Environmental Science & Technology

ARTICLE

pubs.acs.org/est

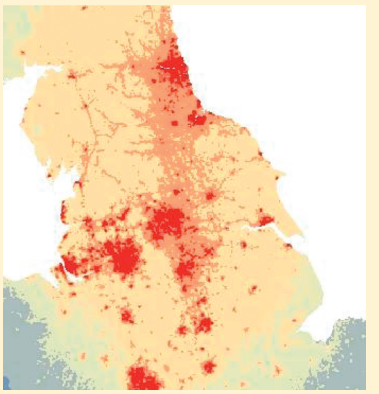
Land Use Regression Modeling To Estimate Historic (1962−1991) Concentrations of Black Smoke and Sulfur Dioxide for Great Britain

John Gulliver,*,† Chloe Morris,† Kayoung Lee,† Danielle Vienneau,† David Briggs,† and Anna Hansell†

†MRC-HPA Centre for Environment and Health, Department of Epidemiology & Biostatistics, Imperial College London, London, U.K.

Supporting Information

ABSTRACT: Land-use regression modeling was used to develop maps of annual average black smoke (BS) and sulfur dioxide (SO_2) concentrations in 1962, 1971, 1981, and 1991 for Great Britain on a 1 km grid for use in epidemiological studies. Models were developed in a GIS using data on land cover, the road network, and population, summarized within circular buffers around air pollution monitoring sites, together with altitude and coordinates of monitoring sites to consider global trend surfaces. Models were developed against the log-normal (LN) concentration, yielding R^2 values of 0.68 (n = 534), 0.68 (n = 767), 0.41 (n = 771), and 0.39 (n = 155) for BS and 0.61 (n = 482), 0.65 (n = 733), 0.38 (n = 756), and 0.24 (n = 153) for SO_2 in 1962, 1971, 1981, and 1991, respectively. Model evaluation was undertaken using concentrations at an independent set of monitoring sites. For BS, values of R^2 were 0.56 (n = 133), 0.41 (n = 191), 0.38 (n = 193), and 0.34 (n = 37), and for SO_2 values of R^2 were 0.71 (n = 121), 0.57 (n = 183), 0.26 (n = 189), and 0.31 (n = 38) for 1962, 1971, 1981, and 1991, respectively. Models slightly underpredicted (fractional bias: 0∼−0.1) monitored concentrations of both pollutants for all years. This is the first study to produce historic concentration maps at a national level going back to the 1960s.



INTRODUCTION

Much of the evidence showing impacts of air pollution on health has come from time-series (i.e., short-term) epidemiological studies.[1] There are far fewer long-term studies - almost none go back much before the 1980s - in part due to the lack of detailed information on historic exposures. The UK has a very long-running network of air pollution monitors for black smoke (BS) and sulfur dioxide (SO_2) with over 3100 monitors in operation from 1955, but the network of monitors has never been sufficiently dense (200−1000 at any one time point) to capture variation in exposures across the whole population. However, cardiovascular, cardiorespiratory, and respiratory mortality, during 1982−1986 to 1994−1998, in the British Census wards containing a monitoring station (totalling <5% of the population) were found to have significant associations between measured concentrations of BS and SO_2, but results suggested attenuation of effect over time.[2]

GIS (Geographical Information Systems) based air pollution modeling offers the potential to improve the spatial resolution in this type of study but has thus far seen limited use in assessing historic exposures to air pollution. Reconstructed emissions data on traffic and heating were used in a dispersion model linked to a GIS to model residential exposures to nitrogen oxides and SO_2 across Greater Stockholm over three points in time: 1960, 1970, and 1980.[3] Air pollution dispersion models were developed to estimate annual average population exposures to PM_{10}, NO_2, SO_2, and O_3 in England for 1994−2003.[4] In The Netherlands, GIS techniques were used to estimate residential exposures to BS, NO_2, SO_2, and $PM_{2.5}$ between 1986 and 1996 as part of a study looking at the effects of traffic air pollution on mortality.[5] It is noteworthy that only one study has attempted to model exposures of the population further back in time than the 1980s.[3]

The Chronic Health Effects of Smoke and SO_2 (CHESS) study in Great Britain aims to establish whether long-term air pollution exposure is associated with cardiovascular and respiratory disease in adults. As the basis for exposure assessment, GIS methods were developed for modeling and mapping historic outdoor concentrations of BS and SO_2 for the whole of the UK, at 1 km resolution. A range of techniques were trialled including land use regression (LUR),[6] kriging,[7,8] dispersion modeling,[9] and focal-sum[10] as well as simple interpolation methods such as Inverse Distance Weighting (IDW). LUR was seen to perform consistently better than the other techniques when model estimates were compared with monitored concentrations. Indeed, LUR methods have seen widespread use in modeling city-wide, annual average concentrations of a range of air pollutants.[11−17] The growth in their application is supported by generally good performance results in model evaluation studies and by being relatively easy to develop and apply. More recently, LUR models have been developed for national-scale[18] and European-wide mapping of air pollution.[8] Only once before, however, has LUR been used to develop maps of air pollution for historical exposure assessment.[5] Other attempts at applying LUR to predict exposure for earlier years have relied on

Received: November 12, 2010
Accepted: March 15, 2011
Revised: February 22, 2011
Published: March 29, 2011

back-extrapolation by calibrating current models to monitored concentrations for earlier years.[19]

This paper reports on the development and evaluation of LUR models for the UK for four time points: 1962, 1971, 1981, and 1991.

