



Article

Development and Evaluation of Spatio-Temporal Air Pollution Exposure Models and Their Combinations in the Greater London Area, UK

Konstantina Dimakopoulou ¹, Evangelia Samoli ¹, Antonis Analitis ¹, Joel Schwartz ^{2,3} , Sean Beevers ⁴, Nutthida Kitwiroon ⁴, Andrew Beddows ⁴, Benjamin Barratt ^{4,5} , Sophia Rodopoulou ¹, Sofia Zafeiratou ¹, John Gulliver ⁶ and Klea Katsouyanni ^{1,4,*}

- ¹ Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, 115 27 Athens, Greece; kdimakop@med.uoa.gr (K.D.); esamoli@med.uoa.gr (E.S.); aanalit@med.uoa.gr (A.A.); srodopoyl@med.uoa.gr (S.R.); sofiazaf33@gmail.com (S.Z.)
- ² Department of Environmental Health, Harvard TH Chan School of Public Health, Boston, MA 02115, USA; jschwartz@hsph.harvard.edu
- ³ Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA 02115, USA
- ⁴ MRC Centre for Environment and Health, Imperial College London, London SE1 9NH, UK; s.beevers@imperial.ac.uk (S.B.); nutthida.kitwiroon@kcl.ac.uk (N.K.); andrew.beddows@kcl.ac.uk (A.B.); b.barratt@imperial.ac.uk (B.B.)
- ⁵ National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Health Impact of Environmental Hazards, Imperial College London, London SW7 2AZ, UK
- ⁶ Centre for Environmental Health and Sustainability, School of Geography, Geology and the Environment, University of Leicester, University Road, Leicester LE1 7RH, UK; j.gulliver@imperial.ac.uk
- * Correspondence: kkatsouy@med.uoa.gr



Citation: Dimakopoulou, K.; Samoli, E.; Analitis, A.; Schwartz, J.; Beevers, S.; Kitwiroon, N.; Beddows, A.; Barratt, B.; Rodopoulou, S.; Zafeiratou, S.; et al. Development and Evaluation of Spatio-Temporal Air Pollution Exposure Models and Their Combinations in the Greater London Area, UK. *Int. J. Environ. Res. Public Health* **2022**, *19*, 5401. <https://doi.org/10.3390/ijerph19095401>

Academic Editor: Paul B. Tchounwou

Received: 31 March 2022

Accepted: 27 April 2022

Published: 28 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Land use regression (LUR) and dispersion/chemical transport models (D/CTMs) are frequently applied to predict exposure to air pollution concentrations at a fine scale for use in epidemiological studies. Moreover, the use of satellite aerosol optical depth data has been a key predictor especially for particulate matter pollution and when studying large populations. Within the STEAM project we present a hybrid spatio-temporal modeling framework by (a) incorporating predictions from dispersion modeling of nitrogen dioxide (NO₂), ozone (O₃) and particulate matter with an aerodynamic diameter equal or less than 10 µm (PM₁₀) and less than 2.5 µm (PM_{2.5}) into a spatio-temporal LUR model; and (b) combining the predictions LUR and dispersion modeling and additionally, only for PM_{2.5}, from an ensemble machine learning approach using a generalized additive model (GAM). We used air pollution measurements from 2009 to 2013 from 62 fixed monitoring sites for O₃, 115 for particles and up to 130 for NO₂, obtained from the dense network in the Greater London Area, UK. We assessed all models following a 10-fold cross validation (10-fold CV) procedure. The hybrid models performed better compared to separate LUR models. Incorporation of the dispersion estimates in the LUR models as a predictor, improved the LUR model fit: CV-R² increased to 0.76 from 0.71 for NO₂, to 0.79 from 0.57 for PM₁₀, to 0.81 to 0.66 for PM_{2.5} and to 0.75 from 0.62 for O₃. The CV-R² obtained from the hybrid GAM framework was also increased compared to separate LUR models (CV-R² = 0.80 for NO₂, 0.76 for PM₁₀, 0.79 for PM_{2.5} and 0.75 for O₃). Our study supports the combined use of different air pollution exposure assessment methods in a single modeling framework to improve the accuracy of spatio-temporal predictions for subsequent use in epidemiological studies.

Keywords: air pollution; exposure modeling; land use regression; chemical transport models; machine learning; particulate matter

1. Introduction

Epidemiological studies have been utilizing various air pollution exposure assessment methods to associate individualized exposure to air pollution concentrations and health

effects [1–3]. Early studies [4–6] of long-term air pollution exposure assigned the concentration measurement from the nearest fixed monitoring site or applied interpolation methods (e.g., inverse distance weighing and kriging) to a geographical point of interest (usually at the participant’s residential address). A limitation of this approach is its limited ability to capture temporal variation in concentrations. More recently, land use regression (LUR) and dispersion/chemical transport models (D/CTMs) are among the most common methods applied to predict air pollution concentrations at a fine spatial scale for use in epidemiological studies [1,2,7–9]. D/CTM models can capture temporal variation, but are limited by the accuracy of the emissions inventories. Moreover, satellite aerosol optical depth (AOD) data have been frequently used for predicting fine particulate matter pollution [3,10], especially in the lack of fixed site monitoring data and in the need of analysis of nation-wide/large population studies.

These more recent methods have been shown to be useful tools [11] for different epidemiological study designs since they can be extended to predict both the spatial and temporal variability of air pollution concentrations. For example, in time series and panel studies they can provide daily predictions, while in studies assessing long-term exposure health effects, the daily predictions obtained can be averaged over the time period of interest. In addition, these developments in exposure modeling provide spatially resolved daily estimates enabling an integrated assessment of health effects arising from both long-and-short-term exposures. However, they also have some limitations. LUR model development relies on air pollution measurements provided from a fixed monitoring network or is based upon specifically designed monitoring campaigns within a study [12]. As a result, a spatially sparse monitoring network or the limited temporal coverage of specifically designed monitoring campaigns may increase exposure measurement error. D/CTMs are based on the description of the physicochemical processes of air pollution, involving pollutant emissions sources and its precursors [13]. Therefore, they require high quality input data to produce accurate predictions. AOD data have contributed to developing national models with a spatial resolution of 1×1 km for particulate matter with an aerodynamic diameter equal or less than $2.5 \mu\text{m}$ (PM_{2.5}) [14], but, as they measure PM within a height of several kilometers above ground, modelling is required to estimate concentrations at the height of the human breathing zone [10,11]. Additionally, values are missing on days with cloud coverage, which may be a significant problem for certain geographical areas and seasons. Therefore, models based only on AOD data may have increased uncertainty and may not allow the adequate assessment of intra-city variations.

Air pollution is a major environmental risk factor for human health [15,16] and it is crucial to provide epidemiological studies with accurate estimates of exposure. To overcome the limitations of single exposure assessment methods and to improve the accuracy of predictions, recent studies are combining methodologies or outputs from different methods. Incorporation of predictions from D/CTMs and/or satellite-derived air pollutant concentrations as predictor variables within a LUR model have been shown to improve the model performance in terms of predicting the spatial variability of air pollution [17,18]. On the other hand, very few studies have developed hybrid models by incorporating the output of different exposure assessment methods into a single spatio-temporal (ST) modeling framework [18–20].

Within the “Comparative evaluation of Spatio-Temporal Exposure Assessment Methods for estimating health effects of air pollution” (STEAM) project we developed two hybrid ST modeling approaches for air pollutants, by combining LUR, dispersion and machine learning modeling, in the Greater London Area for the years 2009–2013. In the first approach, we incorporate the predictions from dispersion modeling of nitrogen dioxide NO₂, ozone (O₃; trioxygen) and particulate matter with an aerodynamic diameter equal or less than $10 \mu\text{m}$ (PM₁₀) and PM_{2.5} into a ST LUR model framework. In the second approach, we apply an ST-generalized additive model (GAM) combination of the predictions of individual models. For NO₂, O₃ and PM₁₀ this is carried out by including the predictions of the two methods (LUR and CTM modeling), while for PM_{2.5} by including the predictions of three methods

(LUR, CTM and an ensemble model using machine learning and models informed also by satellite data) as independent variables in the GAM.

2. Methods

2.1. Study Area

The Greater London study area has a total of 9,784,200 inhabitants (Census 2011 data; <https://data.london.gov.uk/dataset/2011-census-demography>, accessed on 1 March 2017) across 5373 census-based aggregation units, the Lower Layer Super Output Areas (LSOAs), of which the centroids (longitude, latitude) are located within the London Orbital Motorway (M25). Each LSOA within the study area has a minimum population of 1000 persons with an average population in 2010 of 1722 [21]. Within the same area there are 219,093 post codes which also have defined centroids which are population weighted. The Greater London Area has an exceptionally extensive air pollution monitoring network which allows testing the performance of various models as will be described in the following sections.

