Atmospheric Environment 301 (2023) 119688

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Development of historic monthly land use regression models of SO_2, NOx and suspended particulate matter for birth cohort ELSPAC

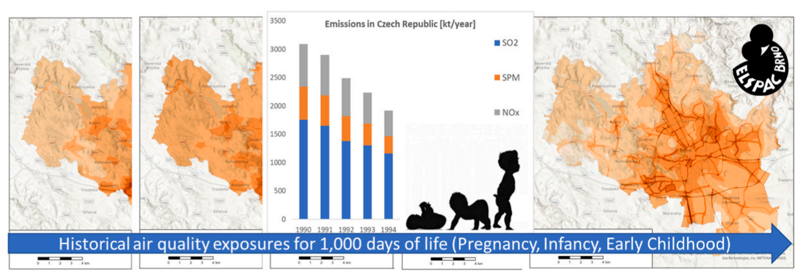
Ondřej Mikeš [a,*], Ondřej Sáňka [a], Aneta Rafajová [a], Jelle Vlaanderen [b], Jie Chen [b], Gerard Hoek [b], Jana Klánová [a], Pavel Čupr [a,**]

[a] RECETOX, Faculty of Science, Masaryk University, Kotlářská 2, Brno, Czech Republic
[b] Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands

HIGHLIGHTS

- Monthly historical LUR models were developed for air pollutants in Central Europe.
- Model performance was moderate.
- Exposures assigned to ELSPAC participants showed temporal and spatial variability.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Land use regression
Air quality
Cohort
Vulnerable windows
Children's health
ELSPAC

ABSTRACT

Vulnerable windows in child development in utero and after birth are critical time points for uncovering the links between environment and health. Particular attention is paid to the first 1000 days of life from conception to the second year of life.

The ELSPAC (European Longitudinal Study of Pregnancy and Childhood) birth cohort, launched in the early 1990s, is a rich source of longitudinal data about health and life events, based mainly in Brno, Czechia. There are currently no air quality concentration maps that can be used to assess exposure to air pollutants for this period of the 1990s in Central Europe. Simply transferring current models to the 1990's is burdened with the error introduced by the temporal change in emission sources and land use of the area. Therefore, Czech air quality monitoring data were used to develop monthly land use regression (LUR) models, which combine collected spatial variables with monitoring data to predict the variation in exposures to pollutants. Monthly pollutant concentrations were regressed against the GIS-based potential predictor variables to develop LUR models, following a supervised forward linear regression, with several predefined constraints.

We constructed 180 LUR monthly models for sulphur dioxide (SO_2), nitrogen oxides (NOx) and suspended particulate matter (SPM) for 1990–1994, that completely cover the first 1000 days for all ELSPAC study participants. The final models showed, on average reasonably good performance (adjusted $R^2 = 0.59$ with hold-out validation (HOV) $R^2 = 0.40$ for SO_2; adjusted $R^2 = 0.75$ with HOV $R^2 = 0.35$ for NOx; and adjusted $R^2 = 0.61$ with HOV $R^2 = 0.31$ for SPM; with a mean number of stations of 74, 38 and 41, respectively). For these models, roads and greenness were predominantly selected as the best predictors.

* Corresponding author. Kamenice 753/5, 625 00, Brno, Czech Republic.
** Corresponding author.
E-mail addresses: ondrej.mikes@recetox.muni.cz (O. Mikeš), pavel.cupr@recetox.muni.cz (P. Čupr).

https://doi.org/10.1016/j.atmosenv.2023.119688
Received 21 October 2022; Received in revised form 26 February 2023; Accepted 28 February 2023
Available online 2 March 2023

The modelled exposures will serve in many subsequent ELSPAC epidemiological studies, but our models may be also used in other Czech and possibly other Central European cities in that period.