METHODS

Monitored Concentrations. Great Britain has continuous electronically available data on monitored concentrations of both BS and SO_2 from April 1961. However, the number and locations of monitoring sites has varied over time. In the early 1970s there were more than 1200 BS and 900 SO_2 sites, followed by a gradual reduction in numbers until by 1991 there were only 285 BS sites and 202 SO_2 sites. Since the early 1990s there have been further minor reductions in the total number of BS and SO_2 sites (but a rapid increase in sites monitoring NO_2 and PM_{10} reflecting the increase in traffic-related air pollution). Summary information about the mean and variability of monitored concentrations of BS and SO_2 for the target years can be found in the Supporting Information (Table S1). The measurement method for both pollutants has remained the same across all years (British Standard 1747).

Average concentrations of both pollutants fell dramatically between 1962 and 1981, with a 7-fold fall in BS and three- to 4-fold fall in SO_2, reflecting a major switch from domestic and industrial coal use to gas and electricity. The distribution of concentrations of BS was highly skewed in all years but markedly skewed in 1981 reflecting the relatively low level of geographical variation in concentrations and only a small number of urban districts with relatively high concentrations. Distributions of SO_2 were less marked in their skewness with the exception, again, of 1981. All distributions are log-normal and monitored concentrations for all years were therefore log-transformed (Ln) prior to model development.

Data on daily average concentrations of BS and SO_2 were obtained from the national air quality archive (www.airquality.co.uk accessed on 21 February 2011) for the years 1962 (instead of 1961 as the full calendar year was not available electronically), 1971, 1981, and 1991 to align with the 10-yearly UK population census. Monitoring stations were retained for modeling if they had at least 75% data capture within each year. Spatial accuracy for monitoring stations is 100 m.

Development of Predictor Variables. A GIS was developed with a range of predictor variables including information on land cover, roads, population, altitude, and X-Y coordinates of each monitoring site. CORINE (CLC1990) for Europe comprises 44 categories of land cover and is provided on a 100-m grid, but the accuracy varies from 100 to 250 m. Data on British roads came from the earliest version of OS Meridian, dating back to the mid-1990s. It includes motorways, A-roads (major trunk roads, dual carriageways and arterial roads), B-roads (roads with significant traffic flows but not in 'A' class), and minor roads (urban side-roads and country lanes). The roads were intersected with the same 100 m grid as CORINE and the length within each grid cell summed. No (digital) data were available on land cover or roads pre-1990s, so these data sets were used in all years. Data on population counts were estimated from Census geography using simple areal-weighting for 1962 and 1971 and by postcode-weighting for 1981 and 1991. (Altitudes were derived from the OS PANORAMA 50 m digital terrain model and assigned to each monitoring site using a point-in-grid function in ArcGIS. The X-Y coordinates of each monitoring site were used to compute the variables for trend surfaces. Trend surface analysis on monitored concentrations was initially explored in ArcGIS using kriging. Predictor variables were subsequently computed for second order trend (i.e., quadratic) on y coordinates (i.e., latitude) and third order trend (i.e., cubic). Although predictors were available for a fine grid, the notional accuracy of CORINE and precision of the monitoring sites (ca. 100 m) precluded modeling at this scale. A minimum resolution for modeling was thus defined as 500 m, with maps presented at 1 km.

"Buffering", using ArcGIS FOCALSUM with the circle option (subsequently referred to as buffers), was used to sum the contribution (i.e., area or length) of each land cover and road variable around each monitoring site, based on the 100 m grids. Both zero-centered and ring variables were constructed.[18] Buffers of 0.5, 1, 2, and 3 km were applied to road variables and 0.5, 1, 2, 3, 4, 5, 7.5, and 10 km buffers were applied to land cover and population, and 24 variables in domains of road, landcover, population, x-y trend, and altitude were considered. Table S2 shows a description of each variable, the buffer sizes applied, and the required direction of effect in the subsequent regression analysis. The complete list of variables, joined with data on annual average monitored concentrations, was imported into SPSS for data analysis and model development.

Model Development. LUR models were developed for BS and SO_2 for 1962, 1971, 1981, and 1991. Models were developed against a stratified random sample (based on X-Y coordinates to avoid geographical bias and site type) of 80% of monitoring sites in each year with 20% of sites retained for model evaluation.