2.2. Air Pollution Monitoring Data and Enhanced PM2.5 Database

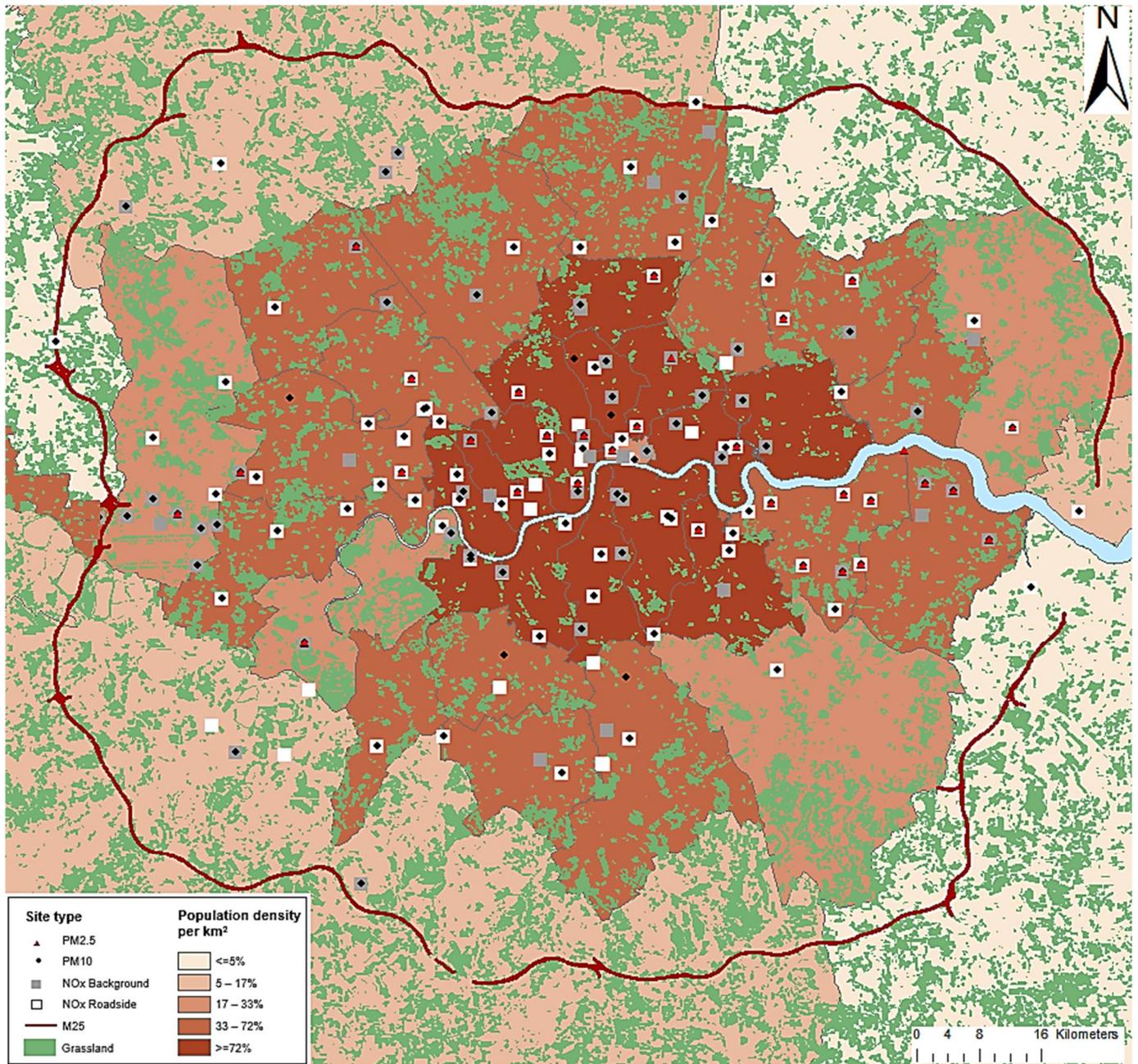
We obtained daily measured concentrations of NO₂, O₃, PM10 and PM2.5 for 2009 through 2013 from the London Air Quality Network [22], the Air Quality England [23] and from the Automatic Urban and Rural Network [24]. For NO₂, PM10 and PM2.5 we formulated a database with the 24 h average measured pollutant concentration (µg/m³), while for O₃ we calculated the daily maximum 8 h (8 h max) average concentration (µg/m³). NO₂ measurements were available from 130 monitoring sites, PM10 from 115 sites and PM2.5 from 33 sites. In order to enhance the representation of PM2.5 monitoring sites in the study area, we combined a regression model and a random forest (RF) approach to predict PM2.5 concentrations at fixed sites with PM10 measurements (but without PM2.5 measurements). More details on the methods applied for PM2.5 can be found in [25]. This procedure was essential for informing the ST LUR and machine learning PM2.5 modeling development with a sufficient number of monitoring locations (n = 104). For O₃, measurements from 62 sites were used located in an extended area including 9688 LSOAs. O₃ is a secondary pollutant whose levels are mostly higher in rural areas than in urban settings. Therefore, we extended our study area to account for its formation properties and transport patterns.

Figure 1 shows the study area and the geographical location of the monitoring sites.

2.3. Independent Exposure Assessment Methods

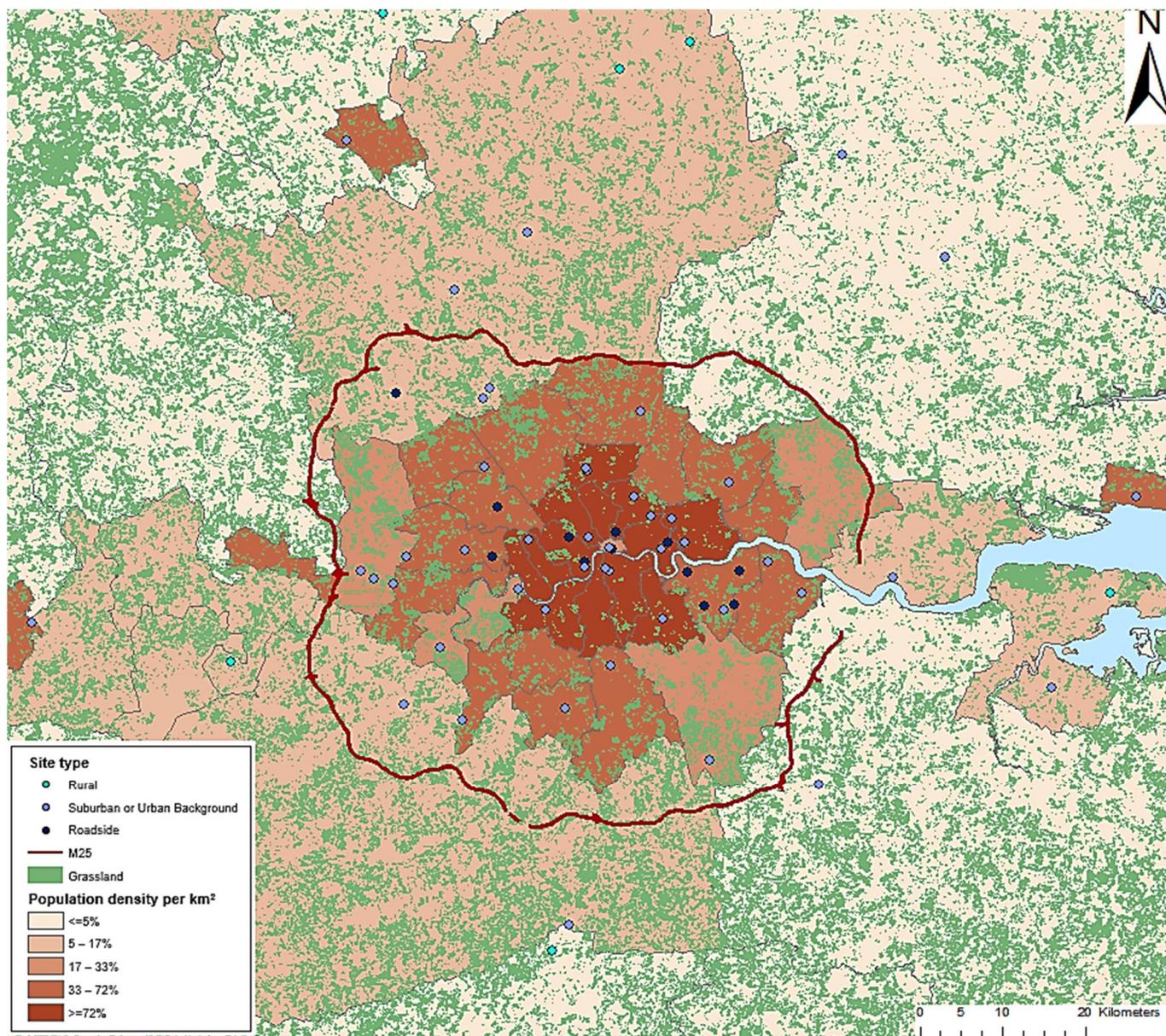
2.3.1. Brief Summary of the First Stage Exposure Assessment

At the first stage ST LUR, dispersion and additionally machine learning models for PM2.5 were developed. The outputs of these approaches were combined under different hybrid model developed as the second stage. “Temporal” refers to daily (24 h) variation, while “spatial” refers to variation within the study area at the coordinates of interest (LSOA centroid).



(a)

Figure 1. Cont.



(b)

Figure 1. Map of the geographical location of the fixed monitoring network operated at the Greater London (a) study area for NO₂, PM₁₀ and PM_{2.5} and (b) extended study area for O₃.