1. Introduction

Air quality is considered to be a major environmental factor affecting human health (WHO, 2013). Both short-term and long-term exposures to air pollution are associated with health impacts (Dalecká et al., 2020; Manisalidis et al., 2020). Air pollution has a particularly negative effect on the elderly, children and those with pre-existing health problems (European Environment Agency, 2018a).

Both the magnitude and timing of exposure to air pollution play a crucial role in the development of many diseases (Gulliver et al., 2018; Luo et al., 2021; Wright, 2017). This is of particular importance in developing fetuses (Baïz et al., 2011; Kajekar, 2007; Klepac et al., 2018; Smarr et al., 2013) and newborns, where many organs and mechanisms (including those acting against toxic factors) are developing (Hogg et al., 2012; Kajekar, 2007). Exposure in pregnancy can also alter immune competence and result in asthma or allergies (Baïz et al., 2011). Studies regarding such timing of exposure refer to this period as "vulnerable windows" (Etzel, 2020; Selevan et al., 2000). Particular attention is paid to the first 1000 days of life from conception to the second year of life (Darling et al., 2020). Insight into much more detailed time scale of exposures can help studies that seek to search for disease origins, these exposure estimates should often be as short as months (Wright, 2017).

Models are needed to estimate past exposures, as monitoring generally lacks sufficient spatial resolution. This becomes highly relevant when longitudinal health data from cohort studies need to be evaluated in the context of potential environmental exposures (Chen et al., 2010a; Gulliver et al., 2013; Levy et al., 2015; Muttoo et al., 2018).

Land use regression (LUR) models are used in many areas, usually with outputs in annual concentrations in different cities and in spatial resolutions, allowing researchers to compare participants' exposures in epidemiological studies (Gulliver et al., 2011; Hoek et al., 2008, 2015). LUR models are developed using air quality data from a limited number of monitoring stations in association with surrounding geospatial characteristics as potential predictor variables (i.e. population density, road classes, usage of the land) estimated using geographic information systems (GIS) (Beelen et al., 2013; Eeftens et al., 2012). The models are then applied to locations (e.g. residential addresses), where no measurements are available (Hoek et al., 2015). LUR models are relatively robust, with lower data requirements than dispersion models (Hoek et al., 2008).

Several studies have used back-extrapolation of the most recent LUR models to cover earlier measurement periods (Chen et al., 2010a; Eeftens et al., 2011; Wang et al., 2013). However, extrapolations relying only on past-to-present ratios may not be valid for earlier years, as may correlations between population, road density and pollutants, and the distribution of emissions may change significantly over time.

LUR models have been used to study the association between birth outcomes and maternal air exposure in many recent studies (Ahmed et al., 2022; Bettiol et al., 2021; Luo et al., 2021; Manisalidis et al., 2020). Unfortunately, most studies have been conducted in Europe, North America, Australia, and a few in East Asia. All European studies are exclusively from Western/Southern/Northern Europe and none from Central or East Europe.

In 1991–1992 a prospective birth study was conducted as part of the European Longitudinal Study of Pregnancy and Childhood (ELSPAC) in the districts of Brno and Znojmo (Piler et al., 2016). A proper environmental assessment for this birth cohort is still lacking.

The aim of this study is to develop spatially and temporally detailed historical LUR models for sulphur dioxide (SO_2), nitrogen oxides (NOx) and suspended particle matter (SPM) covering 1000 days of life for all ELSPAC children. We focus on the city of Brno, where most participants lived and where at least some monitoring stations were located. Models for SO_2 and SPM are often missing in the literature, and these pollutants were the ones measured at the most stations in the early 1990s. This research addresses a critical gap in our understanding of the impact of air quality on acute and long-term children's health in the understudied Central European region. By providing comprehensive estimates, it will be significant resource for future assessments of air pollutant exposures and related health impact analysis of ELSPAC cohort. In addition, the results of this study may apply to even broader Central European region than this study.