A number of rules were used to determine the entry and order of predictor variables: 1) each variable had a significant correlation with the monitored concentration ($p < 0.05$), 2) the direction of correlation (i.e., effect) met predetermined expectations (see Table S2), 3) the direction of effect of predictors already in the model did not change as subsequent predictors were added, 4) predictors already in the model remained significant at $p < 0.1$, and 5) land use variables and population were not allowed in the same model (i.e., double-counting of population). A supervised forward approach was used.[5,18,20] Finally, residuals of the (natural log of) concentrations of BS and SO_2 were tested for non-normality and spatial autocorrelation in ArcGIS using Moran's 'I'.

Model Evaluation. Models were evaluated by comparing predicted and observed concentrations for the retained 20% of monitoring sites, on a year-by-year basis. Log-transformed concentrations, for direct comparison with measures of fit obtained from model development, and 'real' concentrations (i.e., by exponentiation of the predicted concentration) were evaluated. Model errors were checked for non-normality, and any sites exhibiting unusually large errors were identified. Model performance was assessed on the basis of adjusted R^2, the root-mean-square error (RMSE), and fractional bias (FB). To further evaluate model performance, a 'leave-one-out' analysis (i.e., Jack-knife) was performed whereby each prediction and observation pair are, in turn, left out of the creation of performance statistics and N-1 (evaluation site) R^2 values are returned for each year by pollutant.

RESULTS AND DISCUSSION

Model Development. The LUR models for BS and SO_2, by year, are summarized in Table 1. Detailed information about incremental values of adjusted R^2, p-values, and Standard Error of

Table 1. Summary of BS and SO_2 LUR Models by Year

pollutant	year	model	P (sig.)	Adj R^2	SEE	N
black smoke	1962	0.97 + [7.99e-006 * X] + [9.89e-006 * Y] − [8.56e-012 * X^2] − [8.90e-012 * Y^2] − [2.48e-012 * XY] − [1.42 * other natural areas and agriculture within 1 km] − [6.39e-008 * forest within 3 km] + [2.23e-006 * minor roads and 'B' road length within 3 km]	0.00	0.676	0.354	534
	1971	−1.97 + [1.68e-005 * X] + [7.84e-006 * Y] − [1.78e-011 * X^2] − [6.07e-012 * Y^2] − [3.40e-012 * XY] + [1.75-e-005 * minor roads and 'B' road length within 3 km] + [1.71e-009 * low density residential within 10 km] + [3.15e-005 * major road length within 1 km]	0.00	0.680	0.388	767
	1981	−1.04 + [1.17e-005 * X] + [5.75e-006 * Y] − [1.13e-011 * X^2] − [3.40e-012 * Y^2] − [5.76e-012 * XY] − [2.75e-007 * forest and other natural areas within 1 km] + [5.23e-008 * other urban areas within 10 km] + [1.13e-005 * minor roads and 'B' road length within 1 km]	0.00	0.408	0.431	771
	1991	−5.77 + [2.41e-005 * X] + [1.75e-005 * Y] − [2.00e-011 * X^2] − [1.04e-011 * Y^2] − [1.97e-011 * XY] − [2.29e-009 * forest and other natural areas within 10 km] − [4.32e-008 * other urban areas within 7.5 km]	0.00	0.390	0.434	155
sulfur dioxide	1962	−0.05 + [9.71e-006 * X] + [1.24e-005 * Y] − [7.30e-012 * X^2] − [1.06e-011 * Y^2] − [9.17e-012 * XY] + [1.94e-006 * minor road length within 3 km] + [2.56e-009 * low density urban within 10 km] + [1.54e-007 * high density urban within 1 km]	0.00	0.605	0.350	482
	1971	0.29 + [1.94e-006 * minor road length within 3 km] + [1.16e-005 * X] + [4.77e-006 * Y] − [1.09e-011 * X^2] − [3.77e-012 * Y^2] − [3.54e-012 * XY] + [4.11e-008 * industrial and commercial land within 10 km [(≤3 km * 0.84) + (>3 km and ≤10 km * 0.16]] + [7.11e-006 * major road length within 3 km] + [1.84e-009 * low density urban within 10 km]	0.00	0.649	0.337	733
	1981	0.975 + [8.28e-006 * X] + [3.70e-006 * Y] − [7.64e-012 * X^2] − [2.98e-012 * Y^2] − [3.41e-012 * XY] + [2.10e-009 * low density urban within 10 km] + [3.64e-008 * other urban within 10 km] + [3.95e-008 * high density urban within 2 km]	0.00	0.378	0.404	756
	1991	2.50 + [4.05e-006 * Y] − [4.62e-012 * Y^2] + [4.57e-008 * other urban within 7.5 km] + [1.39e-009 * low density urban within 10 km]	0.00	0.243	0.356	153

