land Use Regression (ST-LUR) model for the study period January 1, 2011 to December 31, 2015. The first stage of the ST-LUR model was formulated as follows:

$$y_{i,t} = \mathcal{N}(\mu_i + \sum_{q=1}^Q f_q(M_{i,t,q}) + g_t(S_i), \sigma_t^2) \quad (1)$$

where $y_{i,t}$ is the daily average concentration for site $i = 1 \dots I$ and day $t = 1 \dots T$ ($T = 365$ in a study year ($T = 366$ for the year 2012)); μ_i is the fixed effects for each site; $f_q(\cdot)$ is the one-dimensional penalised spline smooth function for Q time-varying covariate (i.e. CTM estimates and meteorology covariates); $M_{i,t,q}$ are time-varying covariates for $q = 1 \dots Q$; $g_t(\cdot)$ accounts for time-varying residual spatial surface; S_i is the spatial location paired with i th site; σ_t^2 is the daily-varying residual variance.

The second stage is a spatial model that predicts the fitted site-specific terms, $\hat{\mu}_i$,

$$\hat{\mu}_i = \sum_{l=1}^L d_l(X_{i,l}) + g(S_i) + b_i \quad (2)$$

where $X_{i,l}$ are time-invariant spatial predictor covariates for $l = 1 \dots L$, $d_l(\cdot)$ is the one-dimensional penalised spline smooth function for L time-invariant covariate (e.g. land use predictors); $g(\cdot)$ accounts for time-

invariant surface; S_i is the spatial location paired with i th site; b_i is the random effect representing unexplained site-specific variabilities.

Both model stages were fitted using the *gam()* function in the *mgcv* package of R v3.5.1 (Wood, 2018). The first stage (Equation (1)) used an iteratively back-fitting approach to estimate site-specific terms adjusting for time-varying covariate and residual daily spatial variability. The second stage (Equation (2)) fitted a spatial model to $\hat{\mu}_i$, which was obtained from the first stage model, using time-invariant covariates and residual time-invariant spatial variability in air pollution. The non-linearity of the variables was accounted for using spline terms. The degrees of freedom were selected using the default estimate. In the event that a smooth term used high degrees of freedom, the smoothing parameter in the *gam()* function was used to reduce the degrees of freedom. This approach forces a smoother function across the range of the covariate to reduce the potential for overfitting to the data. NO_2 , $\text{PM}_{2.5}$ and PM_{10} measurements and CTM estimates were transformed with the natural logarithm since data were highly right-skewed. Data on O_3 followed a normal distribution pattern and was therefore analysed un-transformed.

We used spatial predictors from the annual models. The predictors were further evaluated in the second stage of the ST-LUR (i.e. fitting a spatial model from time-invariant and residual time-invariant spatial variability). Predictors were only included in the final models if they improved the explained variance, had expected positive/negative direction of effect, and were statistically significant ($p\text{-value} < 0.05$).

2.5. Model evaluation

Five-fold cross-validation (CV) was used to assess the performance of the annual and daily models. Briefly, the sites were randomly allocated into five groups (folds), and the groups were checked to have a similar number of sites from each area (we divided Great Britain into five areas: South, Midlands, North, Scotland and Wales) and site type. Each group was removed from the dataset as a testing set sequentially, with the remaining four groups being fitted to a model. Predictions were made for the held-out dataset. The accuracy and precision of the models were quantified by root mean squared error (RMSE), root mean squared deviation (RMSD) and normalised root mean square error (NRMSE, normalised by the range of measured concentrations), which measure the magnitude of residual error; and cross-validated R^2 (CV- R^2), which is the coefficient of determination between the held-out measurements and model predictions. The daily average models were also assessed on prediction bias (i.e. mean difference between measurements and cross-validation predictions), and slope was obtained by regressing predicted against monitored concentrations in a linear regression model.

2.6. Sensitivity analysis

We applied two sensitivity analyses within the context of the spatio-temporal models. Firstly, we removed the smooth terms in both $f_q(M_{i,t,q})$ and $d_l(X_{i,l})$ for the time-varying and time-invariant covariates, respectively. This was to test models' predictive ability when using linear terms in place of smooth regression. Secondly, we left out all covariates (time-varying and time-invariant) and used spatial smoothing only to test the contribution of the covariates to model performance.

3. Results

3.1. Annual average models

The annual models for NO_2 , $\text{PM}_{2.5}$, PM_{10} and O_3 included similar types of variables. Traffic-related variables were most frequently included in models indicative of road transport emissions. Traffic load in a 50 m buffer was included in all models, followed by (major) road length in a 1000 m buffer. NO_2 , $\text{PM}_{2.5}$ and PM_{10} models also included building volume in a 50 m buffer, accounting for human activity and

street canyon effect formed by tall buildings. Each model (except for NO₂) included a 'pollution sink' variable (natural land or greenspace in various buffer sizes), which fit the pre-defined direction of effect. The full list of the LUR variables is presented in the Supplementary Material (Table A.3).

Model performance was calculated for individual years. The statistics are summarised in the Supplementary Material (Table A.5). Briefly, NO₂ and O₃ showed consistent performance across years, with CV-R² ranging from 0.75 to 0.78 and RMSE from 6.8 to 7.4 µg/m³ for NO₂; and CV-R² ranging from 0.74 to 0.94 and RMSE from 3.1 to 4.8 µg/m³ for O₃. The PM models performed less well, with CV-R² ranging from 0.21 to 0.62 and RMSE from 1.2 to 2.9 µg/m³ for PM_{2.5}; and CV-R² ranging from 0.47 to 0.69 and RMSE from 1.6 to 3.3 µg/m³ for PM₁₀. Unlike NO₂ or O₃, the PM annual models were very sensitive to variable selection, and the same set of variables performed more or less well for specific years. We further analysed the PM annual LUR models using the same monitoring sites and selection criteria but allowing each year to have different variables to maximise explained variance. The results showed a significant improvement in R², ranging from 0.61 to 0.68 for PM_{2.5} and from 0.66 to 0.76 for PM₁₀. We explored various combinations of variables, but none could well explain the spatial variation of the measured concentrations for all study years. This may be due to the smaller number of monitoring sites and low variability in the PM annual mean measurements (standard deviation (SD) ranging from 2.0 to 3.4 for PM_{2.5}, and from 2.8 to 4.7 for PM₁₀).