2. Methods

The Czech part of the European Longitudinal Study of Pregnancy and Childhood (ELSPAC) is one of six prospective birth cohort studies initiated by the World Health Organization (WHO). In the Czech Republic, the target population was all eligible pregnant women who were residents of two South-Eastern districts of the country – Brno and Znojmo – and who were due to deliver between March 1st' 1991 and June 30th' 1992 (Piler et al., 2016). At the outset 5151 mother-child pairs consented to participate out of 7589 births in the study regions. The cohort has so far been assessed for diet, household chemicals, and life stress events (Grulichová et al., 2020; Mikeš et al., 2019; Stepanikova et al., 2021), but a proper environmental assessment is still lacking. The period of our research overlaps with the socio-economic transformation of the Central European region that occurred after the end of the communist regime. Post-communist Central European countries were transitioning from central planning to market economies. Decrease in industrial productivity on the order of tens of percentages and the introduction of the first air protection laws also dramatically altered pollutant emissions in the region (Fig. S1) (Kunc and Tonev, 2022).

In contrast to the more common annual models, we decided to develop monthly LUR models because of the application to the birth cohort. Other reasons for developing monthly models were the expected large seasonal differences in emissions from different sources in a given period.

2.1. Monitoring stations and study area

Air quality monitoring in the Czech Republic started in the late 1960s with SO_2 and SPM. Since the 1990s, NOx has also been measured. The number of air quality (AQ) monitoring stations changed from year to year and by pollutant measured. The methods predominantly used were UV-fluorescence, West-Gaek spectrometry and coulometry for SO_2, chemiluminescence and spectrophotometry-triethanolamine method for NOx, and gravimetry and radiometry for SPM.

Monthly averaged concentration data were obtained for selected monitoring stations present in the period 1990–1994 for the Czech Republic from Czech Hydrometeorological Institute (CHMI) (Fig. 1). As there were insufficient stations in Brno and close surroundings, we initially included all Czech stations, and then filtered the stations according to three selection criteria to derive a model relevant to the Brno urban area and to improve the specificity of the model. First, stations in urban and sub-urban areas, that are relevant to our LUR model for larger cities were selected. Rural stations were also included if located in the South Moravian Region or officially assigned as Czech background monitoring stations to increase the contrast for modelling and provide a regional background. The second criterion originates from the situation in the 1990s in the Czech Republic (then Czechoslovakia), where the coal mining industry and associated coal processing facilities were a large source of SO_2 and SPM pollution (see Fig. S1). Coal mining was predominantly present in the northwest and northeast of the country (marked in shading on Fig. 1). Stations from these areas were excluded from our analysis, as they represent an exposure scenario not relevant for our cohort. Third, months where stations had less than 75% valid measurements were also excluded from the analysis. In total, 181 stations were used where at least one valid monthly average value was available. The position of the monitoring stations and the percentage of valid monthly averages in the area of interest is summarized in Fig. 1. A basic description of monitoring stations is provided in Table S1.

2.2. Geographical and meteorological data

We have collected the maximum available data, used in LUR models, for the period in question, although this is quite challenging with 30-year-old data.

All spatial data derivations were processed in ArcGIS Pro software (ver. 2.5.0 and later) in predefined buffers for each monitoring site. Buffers were constructed taking into account known dispersion patterns (Beelen et al., 2013; Eeftens et al., 2012; J. G. Su et al., 2009). First, 1990 estimates from the Global Human Settlement database in a spatial resolution of 250 × 250 m were used to obtain population and percentage of the built-up area. As there was insufficient local spatial data, we had to use this global dataset. The data were resampled to a 10 × 10 m raster and converted to points. Buffer statistics were calculated for each point (according to Table 1). The total population was calculated in buffers of 100, 300, 500, 1000 and 3000 m around each site (Florczyk et al., 2019).