Table 2. Performance Statistics from the Evaluation Analysis

		LN concentration		concentration						
pollutant	year	Adj R^2	RMSE	Adj R^2	RMSE	FB	Beta	constant	P(sig.)	N[a]
black smoke	1962	0.640	0.356	0.558	57.1	−0.10	1.12	−1.96	0.000	133
	1971	0.684	0.424	0.413	28.8	−0.09	1.04	3.02	0.000	191
	1981	0.504	0.413	0.382	8.8	−0.09	1.29	−3.54	0.000	193
	1991	0.414	0.518	0.339	5.9	−0.03	0.83	3.09	0.000	37
sulfur dioxide	1962	0.645	0.368	0.706	44.8	−0.04	1.12	−12.17	0.000	121
	1971	0.625	0.317	0.573	28.4	−0.05	0.83	20.92	0.000	183
	1981	0.408	0.345	0.255	18.9	−0.06	0.85	9.20	0.000	189
	1991	0.328	0.386	0.309	12.4	−0.05	1.18	−3.93	0.000	38

[a] The number of independent evaluation sites.

the Estimate (SEE) for each model variable are available in the Supporting Information (Tables S3 and S4 for BS and SO_2, respectively).

With the exception of the 1971 SO_2 model, trend surface explained the highest amount of variation in monitored concentrations. Trend surfaces tended to have higher values of R^2 for BS than for SO_2; for BS, between 66% and 88% of the total explained variation. A third-order (i.e., cubic) trend surface was used in all cases except 1991 for SO_2 where a second-order trend surface (i.e., quadratic) was used. Between two and three variables were obtained in addition to the trend surface depending on the year and pollutant, but population and altitude did not come into any of the models.

To further explore variable selection, analysis was undertaken in ArcGIS to determine the proportion of grid cells where discrete variables (i.e., all those apart from trend) had influence. In other words, the number of cells which had predictions based only on trend and the regression constant. The 1971 black smoke map has 13% of 1 km grid cells without contributions from discrete (i.e., non- trend) variables, but in terms of cells with a population >0 (using the estimated population surface for 1971) this was 0.03% with <0.04% for both pollutants in 1962 and 1971 and <1% in all other cases.

High levels of explained variation (>60%) in monitored concentrations were achieved for both pollutants in 1962 and 1971, with comparably weaker models for later years (Table 1). A low level of clustering (i.e., spatial autocorrelation) was seen in the examination of residuals: values of Moran's I were in the range 0.06 to 0.12 for BS and 0.03 to 0.09 for SO_2, with $p < 0.05$ in all cases. This is relevant information for epidemiological analyses if the exposures are treated as continuous variables but not if the exposure is classified into tertiles or quintiles.

The model variables differ between years for each pollutant. This may reflect changes in the spatial pattern of pollutants and source contributions over time. To explore this, concentrations between the different time periods at common monitoring sites were correlated (see Table S5). Values of Pearson's R were between 0.41 and 0.76 for BS and between 0.26 and 0.83 for SO_2 ($p < 0.01$ except between 1962 and 1991 where $p = 0.201$) but with less than 30 common sites for BS in two comparisons and