The CTM-incorporated annual LUR models were compared with the aggregated daily ST-LUR models, and results are presented in Section 3.2. We produced air pollution maps based on the annual average LUR models, and details can be found in Supplementary Material (Appendix B).

3.2. Daily average models

Table 1 shows summary statistics of the measured daily air pollution concentrations. The full list of the time-varying and time-invariant covariates included in the final models is presented in the Supplementary Material (Table A.4). The performance statistics of the ST-LUR model for all pollutants are shown in Table 2. Overall, the results indicate strong performance. The NO₂ model, including both space-time predictors (CTM estimates MACC and meteorology) and spatial predictors (traffic load in a 50 m buffer, major road length in a 1000 m buffer and building volume around main roads in a 50 m buffer), explained 63% of the variation in measured NO₂ concentrations (CV-R²: 0.63), a decrease of 9.2% compared to the Model R². The RMSE was 12.7 µg/m³, bias was 8.5 µg/m³ and slope was 0.79. Both PM_{2.5} and PM₁₀ models included MACC, meteorology covariates, traffic load in a 50 m buffer, natural land in a 1000 m buffer and building volume around main roads in a 50 or 100 m buffer. PM_{2.5} and PM₁₀ models showed similar validation performance, with CV-R² of 0.77 and 0.80, respectively (RMSE 5.0 µg/m³, bias 3.0 µg/m³ and slope 0.81 for PM_{2.5}; RMSE 5.0 µg/m³, bias 3.4 µg/m³ and slope 0.91 for PM₁₀). The spatial predictors of the O₃ model were similar to the NO₂ models, including

traffic load (50 m buffer), major road length (1000 m buffer) and green space (300 m buffer). The O₃ models explained the most variation in the measured concentrations (CV-R² = 0.85) compared to the other pollutants. The RMSE was 7.6 µg/m³, bias was 5.6 µg/m³ and the slope was close to 1 (0.96). In the sensitivity analysis, we showed that our ST-LUR using smooth functions for the spatial and spatio-temporal predictors is better than using linear terms at capturing space-time variability in measured concentrations. Models with linear terms reduced CV-R² by 0.17 for NO₂, and 0.09–0.12 for PM_{2.5}, PM₁₀ and O₃, suggesting that smoothing the covariates was more effective for NO₂ than the other pollutants. The spatial smoothing only models had a lower agreement when leaving out covariates, with CV-R² dropping by 0.12–0.27, depending on the pollutants.

Table 3 shows the model performance by site type. Scatterplots of measurements versus predicted concentrations at each site are presented in the Supplementary Material (Figure A.2-A.5). For NO₂, the performance varied by site type. Larger errors were found for some urban traffic sites, and the model tended to underestimate at background sites (Figure A.2-A). Most traffic sites with a 50 m buffer contain both major and minor roads. Since the traffic counts on minor roads were estimated, this could induce errors in exposure predictions. Moreover, the poorer performance at some sites (R² < 0.50) was because the model could not capture some measurements with extremely high concentrations. For example, the highest daily measurements for London Marylebone and Glasgow Kerbside (both traffic urban sites) during the study period were 136 µg/m³ and 169 µg/m³, respectively, whereas the medians were 23 µg/m³ and 33 µg/m³, respectively. In addition, a few traffic sites have low daily mean concentrations of less than 30 µg/m³ but relatively high values for the traffic load covariate, which resulted in a large bias in the prediction. The PM_{2.5} and PM₁₀ models performed well at most sites. Model prediction for O₃ also exhibited fewer errors, with good performance at all background sites. The O₃ model only included two traffic sites, of which Birmingham Tyburn Roadside yielded a good fit, whereas the prediction for London Marylebone was poor. The London Marylebone site was not an outlier as was checked by Cook's D (influence statistics < 1), and the poor prediction may also be explained by the large values of the traffic load and road length covariates.

Fig. 1 shows the mapped model output for NO₂ on six consecutive days from 1st to 6th January, 2015 for Greater London, as an example. The maps illustrate the space and time heterogeneity of NO₂ concentrations captured by the models. Spatial patterns are clearly visible and predominantly driven by heavily trafficked major roads. The estimated daily mean concentrations across all London postcodes are lowest on 1st January, a public holiday (23.0 µg/m³), and increase over the next days until 4th January (mean 39.7 µg/m³), before slightly falling.

In order to test each spatio-temporal model's ability in long-term prediction, we calculated the annual average concentration from the daily estimates of ST-LUR at each monitoring site for each study year (referred to as "long-term ST-LUR"). The annual average estimates were regressed against annual measurements at monitoring sites, and results are summarised in Table 4. For the individual study years, the NO₂ and O₃ models predicted long-term spatial trends well, with a mean R² of

Table 1

Descriptive statistics for daily monitoring data of NO₂, PM_{2.5}, PM₁₀ and O₃ from 2011 to 2015. Q1, median, mean, Q3 and SD are in µg/m³.

	N ^a	Q1 ^b	Median	Mean	Q3 ^c	SD ^d	CV ^e	Skewness	Completeness ^f
NO ₂	130,906	11.6	22.8	27.4	37.6	21.1	77.0	1.4	94.4
PM _{2.5}	83,783	7.0	9.8	12.7	15.4	9.4	74.1	2.2	85.8
PM ₁₀	67,294	11.2	15.4	18.3	22.0	10.9	59.4	2.0	84.0
O ₃	97,791	34.0	48.0	47.1	60.8	19.2	40.9	-0.1	95.7

^a N: number of daily mean observations.

^b Q1: 1st quantile (unit: µg/m³).

^c Q3: 3rd quantile (unit: µg/m³).

^d SD: standard deviation (unit: µg/m³).

^e CV: coefficient of variation = SD/Mean * 100%.

^f Completeness: % day with valid measurement from monitoring stations.

Table 2
ST-LUR model performance with five-fold cross-validation. RMSE, RMSD and bias are in $\mu\text{g}/\text{m}^3$.