2.3.2. LUR Models

We developed semi-parametric ST LUR models to predict daily NO₂, O₃, PM₁₀ and PM_{2.5} concentrations at any point of interest in the study area, for the years 2009 to 2013. Similar to the approach previously applied in Athens, Greece [26,27], the log-transformed air pollutant measurements (except for O₃ as it was normally distributed) at location i on day t was modeled using a set of covariates that had either a linear or a smooth effect on the pollutant; plus a bivariate smooth function of the fixed monitoring sites' geographical location (longitude, latitude) in order to account for remaining spatial correlation. We used available air pollution measurement data, while 97 potential predictors of air pollutants' spatio-temporal variability were tested. These variables can be classified in four categories: (a) land use type (Land Cover Map—LCM of Great Britain from 2007) in a buffer range of 100 to 5000 m (m) around each fixed monitoring site (precisely, 100, 300, 500, 1000, and 5000 m); (b) meteorological variables, obtained from the UK Meteorological

Office (daily mean): temperature ($^{\circ}\text{C}$), relative humidity (%), wind direction ($^{\circ}\text{N}$), wind speed (m/s), cloud coverage (okta), barometric pressure (mBar/hPa) and solar radiation (W/m^2); (c) traffic-related: total length of major roads (m) in buffers of 25, 50, 100, 300, 500 and 1000 m, inverse distance of the fixed monitoring sites to the nearest major road (m^{-1}), and traffic intensity on the nearest major road (veh day^{-1}) to the fixed monitoring site and total traffic load within each buffer of 25 to 1000 m ($\text{veh day}^{-1} * \text{m}$). Traffic counts were obtained from the Department of Transport in the United Kingdom; (d) indicators: of linear trend within a year (a day count variable accounting for trends within each year coded from 1 to 365 or 366), of linear trend over the years of the study period (4 dummy variables with year 2009 as reference category) and the day of the week (6 dummy variables with Sunday as the reference category). For the smooth function we used a penalized spline with degrees of freedom (df) estimated via Restricted Maximum Likelihood (REML). Regarding all continuous variables, the final model included the term (linear or smoothed) that provided the better model fit. The final set of explanatory variables was selected based on the model's adjusted- R^2 value. Apart from the adjusted- R^2 value, the coefficient of spatial covariates had to conform to the pre-defined direction of effect. The addition of variables was continued until none added more than 1% to the value of adjusted- R^2 . The final variables included in each of the ST LUR models are shown in Figure 2. All the predictor variables of the spatial variation of O_3 , PM10 and PM2.5 concentrations are traffic related. For NO_2 , the spatial variables included in addition to traffic-related variables, also a land use type (area characterized as urban within a buffer of 300 m) variable.

Regarding temporal variables, all final ST LUR models included meteorological and indicator variables for study years. Moreover, all final models included the geography (longitude, latitude) of the fixed monitoring sites. We validated our developed models using ten-fold cross-validation (10-fold CV). All land use types and traffic-related variables were extracted by conducting GIS analysis via ArcGIS Desktop, Release 10 [28]. All statistical analysis was conducted using the R statistical software (version 3.3.3; R Core Team, 2017, sourced from Athens, Greece) [29] and the R library "SemiPar" version 1.0-2 [30].

2.3.3. Dispersion Models

We used a modeling system that combines the anthropogenic and natural emissions data, with the Weather Researching and Forecasting (WRF) meteorological model [31] and the Community Multiscale Air Quality (CMAQ) model [32], which has been coupled to the Atmospheric Dispersion Modelling System (ADMS) roads model [33,34] to predict hourly NO_2 , O_3 , PM10 and PM2.5 spatially at a 20 m grid resolution over the study area, for the time period 2009 to 2013. The anthropogenic emissions data were obtained by combining the UK National Atmospheric Emissions Inventory (NAEI) [35], the London Atmospheric Emissions Inventory [36], King's road transport emissions model [37] and EMEP European emissions (<https://www.ceip.at/>, accessed on 15 June 2017). The biogenic emissions from vegetation and soils were estimated using the Biogenic Emission Inventory System version 3 (BEIS3) model [38]. Sea-salt emissions were calculated in-line in CMAQ. Bias in the 2×2 km CMAQ PM2.5 and PM10 hourly output was corrected using a sample of background sites before the local scale dispersion modelling stage. The discrepancies between the model output and the measurements at a random sample of 50% of background sites in the case of PM10, and 5 sites in the case of PM2.5, was interpolated onto the 2×2 km grid to create a correction surface. This interpolation was carried out using two iterations of a multilevel B-spline algorithm [39], which normally takes around eight iterations to interpolate points exactly, so that the resultant error surface provided smoothly varying bias correction across the domain, rather than fixing the model output to the measurements. The results from CMAQ-urban model were evaluated at 152 fixed monitoring stations from the UK and London monitoring networks, using methods described in the UK Department of the Environment, Food and Rural Affairs (DEFRA) model evaluation protocol [40].

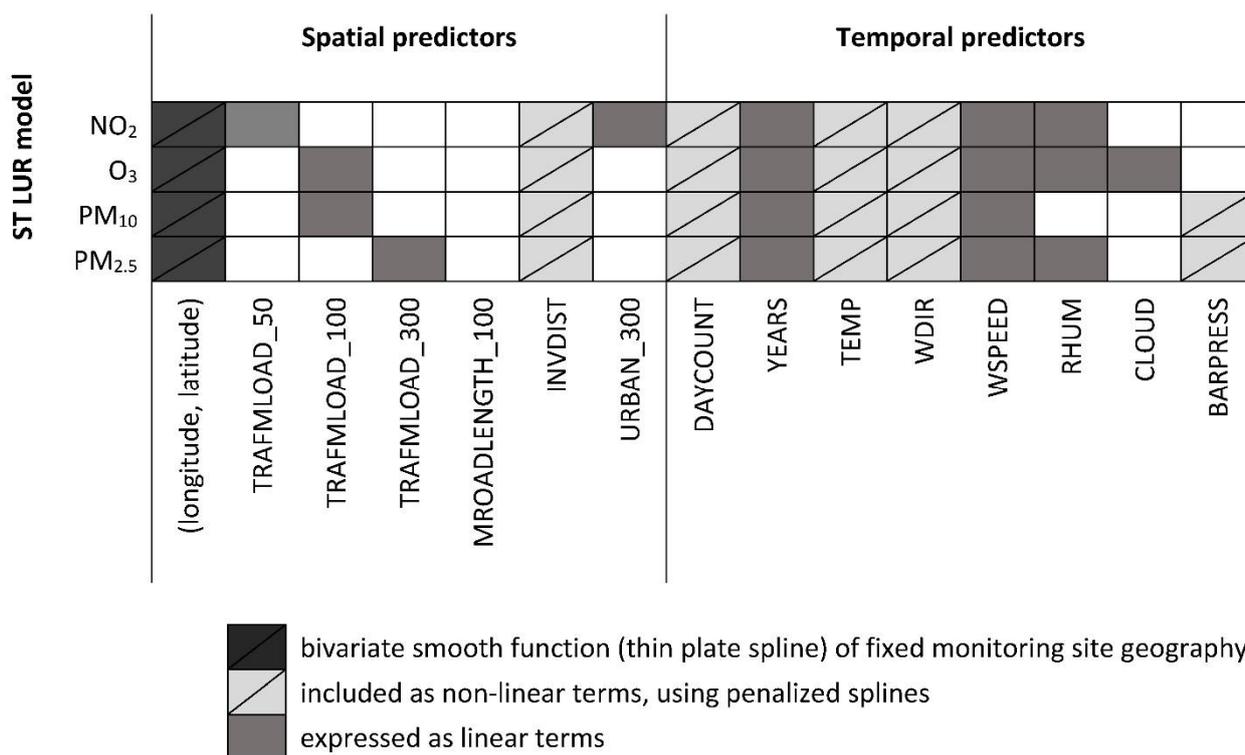


Figure 2. Predictor variables included in the final spatio-temporal land use regression (ST LUR) models developed for NO₂, O₃, PM₁₀ and PM_{2.5} ($\mu\text{g}/\text{m}^3$), in the Greater London Area for the years 2009 to 2013. TRAFMLOAD_50; TRAFMLOAD_100; TRAFMLOAD_300: traffic load of major roads (veh*m/day) in a buffer of 50m, 100m and 300m around each fixed monitoring site, respectively; MROADLENGTH_100: total length of major roads (m) in a buffer of 100m around each fixed monitoring site; INVDIST: inverse distance of fixed monitoring sites to the nearest major road (m^{-1}); URBAN_300: urban areas (m^2) in a buffer of 300m around each fixed monitoring site; DAYCOUNT: day count variable accounting for trends within each year coded from 1 to 365 or 366 (included penalized splines with 6 degrees of freedom (df)); YEARS: years of the study period (4 dummy variables with year 2009 as reference category); TEMP: daily mean temperature ($^{\circ}\text{C}$, included penalized splines with 3 df); WDIR: daily mean wind direction ($^{\circ}\text{N}$, included penalized splines with 3 df); WSPEED: daily mean wind speed (m/s); RHUM: daily mean relative humidity (%); CLOUD: daily mean cloud coverage (okta); BARPRESS: daily mean barometric pressure (mBar/hPa, included penalized splines with 3 df).

2.3.4. PM_{2.5} Prediction Model Based on an Ensemble Machine Learning ST Approach

We applied an ensemble machine learning approach including AOD, land use and meteorological data in order to predict daily PM_{2.5} concentrations in the study area, on a 1km × 1km scale (consisting in a total of 3960 grid cells). Details on the prediction model development can be found in [41]. In brief, the machine learners used in the process were the gradient boosting machine (GBM), the random forest (RF) and the k-nearest neighbor (KNN). AOD data were provided by the MAIAC algorithm for Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on the Aqua and Terra satellites [42]. Predictors of the spatial variation of PM_{2.5} were: population density (persons/km²), land use type (LCM 2007), distance to water (km), distance to Heathrow airport (m), normalized difference vegetation index (NDVI), traffic counts, average daily PM_{2.5} across the greater London area ($\mu\text{g}/\text{m}^3$), light at night, elevation (m), distance to nearest major road (km), distance to nearest bus stop (km), average building height (m) and length of major road (km), number of bus stops and number of buildings, in the grid cell. The included meteorological covariates were: cloudiness (okta), barometric pressure (mBar/hPa), wind direction ($^{\circ}\text{N}$), wind speed (m/s), dew point temperature ($^{\circ}\text{C}$), temperature ($^{\circ}\text{C}$) and

inverse of the height of the planetary boundary layer (m^{-1}). Additionally, variables on the temporal scale were included to account for seasonal variations of PM_{2.5} (sine of day of the year, cosine of day of the year and day of week) and to account for long-term trends (number of days from time of origin and year). Model training was based using a grid search to optimize the hyper-parameters for the algorithms, and by taking into account the mean square error (MSE) and cross-validated R^2 values as selection criteria. Following, we obtained the final ensemble-averaged PM_{2.5} predictions from a GAM, with independent variables a smoothed function (using a penalized spline with degrees of freedom estimated via REML) of the predictions obtained from each machine learning methods and by including a bivariate smooth function of latitude and longitude. Ten-fold CV was applied to evaluate model performance.