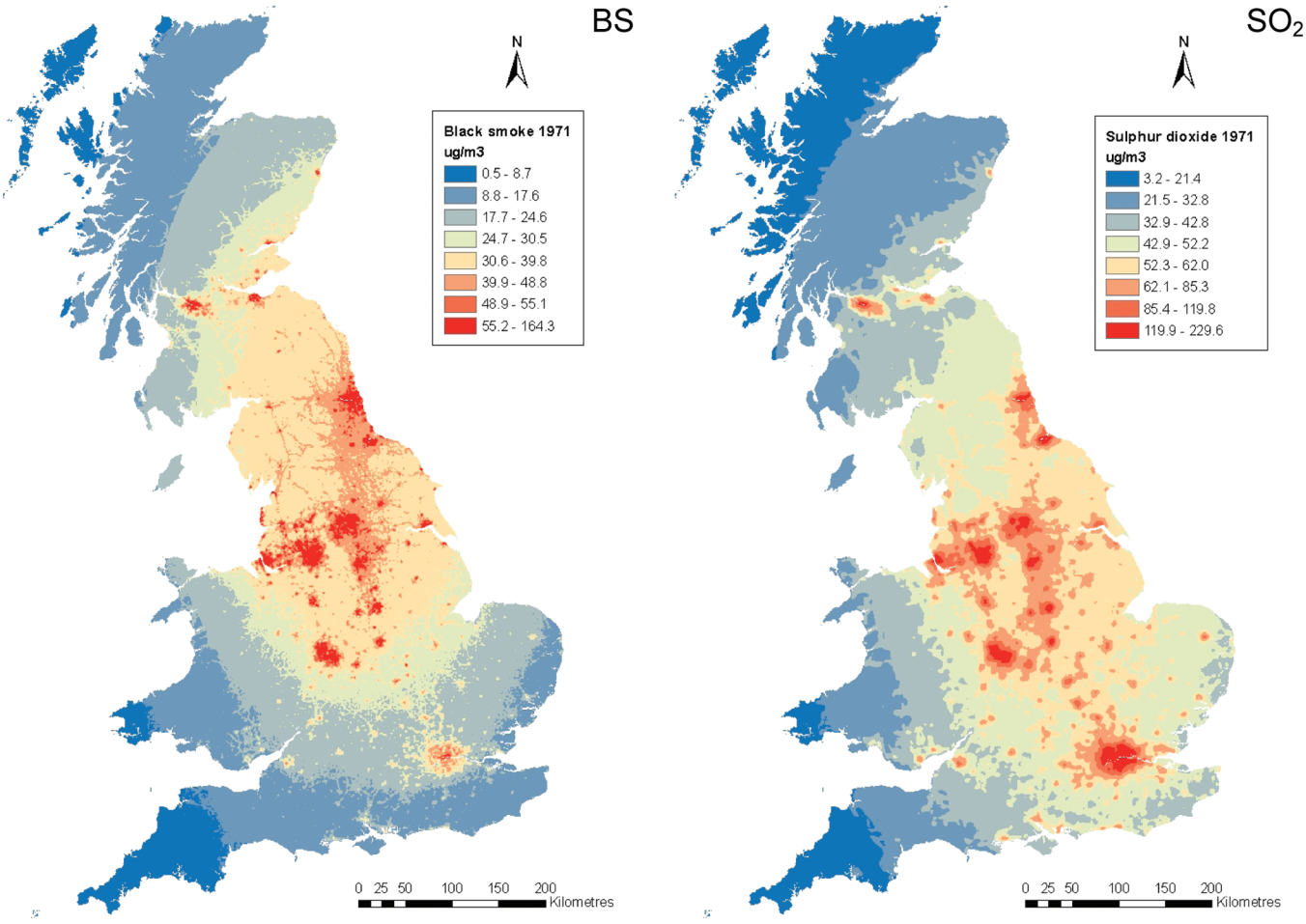


Figure 1. Modeled BS and SO_2 concentrations for 1971.

SO_2 in one comparison. Correlations are generally good, but not high, suggesting that spatial patterns of each pollutant are not exactly the same between the time periods, particularly at the local scale, and that model variables might be expected to change.

Model Evaluation. Results from a model evaluation are shown in Table 2. Scatter-plots of predicted versus monitored concentrations, set against one-to-one lines representing perfect correlation, can be seen in the Supporting Information (Figure S1 and Figure S2 for BS and SO_2, respectively).

Evaluation tests using log-transformed data produced R^2 values broadly similar to those yielded in model building. Models were seen to perform well for both pollutants in 1962 and 1971 with values of R^2 in the range 0.63 to 0.68. Results for other years were good ($R^2 > 0.4$) except for SO_2 in 1991 with a moderate R^2 of 0.33. For the concentrations (i.e., exponentiation of model predictions), overall model performance was weaker. For BS, R^2 values were 0.56, 0.41, 0.38, and 0.34, and for SO_2 were 0.71, 0.57, 0.26, and 0.31 in 1962, 1971, 1981 and 1991, respectively. Very similar performance results were obtained using the 'leave-one-out' analysis (Table S6), but a wider range of R^2 was seen for 1991 than other years reflecting the sensitivity of the test to the smaller number of sites.

Values of RMSE for both SO_2 and BS were broadly similar and proportional to standard deviations in monitored concentrations (see Table S1). Values of FB were negative and relatively small for both pollutants in all years (Table 2) representing consistent slight underprediction of the models.

Model performance, as indicated by values of R^2, is weaker in later years. The weaker models for later years can largely be explained by the lack of variation in monitored concentrations. In 1991, for example, 31% of monitoring sites recorded a BS concentration between 14 and 16 $\mu g/m^3$, and BS was in the range of 10 to 20 $\mu g/m^3$ for 82% of monitoring sites. By the early 1990s a significant proportion of BS emissions in the UK were from traffic-related sources, but the model presented here does not include any road variables. One reason might be the 1 km^2 spatial scale of our model as traffic-related particles have a distance-decay gradient of 100−400 m.[21]

Concentration Mapping. The LUR models were used to map concentrations on 1 km grids across Great Britain for each of the four years. The coefficients (see Table 1) were applied to their respective target variables in the 100 m grids used in model building. The resulting grids were then summed and aggregated to a 1 km grid for mapping. The coefficients for trend surfaces were applied to the x-y coordinates of the geometric centroid of the same 1 km grid, and then all 1 km grids were finally summed and the model constant added to make a single map of concentrations for each pollutant, for each year. Figure 1 shows mapped concentrations for BS and SO_2 for 1971. Maps of concentrations for all years are available in the Supporting Information (Figure S3 and Figure S4 for BS and SO_2, respectively). The maps are on different scales due to the large

differences in concentrations of both pollutants over time, therefore not directly comparable.