	N ^a	n ^b	Model-R ²	CV-R ²	RMSE	RMSD	NRMSE (%)	Bias	Intercept	Slope
NO ₂	130,906	85	0.72	0.63	12.8	0.5	3.9	8.5	6.32	0.79
PM _{2.5}	83,783	56	0.93	0.77	5.0	0.6	3.3	3.0	2.32	0.81
PM ₁₀	67,294	48	0.92	0.80	5.0	0.5	3.0	3.4	1.95	0.91
O ₃	97,791	57	0.92	0.85	7.6	0.3	4.9	5.7	2.28	0.96

^a N: number of daily mean predictions.

^b n: number of monitoring sites.

Table 3
Model performance by site type.

Pollutant	Site Type	N ^a	R ² (range, median) ^b	RMSE (range, median)
NO ₂	Traffic Urban	28	0.55 (0.11–0.86, 0.58)	1.0 (0.5–2.4, 0.7)
	Background Urban	38	0.67 (0.25–0.88, 0.73)	0.7 (0.3–1.7, 0.6)
	Background Suburban	4	0.70 (0.39–0.85, 0.79)	0.8 (0.4–1.3, 0.8)
	Background Rural	15	0.65 (0.37–0.85, 0.65)	1.0 (0.5–1.7, 1.0)
PM _{2.5}	Traffic Urban	17	0.86 (0.50–0.95, 0.89)	0.5 (0.3–1.0, 0.4)
	Background Urban	34	0.87 (0.49–0.95, 0.89)	0.5 (0.3–1.0, 0.5)
	Background Suburban	2	0.93 (0.91–0.94, 0.93)	0.4 (0.4–0.4, 0.4)
	Background Rural	3	0.83 (0.71–0.89, 0.88)	0.8 (0.4–1.3, 0.6)
PM ₁₀	Traffic Urban	17	0.81 (0.53–0.93, 0.86)	0.6 (0.3–2.9, 0.4)
	Background Urban	27	0.87 (0.58–0.94, 0.90)	0.4 (0.3–0.9, 0.4)
	Background Rural	4	0.80 (0.66–0.93, 0.81)	0.8 (0.3–1.9, 0.5)
O ₃	Traffic Urban	2	0.66 (0.42–0.90, 0.66)	0.4 (0.3–0.5, 0.4)
	Background Urban	32	0.88 (0.68–0.94, 0.89)	0.3 (0.2–0.7, 0.3)
	Background Suburban	3	0.90 (0.88–0.92, 0.89)	0.3 (0.2–0.3, 0.3)
	Background Rural	20	0.86 (0.65–0.95, 0.88)	0.4 (0.2–0.8, 0.4)

^a N: number of monitoring sites.

^b R²: average coefficient of determination, calculated by regressing the daily predictions and daily measurements for each site.

0.70 (range, 0.66–0.72) and a mean RMSE of 8.1 $\mu\text{g}/\text{m}^3$ (range, 7.4–8.9 $\mu\text{g}/\text{m}^3$) for NO₂; and a mean R² of 0.77 (range, 0.66–0.88) and a mean RMSE of 4.9 $\mu\text{g}/\text{m}^3$ (range, 3.8–5.6 $\mu\text{g}/\text{m}^3$) for O₃. The R² was moderate for PM₁₀, ranging from 0.57 to 0.64, and RMSE ranging from 1.8 to 3.2 $\mu\text{g}/\text{m}^3$. The performance varied largely by year for PM_{2.5}, with the lowest R² in 2011 (R² = 0.32, RMSE = 3.2 $\mu\text{g}/\text{m}^3$) and highest in 2015 (R² = 0.58, RMSE = 1.6 $\mu\text{g}/\text{m}^3$). A comparison of the model performance of the long-term ST-LUR and CTM-incorporated annual LUR models (referred to as “annual LUR”) are presented in the Supplementary Material (Table A.5). The annual LUR outperformed the long-term ST-LUR for NO₂ and O₃ across all study years. However, the difference of model performance between the two was small in terms of R² and RMSE, with differences (model performance parameter for annual LUR minus long-term ST-LUR) of R² ranging from 0.05 to 0.09 and RMSE from 0.55 to 1.52 $\mu\text{g}/\text{m}^3$ for NO₂; and R² ranging from 0.03 to 0.08 and RMSE ranging from 0.68 to 1.15 $\mu\text{g}/\text{m}^3$ for O₃. For PM_{2.5} and PM₁₀, performance varied by year for both sets of models, though the estimates from the long-term ST-LUR exhibited slightly more bias than the annual LUR models (differences of average bias were 1.5 $\mu\text{g}/\text{m}^3$ for PM_{2.5} and 2.0 $\mu\text{g}/\text{m}^3$ for PM₁₀).

4. Discussion

We developed spatio-temporal air pollution models using generalised additive mixed models to predict daily average concentrations for NO₂, PM_{2.5}, PM₁₀ and O₃. This is the first study to combine spatial and spatio-temporal predictors in a modelling framework to predict daily concentrations of multiple air pollutants for every 25 m grid cell for the whole of Great Britain. The results demonstrated good overall model performance and the ability to use the daily models as the basis for annual average concentration predictions. The models’ key feature is to estimate daily NO₂, PM_{2.5}, PM₁₀ and O₃ exposures at any point location in Great Britain. The models can be efficiently reproduced for another time period or region, as the modelling process was written in R (v3.5.1), with the option to use data on measured concentrations for local calibration.

4.1. Predictor variables

The selected GIS-derived spatial covariates were consistent with the findings in most LUR studies (Clougherty et al., 2013; de Hoogh et al., 2018; Vienneau et al., 2013). The covariates were dominated by four types of variables: traffic load, building volume, road length and vegetation cover (green space and natural land). The anthropogenic emissions were, in most cases, represented by road length and traffic load-related variables. Population/household count, altitude and other ‘land cover’ variables were selected in the first stage but were not included in the model as they did not meet the criteria (e.g. not statistically significant or improving explained variance). The covariate selection process accounted for the spatial trends of residuals and allowed non-linearity in the effects of the covariates. The process is computationally feasible for large-area applications and has the potential to include other land use variables that have the potential for improving model performance in different geographic locations.