Landcover was obtained from the CORINE database (European Environment Agency, 2018b); we used Corine Land Cover (CLC) 1990 (Version, 2020_20u1) as the most appropriate for our study. To reduce the number of classes we merged some similar categories (i.e. *continuous urban fabric* and *discontinuous urban fabric* into *urban area*). The following land cover areas were derived: Urban, Industrial, Ports and Airports, Mines and dumpsites, Urban green, Nature, Agriculture and Water. Intersection areas were calculated for each buffer separately. The surface area (in m^2) of each land use was calculated in buffers of 100, 300, 500, 1000 and 5000 m.

The 2016 digital elevation model (ARCDATA, 2016) was used, following the reasonable assumption that terrain does not change significantly over the years. The mean elevation was extracted for each buffer to reflect the effect of the altitude on the results, as well as the range of heights, which can serve as a proxy for terrain roughness.

While traffic intensities data are generally more desirable in LUR, this feature was not available for all monitoring stations in our study. In the absence of traffic data, several LUR studies have successfully used the length of specific road types without traffic intensity data (Ahmed et al., 2022; Novotny et al., 2011; Wang et al., 2013). Road data were not available in a reliable digitised historical map and thus we have used Open Street maps (OSM) and a historic road atlas (1990) and manually updated the road network around monitoring stations and in the city of Brno from 2020 OSM to the state as in 1990. All the motorway, primary and trunk class roads with their links were included in the first class to include fast, national, and international connection roads. The second class was defined to include important and quite heavily trafficked roads (in OSM depicted as secondary, secondary link, tertiary, tertiary link). The third class included more minor roads including categories of living street, residential, service roads, and tracks grade 1–5. We calculated the length of all road types within specific buffers. The influence of the road segment on the surrounding air concentrations often follows an exponential decrease with increasing distance (Zhu et al., 2002). Therefore, we have included the inverse distance and the quadratic inverse distance to the nearest road as variables.