2.3.5. Agreement between First Stage (Independent) Exposure Assessment Methods

Lin's concordance correlation coefficient was calculated as a measure of agreement, at a temporal and spatial level, between the independent exposure assessment methods [43]. Agreement at the spatial level was investigated by the comparison of annual means of estimates (from the 3 methods) for each Lower Layer Super Output Area (LSOA), centroid. Agreement at the temporal level was investigated by taking into account the daily estimates, over the whole study area. Moreover, Bland–Altman method was applied in order to evaluate the mean differences and to estimate an 95% agreement interval of the differences (LoA) between the independent exposure assessment methods (ST LUR models, Dispersion models and the ensemble machine learning approach [44]).

2.4. Spatio-Temporal Air Pollutant Modeling and Validation within a Hybrid Framework

2.4.1. Brief Summary of the Second Stage Exposure Assessment

At the second stage we combined estimates from the first-stage approaches either by incorporating the dispersion model output in the LUR or by combining pollutant-specific output in a GAM mode and applying a smooth term per first stage approach estimate.

2.4.2. Incorporation of Predictions from Dispersion Models and from an Ensemble Machine Learning Approach in the Case of PM_{2.5} within the LUR Model

Depending on the availability of the independent exposure assessment approaches, we developed hybrid models by incorporating estimates derived from dispersion modeling and the prediction model for PM_{2.5} from an ensemble machine learning approach, of the following form:

$$f_p(\text{poll}_{it}) = W_{it}^T \beta + \sum_{l=1}^q f_l(s_{l,it}) + h(\text{geog}_{ij}) + \sum_{k=1}^n g_k f_l(M_{l,it}) + \varepsilon_{it} \quad (1)$$

where poll_{it} is the air pollutant concentration measurement at location i on day t , $f_l(\cdot)$ $l = 1, 2, \dots, q$ is a smooth function reflecting the non-linear effect of covariate $S_{l,it}$ on the pollutant's concentration poll_{it} , $S_{l,it}$ stands for the l th smoothed covariate, h is a bivariate smooth function of the fixed monitoring sites geographical coordinates $\text{geog}_{ij} = (\text{longitude}_i, \text{latitude}_j)$ that account for residual correlation between locations i and j , W_{it} is the vector of covariates that have a linear effect on poll_{it} and β is the corresponding vector of regression coefficients. M_{it} is a ST smooth function of the predictions from the $k =$ dispersion model and the additional prediction model for PM_{2.5}, with coefficient g . For $g_k = 0$ model (1) is equivalent to the non-integrated ST LUR model. The errors (ε_{it}) are assumed to be independent and normally distributed, with a mean value of zero and a constant variance (σ_ε^2). For NO₂, PM₁₀ and PM_{2.5} the air pollutant concentration measurement was log-transformed. We did not log-transform the O₃ measurements since they were normally distributed. Therefore, for NO₂, PM₁₀ and PM_{2.5}: $f_p(\text{poll}_{it}) = \log(\text{poll}_{it})$, while for O₃: $f_p(\text{poll}_{it}) = \text{poll}_{it}$. Based on the above, the specific models constructed were: A hybrid ST LUR framework by incorporating predicted concentrations of NO₂, O₃ and PM₁₀ from dispersion modeling and by incorporating predicted concentrations of PM_{2.5} from the dispersion the ensemble machine learning

approach, as covariates within the ST LUR. These models are thereafter referred to as “Hybrid 1” and “ensemble”, respectively.

2.4.3. Combination of Predictions Derived from LUR, Dispersion and for PM2.5 also Based on an Ensemble Machine Learning Modeling within a GAM

Similarly, depending on the availability of independent exposure assessment approaches we constructed a GAM, by fitting a smooth function (penalized splines) of the predictions from each method (LUR, dispersion and the prediction model based on an ensemble machine learning approach). The predictions reflect the daily estimated air pollutant concentrations at the fixed monitoring sites located in the study area. Since both the LUR model and the prediction model based on the machine learning use the air pollutant measurements during the model development procedure (dependent variable), the 10-fold CV predictions obtained from each method were used in the GAM. Our GAM has the following form:

$$\begin{aligned} poll_{it} = & s\left(LUR\ pred_{10fold\ CV}\cdot poll_{it}\right) + s(dispersion\ pred.\ poll_{it}) + \\ & + s\left(ML\ pred_{10fold\ CV}\cdot poll_{it}\right) + \epsilon_{ij} \end{aligned} \quad (2)$$

where $poll_{it}$ is the air pollutant concentration measurement at the fixed monitoring site location i on day t , s is a penalized splines with the degree of smoothness based on the generalized cross validation criterion—GCV), $LUR\ pred_{10fold\ CV}\cdot poll_{it}$ are the 10-fold CV predictions of the air pollutant concentrations obtained from the ST LUR models, $dispersion\ pred.\ poll_{it}$ are the estimated air pollutant concentrations obtained from dispersion modeling and $ML\ pred_{10fold\ CV}\cdot poll_{it}$ are the 10-fold CV predictions of the air pollutant (available only for PM2.5) concentrations obtained from the ensemble machine learning approach, on location i on day t . The final output from the GAM framework is the weighted-average daily predictions of air pollutants. The weighing of single methods was carried out using the smoothed function which allows each method to vary along the concentration range of the pollutants in case one method performs better than another in a specific concentration range. The specific models constructed were: a hybrid GAM by combining predicted concentrations of NO₂, O₃, PM10 and PM2.5 from the LUR, dispersion and PM2.5 prediction model based on an ensemble machine learning approach—hereafter, “hybrid 2”.

2.4.4. Validation

The hybrid 1 and 2 models’ performance was evaluated using a 10-fold CV method. In this method, all models were fitted to N-10% fixed monitoring sites and the predicted concentrations were compared with the measured (observed) concentrations at the left-out sites. This procedure was repeated 10 times. Finally, the overall level of fit (based on the model’s R² value) between the predicted and observed concentrations, across all sites, was calculated as a measure of our hybrid models’ performance. Moreover, we separately investigated the temporal and spatial validity of the developed hybrid models. To assess each model’s temporal validity, we regressed the difference between the daily predicted and the mean annual predicted pollutant concentrations over the difference between measured pollutant concentrations and their mean annual levels. In order to assess spatial validity, we regressed the mean annual predicted concentrations of each pollutant over the mean annual measured concentrations at the fixed monitoring sites. The level of temporal and spatial fit was evaluated by obtaining the models’ 10-fold CV R² value.

2.4.5. Application

Subsequently, the hybrid models were used to predict daily concentrations of NO₂, O₃, PM10 and PM2.5 concentrations per LSOA, by averaging the predictions in all post code centroids located within the LSOA. Thus, for the 5373 LSOAs included in the Greater London study Area we predicted pollutant concentrations for 219,093 postcode centroids.

3. Results

Table 1 presents the distribution of the estimated NO₂, O₃, PM10 and PM2.5 concentrations predicted from the first-stage (independent) and the ensemble/hybrid exposure assessment methods. The data shown are long-term (years 2009 to 2013) predicted concentrations of pollutants per LSOA, after averaging the predictions in all post code centroids within the area of the LSOA. The range of the number of postcode centroids per LSOA was 1 to 1585, with a median value of 32 postcodes. The ST- LUR predicted pollutant concentrations are slightly underestimated compared to measurements from fixed sites. Similarly regarding NO₂ and PM10 predictions obtained from dispersion modeling.

Table 1. Distribution of air pollutant concentrations measured from the fixed monitoring network and of estimated concentrations by the independent and hybrid exposure assessment methods, by pollutant, in the Greater London Area (n = 5373 LSOAs) for the years 2009–2013.

Concentrations		Pollutant Measurements (µg/m ³)			
		NO ₂ (24 h Average)	O ₃ (8 h Max)	PM10 (24 h Average)	PM2.5 (24 h Average)
Fixed monitoring network	# sites	130	62	115	104
		52.1 (24.39)	53.2 (11.62)	24.2 (5.18)	14.5 (2.48)
		46.9 (36.7–57.3)	54.5 (47.0–62.1)	23.6 (20.5–27.7)	14.3 (12.9–16.1)
		Estimations of Pollutant Concentrations (µg/m ³)			
		Mean (SD) and Median (25th–75th p)			
Independent exposure assessment methods	LUR ¹	41.4 (20.77)	50.8 (8.23)	20.3 (4.01)	12.9 (2.12)
		38.2 (28.2–75.5)	51.9 (45.8–56.3)	19.7 (17.6–22.4)	11.9 (9.2–15.5)
	Dispersion ²	37.7 (12.76)	59.8 (6.66)	19.7 (2.56)	14.2 (1.80)
		35.8 (29.4–43.0)	60.9 (56.0–64.9)	19.1 (17.9–20.9)	13.8 (12.9–15.3)
	Ensemble ³	-	-	-	15.8 (1.28)
					15.7 (15.0–16.5)
Hybrid exposure assessment methods	Hybrid 1 ⁴	43.8 (34.31)	57.7 (12.77)	21.9 (8.67)	14.1 (2.47)
		35.1 (25.5–48.7)	54.1 (49.2–58.8)	20.0 (17.5–23.6)	13.7 (12.3–15.3)
	Hybrid 2 ⁵	35.2 (13.47)	51.7 (7.27)	19.5 (2.29)	14.4 (1.18)
		33.6 (26.7–41.5)	52.6 (47.2–56.8)	18.9 (17.9–20.5)	14.2 (13.6–15.1)

¹ spatio-temporal Land Use Regression (LUR) model. ² spatio-temporal dispersion model. ³ PM2.5 prediction model based on an ensemble machine learning spatio-temporal approach. ⁴ Incorporation of estimates derived from independent exposure assessment methods, within the LUR model. ⁵ Combination of estimates derived from independent exposure assessment methods, within a GAM.