Although industrial emissions (including ports and airports) can contribute substantially to ambient air pollution concentrations at some locations, industrial variables are not common in LUR models (Beelen et al., 2013). This is partly due to the lack of detailed industrial emission data. Some local/regional models contained industrial land-use variables indicating local industrial sources (Eeftens et al., 2016; Meng et al., 2015), but this is not common in national-scale models. In essence, LUR does not lend itself to dealing with infrequent, discrete sources. Typically, this is the case where a source is related to between none and a maximum of two monitoring sites used to train models. Our models, for example, do not recognise Heathrow airport as an emission source. In this study, the land cover data included two industrial variables (i.e. ports and industrial land) but were not offered in the final model as they were not statistically significant. The industrial sources are included in CTM as diffuse sources, but the MACC CTM is not granular enough to capture individual sources and localised concentrations gradients.

PM_{2.5}, PM₁₀ and O₃ concentrations were explained mainly by the daily CTM as it captures the space and time patterns of pollutants operating on regional scales. For NO₂, the CTM estimates or meteorological data made a relatively small contribution to modelling. NO₂ concentrations were mostly captured by the time-invariant land use variables. We did not explore the use of emission inventories and

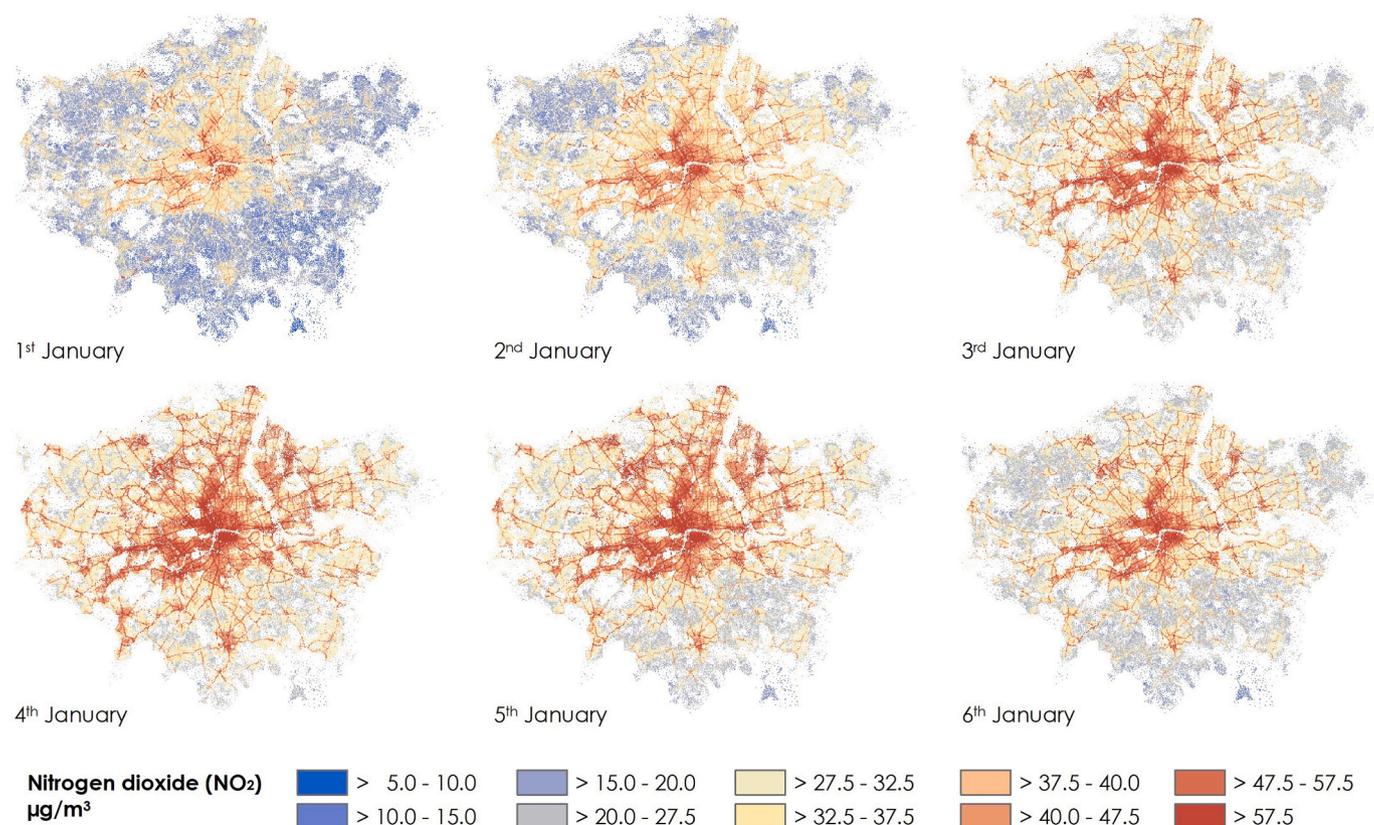


Fig. 1. NO₂ surface for postcode centroids in Greater London from Thursday, 1st January to Tuesday, 6th January, 2015.

Table 4

Summary statistics for long-term (annual) ST-LUR models. RMSE, RMSD and bias are in $\mu\text{g}/\text{m}^3$.

	Year	N	R ²	RMSE	RMSD	NRMSE (%)	Bias	Intercept	Slope
NO ₂	2011	79	0.70	8.3	0.4	14.5	5.8	5.24	0.82
	2012	73	0.66	8.9	0.5	14.0	6.3	5.54	0.82
	2013	74	0.72	8.0	0.5	14.2	5.4	4.01	0.89
	2014	74	0.72	8.1	0.4	14.4	5.6	3.55	0.89
	2015	76	0.72	7.4	0.5	13.6	5.2	3.57	0.89
PM _{2.5}	2011	49	0.32	3.2	0.9	13.3	2.3	7.93	0.48
	2012	54	0.38	2.8	0.9	12.8	1.9	6.53	0.51
	2013	55	0.46	2.5	0.8	12.0	1.8	6.45	0.52
	2014	52	0.45	2.2	1.0	13.4	1.6	6.39	0.49
	2015	51	0.58	1.6	0.8	13.3	1.2	4.08	0.61
PM ₁₀	2011	43	0.57	3.2	0.5	11.1	2.4	4.09	0.82
	2012	47	0.58	2.7	0.8	13.6	2.3	3.41	0.83
	2013	46	0.61	2.6	0.7	12.7	2.1	3.77	0.80
	2014	39	0.64	2.3	0.8	12.9	1.9	4.52	0.74
	2015	39	0.63	1.8	0.7	8.6	1.6	3.72	0.78
O ₃	2011	55	0.66	5.6	0.5	11.0	4.3	5.43	0.89
	2012	57	0.75	5.2	0.4	9.2	3.9	7.69	0.83
	2013	55	0.88	3.8	0.3	6.8	3.0	5.00	0.90
	2014	56	0.80	4.8	0.3	9.0	3.6	3.94	0.92
	2015	55	0.78	4.9	0.4	9.0	3.7	6.33	0.88