To explore spatial patterns, population-weighted exposures were correlated at District level ($n = 406$) between 1991 and each of the earlier years. At this more coarse scale the correlations were higher, as expected: 0.77, 0.80, and 0.91 for BS ($p = .000$) and 0.75, 0.58, and 0.68 for SO_2 ($p = .000$) pairwise between 1991 and 1962, 1971, and 1981. This suggests for BS that spatial patterns of exposure are broadly similar over time but for SO_2 there is more variation, hence the different trend surface in 1991.

Comparisons with Other Studies. This study is the first to use LUR to model concentrations, producing national-level maps for historic air pollution exposure assessment this far back in time. Furthermore, models were developed using contemporaneous monitoring data (rather than extrapolating backward using input data from more recent years), and all models were evaluated, with performance statistics equal to those obtained in models for recent years.[5,18]

The most directly comparable study comes from The Netherlands where LUR methods in conjunction with other GIS-based interpolation methods were used to model historic exposures to a range of pollutants including BS and SO_2 at national scale.[5] Modeling of BS was undertaken for 1985 to 1996 and SO_2 for 1977 to 1996. The methodology was different from the one presented here: a 'regional' component was based on data from the nearest air pollution monitoring site, with separate regression methods modeling an 'urban' component, using land-cover and population, and a 'local' component based on assignment of traffic intensities within a 100-m buffer around monitoring sites. 'Urban' models for BS and SO_2 had one variable each: BS used population within a 1 km buffer, whereas SO_2 sites were categorized as either in a rural area, an urban area, or an industrial area. The models (regional + urban + local) explained 59% and 56% of the overall variation in monitored concentrations of BS and SO_2, respectively. For the urban component, the models explained 49% and 35% of monitored concentrations of BS and SO_2, respectively. The results are similar to those presented here for the 1980s, but better in the 1990s, perhaps because they modeled to a finer spatial scale, thereby better picking up traffic-related variability.

A range of air pollutants (PM_{10}, O_3, SO_2, and NO_X) were modeled in a UK-wide study, going back to the mid-1990s, on a 1 km grid using dispersion modeling techniques. Models were, however, only evaluated against data on monitored concentrations from 2003 and then applied to earlier years using historical data on emissions sources. Models were seen to perform well for all pollutants except for 'background' PM_{10}: models were calibrated on sites from the national monitoring network ($R^2 = 0.25$; $n = 47$) and verified against a separate set of monitoring sites ($R^2 = 0.37$; $n = 20$).[4]

On a regional scale, historic exposures to SO_2 and NO_2 were modeled using a dispersion model for the population of Greater Stockholm from 1955 to 1990. All the available monitoring data were used for model calibration, thus there was no model evaluation.[3] In Newcastle-Upon-Tyne and surrounding area (northeast England), LUR was used to model historic exposures to BS back to the early 1960s (1961 to 1992).[22] The novel approach used included a two-stage, spatiotemporal model, in which first the long-term and seasonal (i.e., within year) trend was modeled, with the spatial covariates handled at a second stage using LUR techniques. The covariates used in the model were chimney count within 500 m, distance to nearest industrial area, binary count of residential or nonresidential land, implementation of the 1956 Clean Air Act (as yes/no), and area of industry within 500 m. Model performance varied considerably across 25 monitoring sites in the study area, but most values of R^2 exceeded 0.5, with about half of the sites having values of R^2 greater than 0.7. Detailed local information e.g. historic land-use derived from aerial photographs had less influence on model performance than inclusion of temporal variables; this approach has since been adapted to model spatiotemporal (i.e., daily) BS concentrations across Northern England between 1985 and 1996 and while their model evaluation yielded R^2 of 0.71, a large proportion of the explained variation came from the temporal variables.[23]

Limitations of the LUR Models. Notwithstanding the generally good performance results, particularly for earlier years, the models as exhibited by the air pollution maps in Figure 1 and the Supporting Information have several weaknesses. First, the trend surfaces may overfit and therefore underpredict concentrations in some areas of the south, west, and northwest due to the relatively small number of monitoring sites in these areas. Likewise, there may be overprediction due to the peaked nature of the trend surface in some rural areas in the north of England. Second, the patterns of air pollution in some years, especially later years, may be overly generalized, due to the low number of land use predictor variables. In 1991, for example, most of the explained variation in SO_2 is from a quadratic, north−south, trend surface. Third, the models could be least accurate in some subregional areas, particularly for the 1962 and 1971 maps, due to effects of misclassification in the land cover and road data sets, which were compiled in the 1990s.