Table 2 summarizes the agreement between the independent exposure assessment methods. Regarding NO₂, O₃ and PM10 we assessed the spatial and temporal agreement by comparing the predictions obtained from LUR and dispersion models, while for PM2.5 we compared per 2 methods at a time. The spatial agreement between LUR and dispersion models is moderate. The temporal agreement is better than the spatial. The pollutant which displays the highest spatial agreement between exposure assessment methods is NO₂. The highest temporal agreement was observed for PM2.5, between the dispersion and ensemble machine learning approach. However, the mean difference is larger for NO₂ and O₃ compared to PM.

Table 2. Summary of agreement between the independent exposure assessment methods.

Agreement	Pollutant ($\mu\text{g}/\text{m}^3$)					
	NO ₂ (24 h Average)		O ₃ (8 h Max)		PM10 (24 h Average)	
	Spatial	Temporal	Spatial	Temporal	Spatial	Temporal
Mean difference ¹ (95% LoA)	−3.7 (−40.9, 33.4)	−3.7 (−18.6, 11.1)	9.0 (−6.7, 24.8)	9.0 (−18.5, 36.7)	−0.6 (−8.4, 7.1)	−0.6 (−12.0, 10.8)
Lin's	0.39 *	0.78 *	0.25 *	0.61 *	0.31 *	0.69 *
r	0.45 *	0.82 *	0.44 *	0.71 *	0.37 *	0.75 *
	PM2.5 (24 h Average; $\mu\text{g}/\text{m}^3$)					
	^a difference		^b difference		^c difference	
	Spatial	Temporal	Spatial	Temporal	Spatial	Temporal
Mean difference ² (95% LoA)	0.4 (−4.4, 5.3)	0.4 (−4.5, 12.4)	−1.6 (−5.0, 1.8)	−1.6 (−5.2, 2.0)	−2.0 (−6.7, 2.7)	−2.0 (−11.9, 7.9)
Lin's	0.26 *	0.67 *	0.24 *	0.96 *	0.12 *	0.71 *
r	0.28 *	0.77 *	0.39 *	0.98 *	0.22 *	0.81 *

Mean difference¹: Applicable to NO₂, O₃ and PM10; difference = Dispersion—LUR predicted concentrations. Mean difference²: Applicable only to PM2.5; ^a difference = Dispersion—LUR predicted concentrations; ^b difference = Dispersion—ensemble machine learning approach model predicted concentrations; ^c difference = LUR—ensemble machine learning approach model predicted concentrations. Lin: Lin's concordance correlation coefficient. r: Pearson correlation coefficient. Ninety-five percent limits of agreement (LoA) Bland–Altman method. * *p*-value < 0.05.

Table 3 presents the performance of the independent and the hybrid modeling approaches for NO₂, O₃, PM10 and PM2.5 concentrations at the fixed air pollution monitoring sites. The ST LUR model for NO₂ performed well (CV-R²: 0.71) with better ability to predict spatially than over time (i.e., for NO₂ spatial R²: 0.67 and temporal R²: 0.33), while for O₃ and particulate matter, the ST LUR model performed moderately well. The dispersion compared to the LUR model performed similarly regarding the prediction of nitrogen dioxides and ozone and better regarding the prediction of particulate matter. Both hybrid modeling approaches (hybrid model 1 and 2) outperformed the independent models and improved the accuracy of predictions for all pollutants in terms of RMSE. Incorporating the estimates derived from dispersion models and the ensemble machine learning approach (only for PM2.5) into the LUR model (hybrid model 1), resulted in an increase in the CV-R² value by 5% to 22%. The combination of estimates derived from the separate exposure assessment methods within a GAM, increased the CV-R² value by 9% to 19%. The largest improvement in terms of CV-R² was for PM10.

The hybrid modeling approaches for NO₂ showed better ability to predict spatially than temporally, while for ozone and particulate matter the predicted better temporally (Table 4).

Figure 3 shows the yearly pattern (2009 to 2013) of the combined estimates derived from independent exposure assessment methods, within a GAM. The distribution of the estimated pollutant concentrations over the study period is similar. PM concentrations have more outliers compared to NO₂ and O₃ series.

Table 3. Model performance evaluated by the value of adjusted R² and 10-fold cross validated (CV) R², root mean square error (RMSE) and mean bias, for the independent and hybrid modeling approaches.

	Pollutant (µg/m ³)			
	NO ₂ (24 h Average)	O ₃ (8 h Max)	PM10 (24 h Average)	PM2.5 (24 h Average)
Number of Fixed Monitoring Sites	130	62	115	104
Independent exposure assessment methods				
ST LUR model ¹				
R ² _{adj} and (CV-R ²)	0.72 (0.71)	0.69 (0.62)	0.61 (0.57)	0.69 (0.66)
RMSE	4.28	13.67	7.42	3.64
Mean bias ²	-5.60	-0.14	0.96	0.59
Dispersion ³				
R ² _{adj} and (CV-R ²)	0.73 (0.70)	0.60 (0.59)	0.71 (0.69)	0.75 (0.74)
RMSE	4.13	15.60	6.41	4.26
Mean bias ²	0.73	-0.02	0.81	0.45
Ensemble ⁴				
R ² _{adj} and (CV-R ²)	-	-	-	0.88 (0.83)
RMSE	-	-	-	-
Mean bias ²	-	-	-	0.058
Hybrid exposure assessment methods				
Hybrid 1 ⁵				
R ² _{adj} and (CV-R ²)	0.84 (0.76)	0.79 (0.75)	0.82 (0.79)	0.84 (0.81)
RMSE	3.71	10.24	2.72	0.20
Mean bias ²	-7.13	-11.12	-0.35	0.13
Hybrid 2 ⁶				
R ² _{adj} and (CV-R ²)	0.81 (0.80)	0.76 (0.75)	0.77 (0.76)	0.80 (0.79)
RMSE	3.64	11.94	4.02	1.91
Mean bias ²	1.64	0.03	0.68	0.36

¹ developed spatio-temporal (ST) LUR models. RMSE: Root Mean Square Error. ² bias = measured concentrations from fixed monitoring sites—10-fold CV predicted concentrations. ³ spatio-temporal dispersion model. ⁴ PM2.5 prediction model based on an ensemble machine learning spatio-temporal approach. ⁵ Incorporation of estimates derived from independent exposure assessment methods, within the LUR model. ⁶ Combination of estimates derived from independent exposure assessment methods, within a GAM R²_{adj}: Adjusted R² value of model. CV: 10-fold cross validation. CV -R²: R² value of cross validated model.

Table 4. Temporal and spatial fit of the hybrid modeling approaches. Results from 10-fold cross validation.

	Pollutant (µg/m ³)							
	NO ₂ (24 h Average)		NO ₂ (24 h Average)		NO ₂ (24 h Average)		NO ₂ (24 h Average)	
	R ² .Spatial	R ² .Spatial						
Hybrid 1 ¹	0.67	0.61	0.59	0.74	0.52	0.70	0.47	0.82
Hybrid 2 ²	0.72	0.63	0.61	0.72	0.62	0.76	0.59	0.87

¹ Incorporation of estimates derived from independent exposure assessment methods, within the LUR model. ² Combination of estimates derived from independent exposure assessment methods, within a GAM. R²: R² value of 10-fold cross validated model.