satellite-derived data in our models. These data have also been used in various spatio-temporal models for short-term air pollution estimates (de Hoogh et al., 2019; Di et al., 2020; Keller et al., 2015; Korek et al., 2016; Li et al., 2021; Meng et al., 2015; Stafoggia et al., 2017; Yanosky et al., 2014). The key advantage of the MACC CTM over these time-varying terms and other CTM data (e.g. GEOS-Chem, CMAQ) is the trade-off between accessibility, spatial resolution and full temporal coverage of the study period.

4.2. Spatio-temporal modelling

Performance of daily models was particularly strong for PM_{2.5}, PM₁₀ and O₃ (CV-R² range, 0.77–0.85) and more modest for NO₂ (R² = 0.63). This is not unexpected, as NO₂ is a pollutant whose main source in Great Britain is road traffic and therefore shows more spatial than temporal variability. The use of non-linear effects for the spatial and spatio-temporal predictors improved the model performance. In contrast, spatial smoothing only models vastly decreased predictive power, suggesting that the covariates in the original models played an important role in capturing the local space-time heterogeneity that cannot be

captured by spatial smoothing alone.

Due to their wide range of values, traffic-related predictors were more likely to generate uncertainty in model performance. For example, values for traffic load within a 50 m buffer ranged from 0 to 16, 344, 189 (unit: vehicle per day \times road length) across monitoring sites. Even though the variable was scaled and log-transformed, model predictions were exceptionally high at sites with high traffic load. To limit the influence of traffic-related predictors, we truncated these variables to predicting no larger than the highest observed values, as in some other studies (Akita et al., 2014; Beelen et al., 2009; Yanosky et al., 2014). However, concentrations might still be overestimated due to the combined effects of the covariates, which poses a limitation of LUR-based models. Such variables are particularly influential for large geographic area models when predicting concentrations for high-traffic locations such as road junctions. Therefore, unrealistic predictions were assessed on an individual basis and truncated, if necessary.

Measured concentrations of PM_{2.5} and PM₁₀ have low spatial variation in the study area (the spatial variance was 2.78 $\mu\text{g}/\text{m}^3$ for PM_{2.5} and 3.96 $\mu\text{g}/\text{m}^3$ for PM₁₀), which may explain models' limited ability in predicting spatial variability. Monitored NO₂ and O₃ have a relatively large spatial variance of 16.68 $\mu\text{g}/\text{m}^3$ and 9.98 $\mu\text{g}/\text{m}^3$, respectively. The unexplained spatial variability could potentially be reduced by a denser monitoring network. However, this cannot be easily achieved for national-scale studies in the short term. The placement of monitoring sites usually takes into account the ambient environment (e.g. population density and traffic density) and convenience. In Great Britain, monitoring sites are much denser in urban areas, and the number of sites varies largely by region. Pollution models unavoidably include more urban sites than suburban and rural sites. This may introduce biased estimates because the monitors in urban areas would capture higher pollution concentrations (lower for O₃).

We used a smaller number of measurement sites from AURN than were available to develop the models. One option was to borrow sites from another large monitoring network, the London Air Quality Network (LAQN). However, this would have resulted in oversampling in London relative to the rest of the UK, especially for roadside sites. Another option was to use monitoring stations (outside London) run by Local Authorities or organisations. However, the data quality cannot be assured, and/or the stations cannot provide continuous measurements for the study period.

The two-stage GAMM approach primarily followed the approaches used in Yanosky et al. (2014), in that they developed two-stage GAMM models for predicting monthly concentrations of PM_{2.5} and PM₁₀ for the conterminous United States (U.S.). Their predictors included geographic data and point source emissions (spatial) and meteorological data (spatio-temporal). The 1999–2007 models could explain 77% (cross-validated) of the variation in daily PM_{2.5} and 58% (cross-validated) of the variation in daily PM₁₀. Compared to Yanosky et al. (2014), our study included a broader range of GIS predictors and used daily CTM estimates in addition to meteorological data as time-varying covariates. The GIS predictors reduced spatial uncertainty in the model while allowing small-scale variation. The daily CTM incorporated meteorological data and accounted for concentrations through transmission, which reflected a more reliable temporal variation of pollution.

An increasing number of studies use machine learning algorithms to develop spatio-temporal air pollution models at a large/national scale. In a model for Great Britain, Schneider et al. (2020) integrated land use, meteorological data, CTM estimates and satellite data in a multi-stage random forest model to estimate daily PM_{2.5} concentrations. The models explained an average of 77.8% of the variation in daily PM_{2.5} measurements across Great Britain for the period 2011–2015, which is very similar to the result of our study (77.1%). Their spatial (annual average) CV-R² yielded about 20% more spatial variation than our long-term ST-LUR. This may be due the larger number of monitoring sites deployed in their study, in which random forest was used to calculate PM_{2.5} concentrations from the more abundant PM₁₀ sites,

which could also increase the variability in the PM_{2.5} measurements. In a similar study, Di et al. (2019) used the same types of predictors in an ensemble strategy rooted in machine learning, including random forest, the gradient boosting machine and neural network, to develop exposure models of PM_{2.5} for the U.S. They found the ensemble strategy was more effective in predicting concentrations across time than individual base learners (cross-validated R² = 0.86), as the approach can remedy the drawbacks from a single method (e.g. random forest tends to over-estimate). The machine learning approach overall showed promising results and often outperformed alternative methods (Shtein et al., 2019; Stafoggia et al., 2020). The methods we presented in this study, which combined a wide range of spatial and spatio-temporal predictors, were less complex and less computationally intensive whilst achieving good model performance.