Poorer performance of models in later years provokes discussion of whether such estimates should be used in epidemiological studies. If models cannot be further improved (see below), grouping exposure, e.g. using quintiles, may better reflect the range of exposure across populations. However, associations with health outcomes have been found using exposure estimates with similar performance to our 1991 estimates. For example, dispersion model estimates of PM_{10} with moderate evaluation statistics ($R^2 = 0.25$ for calibration sites and $R^2 = 0.37$ for verification sites)[4] were associated in cross-sectional studies with chronic systemic inflammation[24] and lung function.[25]

Models could potentially be improved with better land use data for the earlier years rather than using that for 1991. This would involve a huge undertaking as data on historic land use patterns would need to be derived from historic maps. The Ordnance Survey produces a digital historic mapping product in the UK, but this is essentially a series of images, from which areas of different types of land would need to be digitized. The UK has some information on traffic composition and land use going back to the 1960s, but these are unfortunately only available for coarse spatial units (i.e., 'regions' - 10 of them in total). As an alternative general indicator of changes in land use patterns we examined population data going back to 1961 (available in relatively small Census units: parishes in the 1960s, wards between the early 1970s and early 2000s). As parish and ward units have different geographical boundaries over time, we translated these to 1 km grids using simple areal weighting in early years (62 and 71) and postcode-weighting in later years (81 and 91). Grid cells were selected with population greater than zero in each year, and pairwise correlations (Pearson's 'R') were produced between all combinations of years. Values of R were in the range 0.76 to 0.99 ($p = 0.000$); values of R were both 0.76 between 1961 and 1981

and 1961 and 1991; for all comparisons not including 1961 R was ≥0.94. As correlations were all either high or very high this suggests that population patterns are broadly similar between the different years, which indirectly supports our assumption that land use patterns have remained relatively constant over time, at least at the spatial scale used in this model.

More recent analysis on land use change looked at percentage changes by land use type across the EU comparing CORINE 1990 with CORINE 2000. For the UK, 1.44% of the total area of land showed a change in land use, with 0.15% of the total related to urbanization, representing relatively small changes.[26]

Future work will look at assessing uncertainties in the current model by creating historic land cover and roads data around a sample of the monitoring sites used here, further investigating traffic-related pollutants, including NO_2, for which data are available from the late 1980s onward, and including use of finer spatial resolution models. On the evidence of some differences in spatial patterns of BS and SO_2 concentrations, further work is also planned to investigate the transferability of the models to intervening years between the 1960s and 1990s.

■ ASSOCIATED CONTENT

Supporting Information. Further, detailed information on the development and evaluation of the LUR models and maps of modeled BS and SO2 for each year: 1962, 1971, 1981, 1991. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +44 (0) 20 7594 5027. E-mail: j.gulliver@imperial.ac.uk. Corresponding author address: Department of Epidemiology & Biostatistics, Imperial College London, St. Mary's Campus, Norfolk Place W2 1PG, U.K.

■ ACKNOWLEDGMENT

The research presented here was developed as part of the Wellcome Trust Intermediate Clinical Fellowship study on Chronic Health Effects on Smoke and Sulphur (CHESS), grant number 075883.

■ REFERENCES

(1) Pope, C. A.; Dockery, D. W. Health Effects of Fine Particulate Air Pollution: Lines that Connect. *J. Air Waste Manage. Assoc.* **2006**, *56* (6), 709–742.

(2) Elliott, P.; Shaddick, G; Wakefield, J; de Hoogh, C.; Briggs, D. J. Long-term associations of outdoor air pollution with mortality in Great Britain. *Thorax* **2007**, *62* (12), 1088–1094.

(3) Bellander., T.; Berglind, N.; Gustavsson, P.; Jonson, T.; Nyberg, F.; Pershagen, G.; Järup, L. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environ. Health Perspect.* **2001**, *109* (6), 633–639.

(4) Stedman, J. R.; Bush, T. J.; Vincent, K. J.; Kent, A. J.; Grice, S.; Abbott, J. UK air quality modelling for annual reporting 2003 on ambient air quality assessment under Council Directives 96/62/EC, 1999/30/EC and 2000/69/EC; *Report AEAT/ENV/R/1790 January 2005*; Didcot, Oxfordshire, AEA Technology, National Environmental Technology Centre. http://www.airquality.co.uk/archive/reports/cat05/0501121424_dd12003mapsrep4.pdf (accessed month day, year).

(5) Beelen, R.; Hoek, G.; Fischer, P.; van den Brandt, P. A.; Brunekreef, B. Estimated long-term outdoor air pollution concentrations in a cohort study. *Atmos. Environ.* **2007**, *41* (7), 1343–1358.

(6) Hoek, G.; Beelen, R.; de Hoogh, K.; 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* (33), 7561–7578.

(7) Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; Morrison, J.; Giovis, C. A review and evaluation of intraurban air pollution exposure models. *J. Exposure Anal. Environ. Epidemiol.* **2005**, *15* (2), 185–204.

(8) Beelen, R.; Hoek, G.; Pebesma, E.; Vienneau, D.; de Hoogh, K.; Briggs, D. J. Mapping of background air pollution at fine spatial scale across the European Union. *Sci. Total Environ.* **2009**, *407* (6), 1852–1867.

(9) Karrpinen, A.; Kukkonen, J.; Elolähde, T.; Konttinen, M.; Koskentalo, T.; Rantakrans, E. A modelling system for predicting urban air pollution: model description and applications in the Helsinki metropolitan area. *Atmos. Environ.* **2000**, *34* (22), 3723–3733.