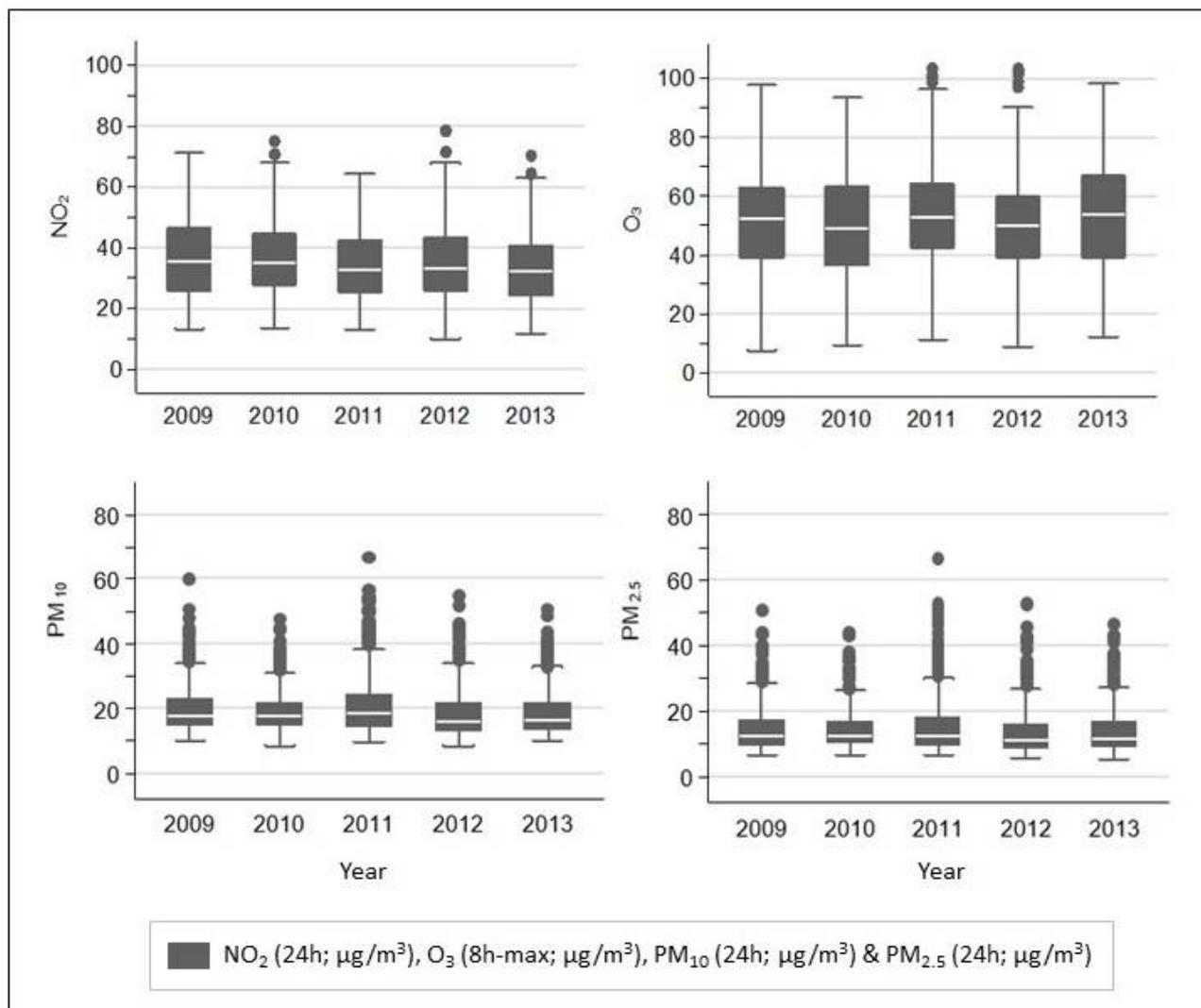


Figure 3. Yearly average (years 2009 to 2013) of estimated NO_2 (24 h; $\mu\text{g}/\text{m}^3$), O_3 (8 h-max; $\mu\text{g}/\text{m}^3$), PM_{10} (24 h; $\mu\text{g}/\text{m}^3$) and $\text{PM}_{2.5}$ (24 h; $\mu\text{g}/\text{m}^3$) concentrations from the hybrid 2 model, in the Greater London Area. Hybrid 2: Combination of estimates derived from independent exposure assessment methods, within a GAM.

Figure 4 displays an application of the combination of estimates derived from independent exposure assessment methods, within a GAM to predict pollutant long-term concentrations, per LSOA. The combination of estimates derived from ST LUR and dispersion models into a GAM framework (hybrid model 2) allows the relative weights for each model to vary spatially and by concentration and, therefore, display better performance.

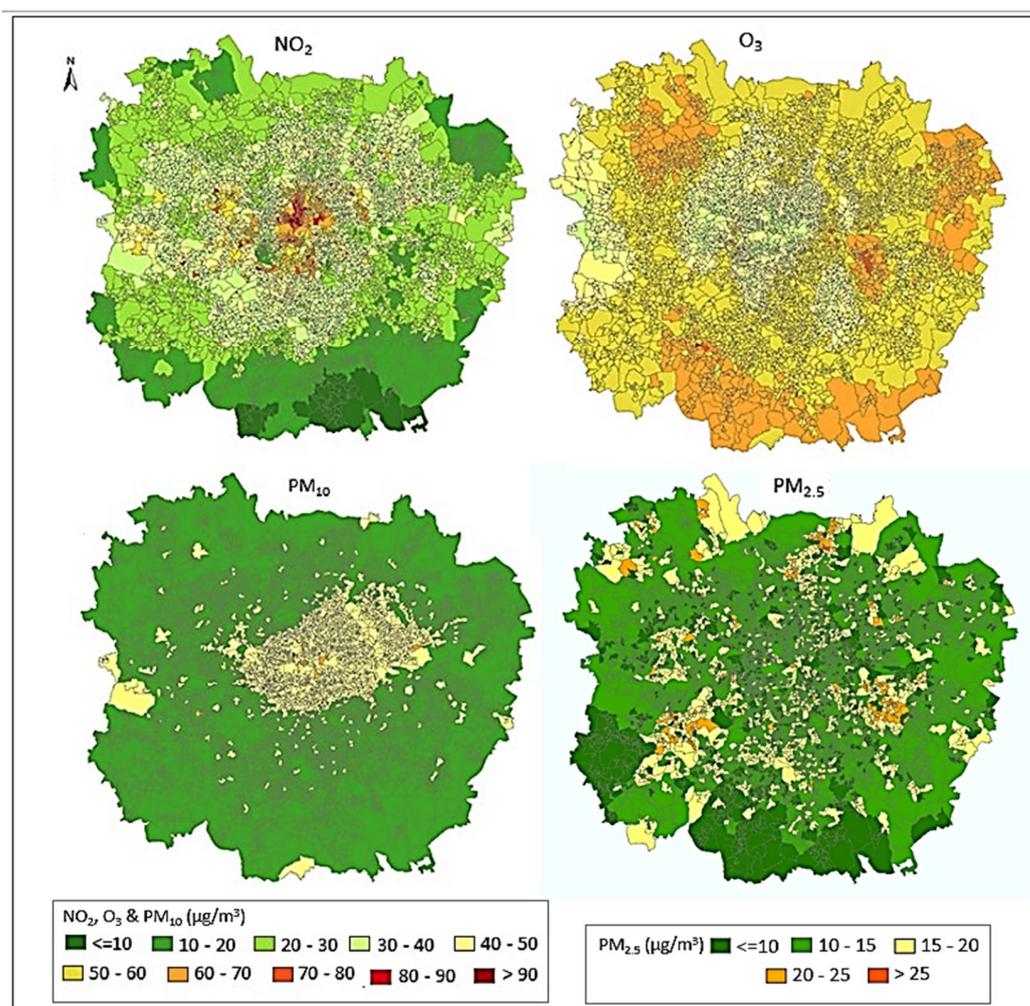


Figure 4. Long-term average (years 2009–2013) of estimated NO₂ (24 h; µg/m³), O₃ (8 h-max; µg/m³), PM₁₀ (24 h; µg/m³) and PM_{2.5} (24 h; µg/m³) concentrations from the hybrid 2 model, per LSOA in the Greater London Area. *Hybrid 2: Combination of estimates derived from independent exposure assessment methods, within a GAM.*

4. Discussion

4.1. Findings

We developed a number of air pollution exposure assessment approaches for the Greater London Area, based on different methodological principles, to estimate concentrations in fine spatial (LSOA level) and temporal (daily) scales. For NO₂, O₃ and PM₁₀ we developed a ST LUR and dispersion models, whilst for PM_{2.5} we additionally developed a model using machine learning algorithms and incorporating satellite data. These independent methods of pollutant concentration estimates are prone to errors from the uncertainty inherent to the measurement of variables used to develop the corresponding models. The errors are likely independent, as the dispersion modelling uses an emission inventory and information on atmospheric transformation processes influenced by the urban-scape, whilst the LUR models are based on air pollution measurements and the spatial and temporal variables determining their magnitude. So, it appears intuitively attractive to combine these methods, expecting that the errors of each separate method will cancel out. In the present project we used two types of combination models for each pollutant: one that incorporates the results of the dispersion model as a covariate in the LUR (hybrid 1) and one that combines the predictions from the 2 or 3 (for PM_{2.5}) independent models with a GAM (Hybrid 2).

The Greater London Area has the advantage of a very dense monitoring network which provides measurements, including 130 sites for NO₂, 115 for PM₁₀, 62 for O₃. The smaller number of sites measuring PM_{2.5} (n = 33) was enhanced by additionally using the data base developed in [25], resulting in the use of 104 sites. This network was used to predict the pollutant concentrations at each site and assess the exposure error and the agreement between methods.

Our results indicate that the combination models performed better in terms of cross-validated R², RMSE and mean bias. It should be noted that for pollutants with high spatial variability and good knowledge on the determinants of this variability, such as NO₂, models explain a larger proportion of this variability. For pollutants that are more spatially homogenous and tend to have larger temporal variability, the models tend to explain the temporal variability better. It is also interesting to note that the combination model 1 performs better for PM_{2.5} and PM₁₀ whilst combination model 2 better for the gaseous pollutants.

4.2. Evaluation of the Combined Modeling Performance

Relatively few studies have compared the performance of model combinations of LUR and dispersion models at daily and fine spatial scales. The temporal scale ranges in the published work from annual to monthly or biweekly and seldom to daily. Additionally, the models have been developed for very different geographical areas and although the concepts are often similar, the models developed and those compared follow very different methodologies. Some studies compared different prediction methods (e.g., regression models and machine learning algorithms) using the same predictors and generally find that the application of machine learning methods yields better predictions [45–47].

De Hoogh et al. [18] compared the performance of LUR and dispersion modelling but did not assess the performance of any combination model. In a later work, De Hoogh et al. [48] developed and evaluated an extended LUR model for predicting annual concentrations in Western Europe for PM_{2.5}, NO₂, BC and O₃, incorporating dispersion model estimates, kriging, satellite observations in addition to the LU variables. The combined full model performed better compared to less sophisticated models. This work was extended to eight elemental PM components [49]. Akita et al. [17] compared the performance of several models with a combined one using Bayesian Maximum Entropy in predicting annual NO₂ concentrations in Catalunya, Spain, and report that their proposed combined framework outperformed the more conventional (LUR, Dispersion and others) approaches based on RMSE and other indices. Wang et al. [19] developed a combined LUR and chemical transport model, using a geostatistical modelling framework, for bi-weekly estimation of O₃ and PM_{2.5} in the Los Angeles Basin and report that the combined model outperformed the initial models especially improving the accuracy of O₃ predictions. Tripathy et al. [50] developed a combined model (“Hybrid”) for PM_{2.5}, BC and metal components for the Pittsburg area, which performed better for PM_{2.5} than for the other components, but its performance was not compared to other models.

We compared the performance of models in terms of accuracy and bias in the predictions in space and time. Other papers evaluate how well the models perform in terms of providing valid and accurate estimates of the pollution exposure association with health outcomes. The optimal methods under these assessments do not necessarily coincide [51]. The exposure assessment methods presented in this paper have been evaluated in terms of their performance in estimating health effects using simulations that indicated the Hybrid 2 model to perform better both for PM and gaseous pollutants [52,53].