4.3. Limitations

Our models have some limitations. First, the models used a relatively small number of monitoring sites, especially for PM_{2.5} and PM₁₀. The poor model performance in the held-out groups is likely due to the lack of sites and/or low variability in pollution concentrations. It is also difficult to identify and eliminate all extreme values caused by local events such as fires. Second, for a national model in Great Britain, the selection of monitoring sites is prone to bias due to the unevenness, and it is hard to balance density and representativeness. Third, the model accuracy is highly dependent on the input data. For example, we used improved traffic data that included estimated traffic counts on minor roads. The data was from modelling and therefore has an inherent error. This could have contributed to the poor performance at some traffic sites. Last, the spatial resolution of the MACC-II Ensemble CTM is relatively coarse (0.1° or approximately 10 \times 10 km²). Recently, processing algorithms have been developed which produce continental/global daily air pollution estimates at improved spatial resolution. This includes Multi-Angle Implementation of Atmospheric Correction (MAIAC), which downscales aerosol optical depth (AOD) retrievals to 1 km resolution and has been successfully implemented in several spatio-temporal models for particulate matter (Di et al., 2019; Li et al., 2021; Schneider et al., 2020; Shtein et al., 2019). However, to our knowledge, there is currently no freely available CTM/satellite information on NO₂ or O₃ with higher resolution than MACC. In this study, we did not have the capacity to run a daily CTM at a finer resolution. Therefore, we used MACC which is freely available and has full spatial and temporal coverage for the four pollutants.

5. Conclusions

We used generalised additive mixed models to predict daily mean concentrations of NO₂, PM_{2.5}, PM₁₀ and O₃ for Great Britain. The models combined spatial and spatio-temporal predictors, including land use variables, CTM estimates and meteorological parameters. Despite the limitations, the models could explain 63%–85% (cross-validated) of the variation in daily concentration measurements with a low implementation cost. In addition, the models allow the flexibility to aggregate to health-relevant exposure windows to facilitate large-scale short- and mid-term exposure studies, such as daily for time-series analyses and pregnancy trimesters and infancy exposure periods for studies on birth outcomes. The air pollution maps we produced are at fine resolution (25 \times 25 m) and capture the space-time variation in concentrations related to most localised emissions sources and background sources. Epidemiological studies that need air pollution exposure of point locations can benefit from this study by simply providing coordinates or postcodes and extracting values from the maps/models.

Credit author statement

Weiyi Wang: Methodology, Software, Formal analysis,

Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Daniela Fecht**: Writing – review & editing, Visualization, Supervision, Project administration. **Sean Beevers**: Writing – review & editing, Supervision, Project administration. **John Gulliver**: Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

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

Acknowledgements

This research was supported by the MRC Centre for Environment and Health, which is currently funded by the Medical Research Council (MR/S019669/1, 2019–2024). Infrastructure support for the Department of Epidemiology and Biostatistics was provided by the NIHR Imperial Biomedical Research Centre. This study is part-funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Chemical and Radiation Threats and Hazards (NIHR-200922), a partnership between UK Health Security Agency and Imperial College London. The views expressed are those of the authors and not necessarily those of the NIHR, UK Health Security Agency or the Department of Health and Social Care.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.apr.2022.101506>.