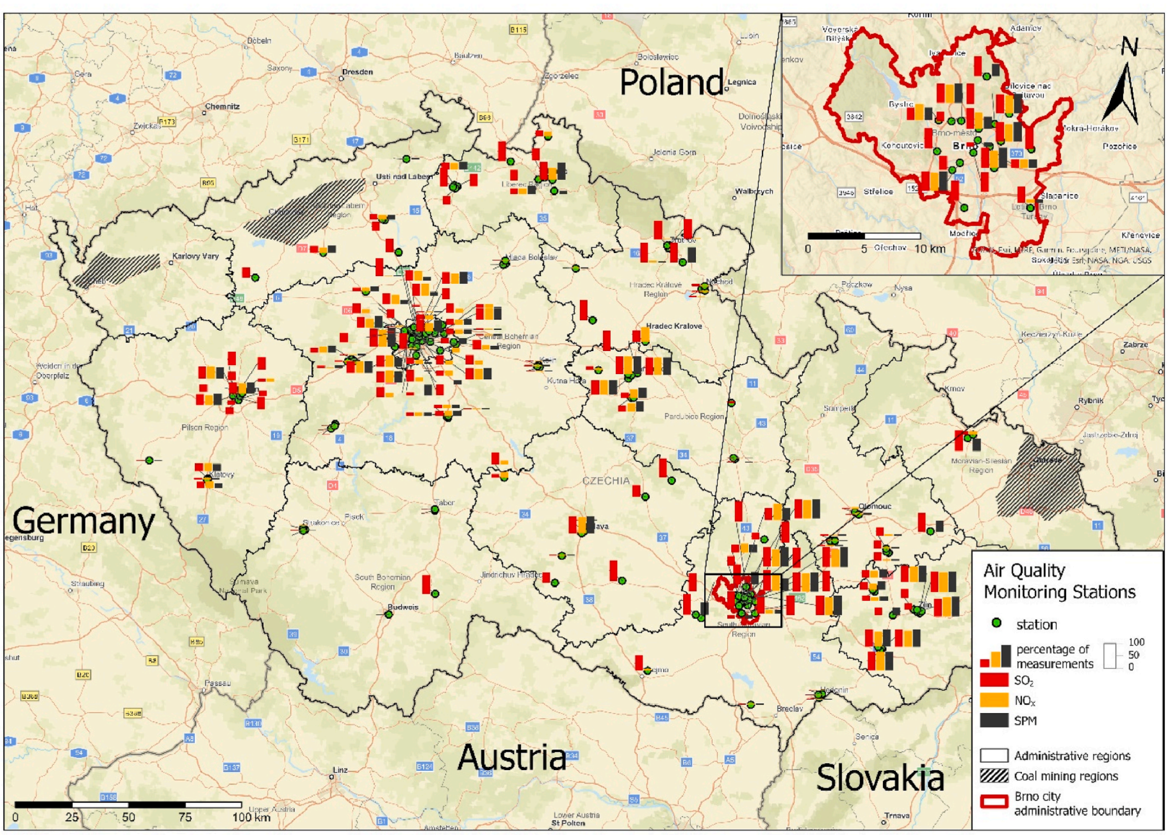


Fig. 1. Map of all monitoring stations from Czech Republic included in the LUR model building for years 1990–1994. In red, city of Brno is depicted where ELSPAC cohort is located. The bar graphs show the percentage of measurements that each station contributed to the models. Major coal regions excluded from the study are depicted as dashed polygons.

Table 1
Variables used in the modelling with description, time characteristics, units, source of the data and defined buffer sizes.

Name	Type of the data	Time	Variable	Units	Type/ Resolution	Source	Reference	Time access	Buffers (meters)
Global human settlement	Population and building density	1990	Population density Density of inhabitants by square kilometre	Absolute number of inhabitants Pop/km^2	Raster/ 250mx250m	https://ghsl.jrc.ec.europa.eu/download.php?ds=bu	Florczyk et al. (2019)	19.12.2020	100, 300, 500, 1,000, 3000
			Built area	0–100%	Raster/ 250mx250m	https://ghsl.jrc.ec.europa.eu/download.php?ds=bu	Florczyk et al. (2019)	19.12.2020	100, 300, 500, 1,000, 3000
Digital elevation model	Elevation	2016	Mean elevation Elevation range	m above sea level	Raster/ 50mx50m	https://www.arcdata.cz/produkty/geograficka-data/arccr-500	(ARCDATA, 2016)	19.12.2020	100, 300, 500, 1,000, 3000
Corine Land Cover	Land use data	1990	Urban Industrial Port Mines and dumpsite Agricultural Urban green Nature Combination of Urban green and Nature Water	m^2	Vector/ accuracy 100 m	https://land.copernicus.eu/pan-european/corine-land-cover/clc-1990?tab=metadata	European Environment Agency (2018b)	10.1.2021	100, 300, 500, 1,000, 5000
Open Street Map with modifications from historical road atlas	Road classes	2020, 1991	Length of road classes Inverse distance to the nearest road Squared inverse distance to the nearest road	meters m^{-1} m^{-2}	Vector	https://www.geofabrik.de/data/download.html		9.6.2020	25, 50, 100, 300, 500, 1000
Coordinates	Longitude, Latitude in WGS84 coordination system	2020		Decimal degrees	–	https://www.arcdata.cz/produkty/geograficka-data/arccr-500	(ARCDATA, 2016)	19.12.2020	–
Monitoring stations	Concentration data	Monthly/ 1990-94	SO_2 NOx SPM	μg/m^3	Points	www.chmi.cz		19.12.2020	–
Meteorological data	Meteorological conditions	Monthly/ 1990-94	Temperature Humidity Wind speed	°C % m/s	Points	http://portal.chmi.cz/historicka-data/pocasi/mesicni-data/mesicni-data-dle-z.-123-1998-Sb		10.1.2021	–

For the monthly meteorological data, we used the official monitoring stations closest to the air quality station. If the data from the nearest station were not reported, the next closest meteorological station was used. The locations of all monitoring stations have been checked and converted to the World Geodetic System 1984 (WGS84). All sources and resolutions of parameters and measurements used are summarized in Table 1.