4.3. Advantages and Limitations

Our work has some advantages: it relies on a very dense and extensive monitoring network allowing many points in space for validation. Additionally, it combines exposure assessment approaches widely used in epidemiological studies. Further it assesses concentration estimates in a very fine spatial and temporal scale. However, it also has a number of

limitations. London has a dense air pollution monitoring network. This may not be the case in other urban settings, especially those suffering from poor air quality. In such cities, the lack of measurements may limit the possible modelling and hybrid approaches. However, this could be overcome by designing a specific monitoring campaign or by applying methods to enhance existing measurement data bases [25]. The methods evaluated are only a subset of those that can be developed for the same area. Thus, other models may incorporate further data as available for example satellite data which were only used for PM_{2.5} in our study and other algorithms for prediction. Other types of combinations to produce hybrid models may be used. Models using land use/cover variables to predict air pollutant concentrations should be periodically updated to capture any changes in land use. In Europe, freely available land use databases (i.e., CORINE, Urban Atlas) are updated every 6 years. Data from local sources could be included to account for intense or fast changes in land use and therefore improve predictive performance of LUR models. Additionally, the transferability of the comparison results to other areas is questionable. As many characteristics determining space and time specific pollutant concentrations depend on the local topography, urban characteristics, population behavior and climate, our results may not be readily transferable to other locations and this aspect should be further investigated.

5. Conclusions

In conclusion, we show that combination or hybrid exposure models combining independent modelling methods based on different methodological principles perform better in terms of valid and accurate estimations of concentrations in time and space. This is broadly in accordance to the sparse and not directly comparable results that have already been published for other geographical locations. Future work should further evaluate methods that combine approaches (often termed “hybrid” to denote a variety of combinations) which appear consistently to outperform the single method approach.

Author Contributions: Conceptualization K.K., E.S., J.S., S.B., B.B. and J.G.; Data curation, K.D., A.A., N.K., A.B., B.B., S.R., S.Z. and J.G.; Formal analysis, K.D., A.A., N.K., A.B., S.R. and S.Z.; Funding acquisition, K.K.; Methodology, K.D., E.S., A.A., J.S., S.B., N.K., A.B., B.B., S.R., S.Z., J.G. and K.K.; Supervision, E.S., J.S., S.B., B.B., J.G. and K.K.; Writing—original draft, K.D. and K.K.; Writing—review & editing, K.D., K.K., E.S., A.A., J.S., S.B., N.K., A.B., B.B., S.R., S.Z. and J.G. All authors have read and agreed to the published version of the manuscript.

Funding: Research in this paper was funded under the MRC UK Grant, ref: MR/N014464/1.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Publicly archived datasets analyzed in this study: Census 2011 data; <https://data.london.gov.uk/dataset/2011-census-demography> (accessed on 1 March 2017); London Air Quality Network (LAQN). King’s College, London. Available online: <http://www.londonair.org.uk/> (accessed on 1 March 2017); London Datastore (accessed on 13 April 2020). Available online: <https://data.london.gov.uk/dataset/lsoa-atlas> (accessed on 18 November 2016); Air Quality England (AQE). Ricardo Energy and Environment. Available online: [http://www.airqualityengland.co.uk/Automatic Urban and Rural Network \(AURN\)](http://www.airqualityengland.co.uk/Automatic-Urban-and-Rural-Network-(AURN)) (accessed on 1 March 2017). Data Archive: Available online: <https://uk-air.defra.gov.uk> (accessed on 6 December 2016); Land Cover Map. Available online: <http://www.ukso.org/static-maps/land-cover-map.html> (accessed on 29 November 2016); UK Meteorological Office, Available online: <https://www.metoffice.gov.uk/research/climate/maps-and-data/historic-station-data> (accessed on 21 November 2016); Department of Transport in the United Kingdom, Available online: <https://data.gov.uk/dataset/d1f9e79f-d9db-44d0-b7b1-41c216fe5df6/national-public-transport-data-repository-nptdr> (accessed on 4 August 2017); EMEP European emissions (accessed on 15 June 2016), Available online: <https://www.ceip.at/>; CERC, ADMS Roads v4 User Guide. (accessed on 15 December 2016) Available online: http://www.cerc.co.uk/environmental-software/assets/data/doc_userguides/CERC_ADMS-Roads4.0_User_Guide.pdf (accessed on 15 December 2016); Greater London Authority (GLA). The London Atmospheric Emissions Inventory 2013. Available online: <https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory-2013> (accessed on 6 December 2016).

Acknowledgments: We are grateful to the UK Met Office for provision of meteorological data, accessed through the Centre for Environmental Data Analysis (CEDA). We are also grateful to the UK Government and Local Authorities providing air pollution measurements used in this study, managed by King's College London and Ricardo Energy and Environment.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Beelen, R.; Raaschou-Nielsen, O.; Stafoggia, M.; Andersen, Z.J.; Weinmayr, G.; Hoffmann, B.; Wolf, K.; Samoli, E.; Fischer, P.; Nieuwenhuijsen, M.; et al. Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet* **2014**, *383*, 785–795. [[CrossRef](#)]
2. Cesaroni, G.; Badaloni, C.; Gariazzo, C.; Stafoggia, M.; Sozzi, R.; Davoli, M.; Forastiere, F. Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ. Health Perspect.* **2013**, *121*, 324–331. [[CrossRef](#)] [[PubMed](#)]
3. Kloog, I.; Koutrakis, P.; Coull, B.A.; Lee, H.J.; Schwartz, J. Assessing temporally and spatially resolved PM_{2.5} exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmos. Environ.* **2011**, *45*, 6267–6275. [[CrossRef](#)]
4. Jerrett, M.; Burnett, R.T.; Kanaroglou, P.S.; Eyles, J.; Brook, J.R.; Giovis, C.; Finkelstein, N. A GIS—environmental justice analysis of particulate air pollution in Hamilton, Canada. *Environ. Plan. A* **2001**, *33*, 955–973. [[CrossRef](#)]
5. Miller, K.A.; Siscovick, D.S.; Sheppard, L.; Shepherd, K.; Sullivan, J.H.; Anderson, G.L.; Kaufman, J.D. Long-term exposure to air pollution and incidence of cardiovascular events in women. *N. Engl. J. Med.* **2007**, *356*, 447–458. [[CrossRef](#)] [[PubMed](#)]
6. Beelen, R.; Hoek, G.; van den Brandt, P.A.; Goldbohm, R.A.; Fischer, P.; Schouten, L.J.; Jerrett, M.; Hughes, E.; Armstrong, B.; Brunekreef, B. Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environ Health Perspect.* **2008**, *116*, 196–202. [[CrossRef](#)] [[PubMed](#)]
7. Tonne, C.; Elbaz, A.; Beevers, S.; Singh-Manoux, A. Traffic-related air pollution in relation to cognitive function in older adults. *Epidemiology* **2014**, *25*, 674–678. [[CrossRef](#)]
8. Fuks, K.B.; Weinmayr, G.; Hennig, F.; Tzivian, L.; Moebus, S.; Jakobs, H.; Memmesheimer, M.; Kälsch, H.; Andrich, S.; Nonnemacher, M.; et al. Association of long-term exposure to local industry- and traffic-specific particulate matter with arterial blood pressure and incident hypertension. *Int. J. Hyg. Environ. Health* **2016**, *219*, 527–535. [[CrossRef](#)]
9. Raaschou-Nielsen, O.; Andersen, Z.J.; Jensen, S.S.; Kettel, M.; Sorensen, M.; Hansen, J.; Loft, S.; Tjønneland, A.; Overvad, K. Traffic air pollution and mortality from cardiovascular disease and all causes: A Danish cohort study. *Environ. Health* **2012**, *11*, 60. [[CrossRef](#)]
10. Van Donkelaar, A.; Martin, R.V.; Brauer, M.; Kahn, R.; Levy, R.; Verduzco, C.; Villeneuve, P.J. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environ. Health Perspect.* **2010**, *118*, 847–855. [[CrossRef](#)]
11. Hoek, G. Methods for assessing long-term exposures to outdoor air pollutants. *Curr. Environ. Health Rep.* **2017**, *4*, 450–462. [[CrossRef](#)] [[PubMed](#)]
12. Hoek, G.; Beelen, R.; Cd, H.; Vienneau, D.; Gulliver, J.; Fischer, P.; Briggs, D. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* **2008**, *42*, 7561–7578. [[CrossRef](#)]
13. Vallero, D.A. Chapter 14—Air pollution dispersion models. In *Air Pollution Calculations*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 429–448. ISBN 9780128149348.
14. Kloog, I.; Sorek-Hamer, M.; Lyapustin, A.; Coull, B.; Wang, Y.; Just, A.C.; Schwartz, J.; Broday, D.M. Estimating daily PM_{2.5} and PM₁₀ across the complex geoclimate region of Israel using MAIAC satellite-based AOD data. *Atmos. Environ.* **2015**, *122*, 409–416. [[CrossRef](#)] [[PubMed](#)]
15. World Health Organization (WHO); Prüss-Ustün, A.; Wolf, J.; Corvalán, C.; Bos, R.; Neira, M. *Preventing Disease through Healthy Environments: A Global Assessment of the Burden of Disease from Environmental Risks*; World Health Organization: Geneva, Switzerland, 2016.
16. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019, a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **2020**, *396*, 1204–1222, Erratum in: *Lancet* **2020**, *396*, 1562. [[CrossRef](#)]
17. Akita, Y.; Baldasano, J.M.; Beelen, R.; Cirach, M.; de Hoogh, K.; Hoek, G.; Nieuwenhuijsen, M.; Serre, M.L.; de Nazelle, A. Large scale air pollution estimation method combining land use regression and chemical transport modeling in a geostatistical framework. *Environ. Sci. Technol.* **2014**, *48*, 4452–4459. [[CrossRef](#)]
18. de Hoogh, K.; Korek, M.; Vienneau, D.; Keuken, M.; Kukkonen, J.; Nieuwenhuijsen, M.J.; Badaloni, C.; Beelen, R.; Bolignano, A.; Cesaroni, G. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environ. Int.* **2014**, *73*, 382–392. [[CrossRef](#)]
19. Wang, M.; Sampson, P.D.; Hu, J.; Kleeman, M.; Keller, J.P.; Olives, C.; Szpiro, A.A.; Vedal, S.; Kaufman, J.D. Combining Land-Use Regression and Chemical Transport Modeling in a Spatiotemporal Geostatistical Model for Ozone and PM_{2.5}. *Environ. Sci. Technol.* **2016**, *50*, 5111–5118. [[CrossRef](#)]