References

- Akita, Y., Baldasano, J.M., Beelen, R., Cirach, M., de Hoogh, K., Hoek, G., Nieuwenhuijsen, M., Serre, M.L., de Nazelle, A., 2014. A large scale air pollution estimation method combining Land Use Regression and Chemical Transport Modeling in a geostatistical framework. *Environ. Sci. Technol.* 48, 4452–4459. <https://doi.org/10.1021/es405390e>.
- Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union. *Sci. Total Environ.* 407 (6), 1852–1867. <https://doi.org/10.1016/j.scitotenv.2008.11.048>.
- Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K.T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., et al., 2013. Development of NO₂ and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe - the ESCAPE project. *Atmos. Environ.* 72, 10–23. <https://doi.org/10.1016/j.atmosenv.2013.02.037>.
- Chen, G., Li, S., Knibbs, L.D., Hamm, N.A.S., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M.J., Guo, Y., 2018. A machine learning method to estimate PM_{2.5} concentrations across China with remote sensing, meteorological and land use information. *Sci. Total Environ.* 636, 52–60. <https://doi.org/10.1016/j.scitotenv.2018.04.251>.
- Clougherty, J.E., Kheirbek, I., Eisl, H.M., Ross, Z., Pezeshki, G., Gorzczynski, J.E., Johnson, S., Markowitz, S., Kass, D., Matte, T., 2013. Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: the New York City Community Air Survey (NYCCAS). *J. Expo. Sci. Environ. Epidemiol.* 23 (3), 232–240. <https://doi.org/10.1038/jes.2012.125>.
- de Hoogh, K., Chen, J., Gulliver, J., Hoffmann, B., Hertel, O., Ketzler, M., Bauwelinck, M., van Donkelaar, A., Hvidtfeldt, U.A., Katsouyanni, K., Klompaker, J., Martin, R.V., Samoli, E., Schwartz, P.E., Stafoggia, M., Bellander, T., Strak, M., Wolf, K., Vienneau, D., et al., 2018. Spatial PM_{2.5}, NO₂, O₃ and BC models for western Europe – evaluation of spatiotemporal stability. *Environ. Int.* 120, 81–92. <https://doi.org/10.1016/j.envint.2018.07.036>.
- de Hoogh, K., Saucy, A., Shtein, A., Schwartz, J., West, E.A., Strassmann, A., Puhon, M., Roösl, M., Stafoggia, M., Kloog, I., 2019. Predicting fine-scale daily NO₂ for 2005–2016 incorporating OMI satellite data across Switzerland. In: *Environmental Science and Technology*, vol. 53. American Chemical Society, pp. 10279–10287. <https://doi.org/10.1021/acs.est.9b03107>. Issue 17.
- Di, Q., Amini, H., Shi, L., Kloog, I., Silvern, R., Kelly, J., Sabath, M.B., Choirat, C., Koutrakis, P., Lyapustin, A., Wang, Y., Mickley, L.J., Schwartz, J., 2019. An ensemble-based model of PM_{2.5} concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* 130, 104909. <https://doi.org/10.1016/j.envint.2019.104909>.
- Di, Q., Amini, H., Shi, L., Kloog, I., Silvern, R., Kelly, J., Sabath, M.B., Choirat, C., Koutrakis, P., Lyapustin, A., Wang, Y., Mickley, L.J., Schwartz, J., 2020. Assessing NO₂ concentration and model uncertainty with high spatiotemporal resolution across the contiguous United States using ensemble model averaging. *Environ. Sci. Technol.* 54 (3), 1372–1384. <https://doi.org/10.1021/acs.est.9b03358>.
- Eeftens, M., Beelen, R., De Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dedele, A., Dons, E., De Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Grazuleviciene, R., Heinrich, J., Hoffmann, B., Jerrett, M., et al., 2012. Development of land use regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM_{coarse} in 20 European study areas; Results of the ESCAPE project. *Environ. Sci. Technol.* 46 (20), 11195–11205. <https://doi.org/10.1021/es301948k>.
- Eeftens, M., Meier, R., Schindler, C., Aguilera, I., Phuleria, H., Ineichen, A., Davey, M., Ducret-Stich, R., Keidel, D., Probst-Hensch, N., Kunzli, N., Tsai, M.-Y., 2016. Development of land use regression models for nitrogen dioxide, ultrafine particles, lung deposited surface area, and four other markers of particulate matter pollution in the Swiss SAPALDIA regions. *Environ. Health* 15. <https://doi.org/10.1186/s12940-016-0137-9>.
- Fecht, D., Beale, L., Briggs, D., 2014. A GIS-based urban simulation model for environmental health analysis. *Environ. Model. Software* 58, 1–11. <https://doi.org/10.1016/j.envsoft.2014.03.013>.
- Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* 42 (33), 7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>.
- Hystad, P., Setton, E., Cervantes, A., Poplawski, K., Deschenes, S., Brauer, M., van Donkelaar, A., Lamsa, L., Martin, R., Jerrett, M., Demers, P., 2011. Creating national air pollution models for population exposure assessment in Canada. *Environ. Health Perspect.* 119 (8), 1123–1129. <https://doi.org/10.1289/ehp.1002976>.
- Keller, J.P., Olives, C., Kim, S.Y., Sheppard, L., Sampson, P.D., Szpiro, A.A., Oron, A.P., Lindström, J., Vedal, S., Kaufman, J.D., 2015. A unified spatiotemporal modeling approach for predicting concentrations of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. *Environ. Health Perspect.* 123 (4), 301–309. <https://doi.org/10.1289/ehp.1408145>.
- Kerckhoffs, J., Wang, M., Meliefste, K., Malmqvist, E., Fischer, P., Janssen, N.A.H., Beelen, R., Hoek, G., 2015. A national fine spatial scale land-use regression model for ozone. *Environ. Res.* 140, 440–448. <https://doi.org/10.1016/j.envres.2015.04.014>.
- Korek, M., Johansson, C., Svensson, N., Lind, T., Beelen, R., Hoek, G., Pershagen, G., Bellander, T., 2016. Can dispersion modeling of air pollution be improved by land-use regression? An example from Stockholm, Sweden. *J. Exposure Sci. Environ. Epidemiol.* 1–7. <https://doi.org/10.1038/jes.2016.40>, 2015.
- Li, J., Garshick, E., Hart, J.E., Li, L., Shi, L., Al-Hemoud, A., Huang, S., Koutrakis, P., 2021. Estimation of ambient PM_{2.5} in Iraq and Kuwait from 2001 to 2018 using machine learning and remote sensing. *Environ. Int.* 151, 106445. <https://doi.org/10.1016/j.envint.2021.106445>.
- Lindström, J., Szpiro, A.A., Sampson, P.D., Oron, A.P., Richards, M., Larson, T.V., Sheppard, L., 2014. A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environ. Ecol. Stat.* 21 (3), 411–433.
- Maréchal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chérroux, F., Colette, A., Coman, A., Curier, R.L., Denier van der Gon, H.A.C., Droouin, A., Elbern, H., Emili, E., Engelen, R.J., et al., 2015. A regional air quality forecasting system over Europe: the MACC-II daily ensemble production. *Geosci. Model Dev. (GMD)* 8 (9), 2777–2813. <https://doi.org/10.5194/gmd-8-2777-2015>.
- Meng, X., Chen, L., Cai, J., Zou, B., Wu, C.F., Fu, Q., Zhang, Y., Liu, Y., Kan, H., 2015. A land use regression model for estimating the NO₂ concentration in Shanghai, China. *Environ. Res.* 137, 308–315. <https://doi.org/10.1016/j.envres.2015.01.003>.
- Meng, X., Fu, Q., Ma, Z., Chen, L., Zou, B., Zhang, Y., Xue, W., Wang, J., Wang, D., Kan, H., Liu, Y., 2016. Estimating ground-level PM₁₀ in a Chinese city by combining satellite data, meteorological information and a land use regression model. *Environ. Pollut.* 208, 177–184. <https://doi.org/10.1016/j.envpol.2015.09.042>.
- Mo, Z., Fu, Q., Lyu, D., Zhang, L., Qin, Z., Tang, Q., Yin, H., Xu, P., Wu, L., Wang, X., Lou, X., Chen, Z., Yao, K., 2019. Impacts of air pollution on dry eye disease among residents in Hangzhou, China: a case-crossover study. *Environ. Pollut.* 246, 183–189. <https://doi.org/10.1016/j.envpol.2018.11.109>.
- Morley, D.W., Gulliver, J., 2016. Methods to improve traffic flow and noise exposure estimation on minor roads. *Environ. Pollut.* 216, 746–754. <https://doi.org/10.1016/j.envpol.2016.06.042>.
- Muttoo, S., Ramsay, L., Brunekreef, B., Beelen, R., Meliefste, K., Naidoo, R.N., 2018. Land use regression modelling estimating nitrogen oxides exposure in industrial south Durban, South Africa. *Sci. Total Environ.* 1439–1447. <https://doi.org/10.1016/j.scitotenv.2017.07.278>, 610–611.
- Ordnance Survey, 2013. *Meridian 2: User Guide and Technical Specification*.
- Rose, N., Cowie, C., Gillett, R., Marks, G.B., 2011. Validation of a spatiotemporal land use regression model incorporating fixed site monitors. *Environ. Sci. Technol.* 45 (1), 294–299. https://doi.org/10.1021/ES100683T/SUPPL_FILE/ES100683T_SI_001.PDF.
- Rushworth, A., Lee, D., Mitchell, R., 2014. A spatio-temporal model for estimating the long-term effects of air pollution on respiratory hospital admissions in Greater London. *Spatial Spatio-Temp. Epidemiol.* 10, 29–38. <https://doi.org/10.1016/j.sste.2014.05.001>.
- Schneider, R., Vicedo-Cabrera, A.M., Sera, F., Masselot, P., Stafoggia, M., de Hoogh, K., Kloog, I., Reis, S., Vieno, M., Gasparrini, A., 2020. A satellite-based spatio-temporal