(10) Vienneau, D.; de Hoogh, K.; Briggs, D. J. A GIS-based method for modeling air pollution exposures across Europe. *Sci. Total Environ.* **2009**, *405* (2), 255–266.

(11) Briggs, D. J.; Collins, S.; Elliott, P.; Fischer, P.; Kingham, S.; Lebret, E.; Pryl, K.; van Reeuwijk, H.; Smallborne, K.; van der Veen, A. Mapping urban air pollution using GIS: a regression-based approach. *Int. J. Geogr. Inf. Sci.* **1997**, *11* (7), 699–718.

(12) Briggs, D. J.; de Hoogh, C.; Gulliver, J.; Wills, J.; Elliott, P.; Kingham, S.; Smallbone, K. A. regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Sci. Total Environ.* **2000**, *253* (1−3), 151–167.

(13) Brauer, M.; Hoek, G.; van Vliet, P.; Meliefste, K.; Fischer, P.; Gehring, U.; Heinrich, J.; Cyrys, J.; Bellander, T.; Lewne, M.; Brunekreef, B. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* **2003**, *14* (2), 228–239.

(14) Gilbert, N. L.; Goldberg, M. S.; Beckerman, B.; Brook, J. R.; Jerrett, M. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. *J. Air Waste Manage. Assoc.* **2005**, *55* (8), 1059–1063.

(15) Arain, M. A.; Blair, R.; Finkelstein, N.; Brook, J. R.; Sahsuvaroglu, T.; Beckerman, B.; Zhang, L.; Jerrett, M. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. *Atmos. Environ.* **2007**, *41* (16), 3453–3464.

(16) Henderson, S. B.; Beckerman, B.; Jerrett, M.; Brauer, M. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environ. Sci. Technol.* **2007**, *41* (7), 2422–2428.

(17) Moore, D. K.; Jerrett, M.; Mack, W. J.; Künzli, N. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. *J. Environ. Monit.* **2007**, *9* (3), 246–252.

(18) Vienneau, D.; de Hoogh, K.; Beelen, R.; Fischer, P.; Hoek, G.; Briggs, D. Comparison of land use regression models between Great Britain and the Netherlands. *Atmos. Environ.* **2010**, *44* (5), 688–696.

(19) Chen, H.; Goldberg, M.; Crouse, D. L.; Burnett, R. T.; Jerrett, M.; Villeneuve, P. J.; Wheeler, A. J.; Labrèche, F.; Ross, N. A. Back-extrapolations of estimates of exposure from current land-use regression models. *Atmos. Environ.* **2010**, *44*, 4346–4354.

(20) Jerrett, M.; Arain, M. A.; Kanaroglou, P.; Beckerman, B.; Crouse, D.; Gilbert, D.; Brook, J. R.; Finkelstein, N.; Finkelstein, M. M. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. *J. Toxicol. Environ. Health, Part A* **2007**, *70* (3−4), 200–212.

(21) HEI Panel on the Health Effects of Traffic-Related Air Pollution. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects; HEI Special Report 17. 2010; Health Effects Institute: Boston, MA, pp 4.18−4.37.

(22) Fanshawe, T. R.; Diggle, P. J.; Rushton, S.; Sanderson, R.; Lurz, P. W. W.; Glinianaia, S. V.; Pearce, M. S.; Parker, L.; Charlton, M.;

Pless-Mulloli, T. Modelling spatio-temporal variation in exposure to particulate matter: a two-stage approach. *Environmetrics* **2008**, *19* (6), 549–566.

(23) Dadvand, P.; Rushton, S.; Diggle, P. J.; Goffe, L.; Rankin, J.; Pless-Mulloli., T. Using spatio-temporal modeling to predict long-term exposure to black smoke at fine spatial and temporal scale. *Atmos. Environ.* **2011**, *45* (3), 659–664.

(24) Forbes, L. J. L.; Patel, M. D.; Rudnicka, A. R.; Cook, D. G.; Bush, T.; Stedman, J. R.; Whincup, P. H.; Strachan, D. P.; Anderson, R. H. Chronic exposure of markers to outdoor air pollution and markers of systemic inflammations. *Epidemiology* **2009a**, *20* (2), 245–253.

(25) Forbes, L. J. L.; Kapetanakis, V.; Rudnicka, A. R.; Cook, D. G.; Bush, T.; Stedman, J. R.; Whincup, P. H.; Strachan, D. P.; Anderson, R. H. Chronic exposure to outdoor air pollution and lung function in adults. *Thorax* **2009b**, *64* (8), 657–63.

(26) Feranec, J.; Jaffrain, G.; Soukup, T.; Hazeu, G. Determining changes and flows in European landscapes 1990−2000 using CORINE land cover data. *Atmos. Environ.* **2010**, *30* (1), 19–35.