20. Di, Q.; Koutrakis, P.; Schwartz, J. A hybrid prediction model for PM_{2.5} mass and components using a chemical transport model and land use regression. *Atmos. Environ.* **2016**, *131*, 390–399. [[CrossRef](#)]
21. London Datastore. Available online: <https://data.london.gov.uk/dataset/Isao-atlas> (accessed on 13 April 2020).
22. London Air Quality Network (LAQN). King's College, London. Available online: <http://www.londonair.org.uk/> (accessed on 1 March 2017).
23. Air Quality England (AQE). Ricardo Energy and Environment. Available online: <http://www.airqualityengland.co.uk/> (accessed on 1 March 2017).
24. Automatic Urban and Rural Network (AURN) Data Archive. © Crown 2017 Copyright Defra Via <https://uk-air.defra.gov.uk>, Licensed under the Open Government License (OGL) v2.0. Available online: <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/> (accessed on 1 March 2017).
25. Analitis, A.; Barratt, B.; Green, D.; Beddows, A.; Samoli, E.; Schwartz, J.; Katsouyanni, K. Prediction of PM_{2.5} concentrations at the locations of monitoring sites measuring PM₁₀ and NO_x, using generalized additive models and machine learning methods, for enhancement of the PM_{2.5} database: A case study in London. *Atmos. Environ.* **2020**, *240*, 117757. [[CrossRef](#)]
26. Dimakopoulou, K.; Samoli, E.; Katsouyanni, K. Spatio-temporal land use regression modelling of ozone levels in Athens, Greece. *Glob. Nest J.* **2020**, *22*, 85–94. [[CrossRef](#)]
27. Gryparis, A.; Dimakopoulou, K.; Pedeli, X.; Katsouyanni, K. Spatio-temporal semiparametric models for NO₂ and PM₁₀ concentration levels in Athens, Greece. *Stoen* **2014**, *479–480*, 21–30. [[CrossRef](#)] [[PubMed](#)]
28. ESRI. *ArcGIS Desktop: Release 10*; Environmental Systems Research Institute: Redlands, CA, USA, 2011.
29. R Core Team (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available online: <https://www.R-project.org/> (accessed on 1 February 2017).
30. Wand, M. SemiPar: Semiparametric Regression. R package Version 1.0-4.2 SemiPar 1.0. R Package. Available online: <http://cran.r-project.org> (accessed on 5 January 2018).
31. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Barker, D.M.; Duda, M.D.; Huang, X.-Y.; Wang, W.; Powers, J.G. *A Description of the Advanced Research WRF Version 3*; NCAR/TN-475+STR; University Corporation for Atmospheric Research: Denver, CO, USA, 2008.
32. Byun, D.W.; Ching, J.K.S. *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modelling System*; EPA/600/R-99/030; U.S. Environmental Protection Agency, Office of Research and Development: Washington, DC, USA, 1999.
33. Beevers, S.D.; Kitwiroon, N.; Williams, M.L.; Carslaw, D.C. One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions. *Atmos. Environ.* **2012**, *59*, 47–58. [[CrossRef](#)] [[PubMed](#)]
34. CERC, ADMS Roads v4 User Guide. Available online: http://www.cerc.co.uk/environmental-software/assets/data/doc_userguides/CERC_ADMS-Roads4.0_User_Guide.pdf (accessed on 15 December 2016).
35. Defra. Emissions of Air Quality Pollutants 1990–2014. Defra, UK. 2016. Available online: https://uk-air.defra.gov.uk/assets/documents/reports/cat07/1609130906_NAEI_AQPI_Summary_Report_1990-2014_Issue1.1.pdf (accessed on 6 December 2016).
36. Greater London Authority (GLA). The London Atmospheric Emissions Inventory 2013. 2016. Available online: <https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory-2013> (accessed on 6 December 2016).
37. Beevers, S.D.; Westmoreland, E.; de Jong, M.C.; Williams, M.L.; Carslaw, D.C. Trends in NO_x and NO₂ emissions from road traffic in Great Britain. *Atmos. Environ.* **2012**, *54*, 107–116. [[CrossRef](#)]
38. Vukovich Jeffrey, M.; Pierce, T.E. The Implementation of BEIS3 within the SMOKE modeling framework. In Proceedings of the 11th International Emissions Inventory Conference, Atlanta, Georgia, 15–18 April 2002.
39. Lee, S.; Wolberg, G.; Shin, S. Scattered data interpolation with multilevel B-splines. *IEEE Trans. Vis. Comput. Graph.* **1997**, *3*, 228–244. [[CrossRef](#)]
40. Defra. Evaluating the Performance of Air Quality Models. Defra, UK. 2010. Available online: https://uk-air.defra.gov.uk/assets/documents/reports/cat05/1006241607_100608_MIP_Final_Version.pdf (accessed on 6 December 2016).
41. Danesh Yazdi, M.; Kuang, Z.; Dimakopoulou, K.; Barratt, B.; Suel, E.; Amini, H.; Lyapustin, A.; Katsouyanni, K.; Schwartz, J. Predicting Fine Particulate Matter (PM_{2.5}) in the Greater London Area: An Ensemble Approach using Machine Learning Methods. *Remote Sens.* **2020**, *12*, 914. [[CrossRef](#)]
42. Lyapustin, A.; Wang, Y.; Korokin, S.; Huang, D. MODIS Collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* **2018**, *11*, 5741–5765. [[CrossRef](#)]
43. Lin, L.I.-K. A Concordance Correlation Coefficient to evaluate reproducibility. *Biometrics* **1989**, *45*, 255–268. [[CrossRef](#)]
44. Bland, J.M.; Altman, D.G. Measuring agreement in method comparison studies. *Stat Methods Med. Res.* **1999**, *8*, 135–160. [[CrossRef](#)]
45. Araki, S.; Shima, M.; Yamamoto, K. Spatiotemporal land use random forest model for estimating metropolitan NO₂ exposure in Japan. *Sci. Total Environ.* **2018**, *634*, 1269–1277. [[CrossRef](#)]
46. Chen, J.; de Hoogh, K.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Bauwelinck, M.; van Donkelaar, A.; Hvidtfeldt, U.A.; Katsouyanni, K. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Environ. Int.* **2019**, *130*, 104934. [[CrossRef](#)]
47. Wong, P.Y.; Lee, H.Y.; Chen, Y.C.; Zeng, Y.T.; Chern, Y.R.; Chen, N.T.; Candice Lung, C.S.; Su, H.J.; Wu, C.D. Using a land use regression model with machine learning to estimate ground level PM_{2.5}. *Environ. Pollut.* **2021**, *277*, 116846. [[CrossRef](#)] [[PubMed](#)]

48. de Hoogh, K.; Chen, J.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Bauwelinck, M.; van Donkelaar, A.; Hvidtfeldt, U.A.; Katsouyanni, K.; et al. Spatial PM_{2.5}, NO₂, O₃ and BC models for western Europe—evaluation of spatiotemporal stability. *Environ. Int.* **2018**, *120*, 81–92. [[CrossRef](#)] [[PubMed](#)]
49. Chen, J.; de Hoogh, K.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Weinmayr, G.; Bauwelinck, M.; van Donkelaar, A.; Hvidtfeldt, U.A.; et al. Development of Europe-wide models for particle elemental composition using supervised linear regression and random forest. *Environ. Sci. Technol.* **2020**, *54*, 15698–15709. [[CrossRef](#)]
50. Tripathy, S.; Tunno, B.J.; Michanowicz, D.R.; Kinnee, E.; Shmool, J.L.; Gillooly, S.; Clougherty, J.E. Hybrid land use regression modeling for estimating spatio-temporal exposures to PM_{2.5}, BC, and metal components across a metropolitan area of complex terrain and industrial sources. *Sci. Total Environ.* **2019**, *673*, 54–63. [[CrossRef](#)]
51. Szpiro, A.A.; Sheppard, L.; Lumley, T. Efficient measurement error correction with spatially misaligned data. *Biostatistics* **2011**, *12*, 610–623. [[CrossRef](#)] [[PubMed](#)]
52. Samoli, E.; Butland, B.K.; Rodopoulou, S.; Atkinson, R.W.; Barratt, B.; Beevers, S.D.; Beddows, A.; Dimakopoulou, K.; Schwartz, J.D.; Yazdi, M.D.; et al. The impact of measurement error in modeled ambient particles exposures on health effect estimates in multilevel analysis. *Environ. Epidemiol.* **2020**, *4*, 094. [[CrossRef](#)] [[PubMed](#)]
53. Butland, B.K.; Samoli, E.; Atkinson, R.W.; Barratt, B.; Beevers, S.D.; Kitwiroon, N.; Dimakopoulou, K.; Rodopoulou, S.; Schwartz, J.D.; Katsouyanni, K. Comparing the performance of air pollution models for nitrogen dioxide and ozone in the context of a multilevel epidemiological analysis. *Environ. Epidemiol.* **2020**, *4*, 093. [[CrossRef](#)] [[PubMed](#)]