- machine learning model to reconstruct daily PM_{2.5} concentrations across great Britain. *Rem. Sens.* 12 (22), 1–19. <https://doi.org/10.3390/rs12223803>.
- Shtein, A., Kloog, I., Schwartz, J., Silibello, C., Michelozzi, P., Gariazzo, C., Viegi, G., Forastiere, F., Karnieli, A., Just, A.C., Stafoggia, M., 2019. Estimating Daily PM_{2.5} and PM₁₀ over Italy Using an Ensemble Model. *Environmental Science and Technology*. <https://doi.org/10.1021/acs.est.9b04279>.
- Silibello, C., Carlino, G., Stafoggia, M., Gariazzo, C., Finardi, S., Pepe, N., Radice, P., Forastiere, F., Viegi, G., 2021. Spatial-temporal prediction of ambient nitrogen dioxide and ozone levels over Italy using a Random Forest model for population exposure assessment. *Air Qual. Atmos. Health* 14 (6), 817–829. <https://doi.org/10.1007/s11869-021-00981-4>.
- Stafoggia, M., Johansson, C., Glantz, P., Renzi, M., Shtein, A., Hoogh, K. de, Kloog, I., Davoli, M., Michelozzi, P., Bellander, T., 2020. A random forest approach to estimate daily particulate matter, nitrogen dioxide, and ozone at fine spatial resolution in Sweden. *Atmosphere* 11 (3), 239. <https://doi.org/10.3390/atmos11030239>.
- Stafoggia, M., Schwartz, J., Badaloni, C., Bellander, T., Alessandrini, E., Cattani, G., de Donato, F., Gaeta, A., Leone, G., Lyapustin, A., Sorek-Hamer, M., de Hoogh, K., Di, Q., Forastiere, F., Kloog, I., 2017. Estimation of daily PM₁₀ concentrations in Italy (2006–2012) using finely resolved satellite data, land use variables and meteorology. *Environ. Int.* 99, 234–244. <https://doi.org/10.1016/j.envint.2016.11.024>.
- Stas, M., Aerts, R., Hendrickx, M., Delcloo, A., Dendoncker, N., Dujardin, S., Linard, C., Nawrot, T., Van Nieuwenhuysse, A., Aerts, J.M., Van Orshoven, J., Somers, B., 2021. Exposure to green space and pollen allergy symptom severity: a case-crossover study in Belgium. *Sci. Total Environ.* 781, 146682 <https://doi.org/10.1016/j.scitotenv.2021.146682>.
- Tang, R., Blangiardo, M., Gulliver, J., 2013. Using building heights and street configuration to enhance intraurban PM₁₀, NO_x, and NO₂ land use regression models. *Environ. Sci. Technol.* 47 (20), 11643–11650. <https://doi.org/10.1021/es402156g>.
- van Nunen, E., Vermeulen, R., Tsai, M.-Y., Probst-Hensch, N., Neichen, A., Davey, M., Imboden, M., Ducret-Stich, R., Naccarati, A., Raffaele, D., Ranzi, A., Ivaldi, C., Galassi, C., Nieuwenhuijsen, M., Curto, A., Donaire-Gonzalez, D., Cirach, M., Chatzi, L., Kampouri, M., et al., 2017. Land use regression models for ultrafine particles in six European areas. *Environ. Sci. Technol.* 51 (6), 3336–3345. <https://doi.org/10.1021/acs.est.6b05920>.
- Vienneau, D., De Hoogh, K., Bechle, M.J., Beelen, R., Van Donkelaar, A., Martin, R.V., Millet, D.B., Hoek, G., Marshall, J.D., 2013. Western European land use regression incorporating satellite- and ground-based measurements of NO₂ and PM₁₀. *Environ. Sci. Technol.* 47 (23), 13555–13564. <https://doi.org/10.1021/es403089q>.
- Wood, S., 2018. Package “Mgcv” Title Mixed GAM Computation Vehicle with Automatic Smoothness Estimation.
- Yang, Y., Lin, Q., Liang, Y., Ruan, Z., Acharya, B.K., Zhang, S., Qian, Z., McMillin, S.E., Hinyard, L., Sun, J., Wang, C., Ge, H., Wu, X., Guo, X., Lin, H., 2020. Maternal air pollution exposure associated with risk of congenital heart defect in pre-pregnancy overweighted women. *Sci. Total Environ.* 712, 136470 <https://doi.org/10.1016/j.scitotenv.2019.136470>.
- Yanosky, J.D., Paciorek, C.J., Laden, F., Hart, J.E., Puett, R.C., Liao, D., Suh, H.H., 2014. Spatio-temporal modeling of particulate air pollution in the conterminous United States using geographic and meteorological predictors. *Environ. Health: A Global Access Sci. Sou.* 13 (1) <https://doi.org/10.1186/1476-069X-13-63>.
- Yu, W., Li, S., Ye, T., Xu, R., Song, J., Guo, Y., 2022. Deep ensemble machine learning framework for the estimation of PM_{2.5} concentrations. *Environ. Health Perspect.* 130 (3) <https://doi.org/10.1289/EHP9752>.
- Zhang, Z., Wang, J., Hart, J.E., Laden, F., Zhao, C., Li, T., Zheng, P., Li, D., Ye, Z., Chen, K., 2018. National scale spatiotemporal land-use regression model for PM_{2.5}, PM₁₀ and NO₂ concentration in China. *Atmos. Environ.* 192, 48–54. <https://doi.org/10.1016/j.atmosenv.2018.08.046>.