2.3. Model development and evaluation

Model development was performed in the program RStudio (R version 4.1.0 (2021-05-18). Average monthly pollutant concentrations were regressed against the GIS-based potential predictor variables to develop LUR models, following a supervised forward linear regression approach described before (Beelen et al., 2013; Chen et al., 2019; Eeftens et al., 2012; Muttoo et al., 2018). Considering the nature of the variables, we predefined negative slopes for the following potential predictor variables: urban green, agricultural areas, natural spaces (other than urban green), water bodies, temperature, and wind speed. Firstly, a univariate linear regression model was built for each potential predictor variable to find the best predictor that explained the maximum variance in the measurements. Secondly, at each subsequent step, the significant predictor variable ($P < 0.1$) that added most to the model adjusted coefficient of determination (adjusted R^2) was added to the model. Predictors only entered the model if they adhered to the pre-defined direction of effect (e.g., positive slopes for roads and negative slopes for nature spaces). Predictor variables were added to the model until the model adjusted R^2 could not be increased anymore. In the final step, predictor variables with a variance inflation factor (VIF) larger than three were removed from the model to avoid multicollinearity. All final models were checked for influential observations (Cook's $D > 1$) and residuals of the concentrations were tested for spatial autocorrelation using Moran's I to assess the independence assumption. The model was re-run using all but the influential observation(s). If the model coefficients changed largely (>5%), we retained the new model (Sanchez et al., 2018).

Five-fold hold-out validation (HOV) was performed to evaluate each model. The full set of measurements was randomly divided into five groups (20% each). Five additional HOV models were developed following the same process described above, each based on 80% of the monitoring stations, with the remaining 20% for validation. HOV R^2 and RMSE were computed by comparing the stacked predictions at the five HOV test sets to the corresponding measurements.

The resulting models were adjusted to concentrations measured directly in Brno in a given month to best fit the exposures in the ELSPAC study; this method was also applied in other LUR models (Beelen et al., 2013).

The above-described regression models allowed the derivation of estimated air concentration values for a rectangular grid with 25 m resolution, covering the area of Brno. All steps related to the extraction and calculation of variables for regression modelling, as well as visualization of results in the form of maps, were done using ArcGIS PRO software.

To finalize the map bases for non-urban areas (e.g. forests, water reservoirs) which do not correspond to the scope of model use, concentrations from the rural monitoring station were used. None of these extrapolated values are related to the addresses of our participants. Background data were also used in other studies to estimate non-urban locations (Beelen et al., 2013; Gulliver et al., 2018; Muttoo et al., 2018).

3. Results

3.1. Air monitoring data

181 stations in total (Table S1) were included in the analysis based on the selection criteria described above (number of valid measurements, outliers) within a given period. The number of stations with valid data used to develop specific monthly models was on average highest for SO_2 (74) than for SPM (41) and lowest for NOx (38).

The mean monthly values ranged from 10.73 μg/m^3 to 102.25 μg/m^3 for SO_2, from 17.79 μg/m^3 to 67.60 μg/m^3 for NOx, and from 41.90 μg/m^3 to 117.52 μg/m^3 for SPM. For all pollutants the highest concentrations were within the heating season between November and March, peaking in February. This was most evident for SO_2, as this pollutant is most associated with coal burning for heating. The highest measured average monthly concentrations in the study period were 243.61 μg/m^3 for NOx, 309.13 μg/m^3 for SO_2 and 228.65 μg/m^3 for SPM. The monthly average measured concentrations from the selected stations in the study period are summarized in Fig. 2. Detailed statistics for all stations included in our analysis are summarized in the Supplementary materials (Table S1).

3.2. Modelling results and validation

In total, 180 models were constructed (5 × 12 months for three pollutants). All models are summarized in the Supplementary material (Table S2). The explained variability (average adjusted R^2) of models were 0.59 (SD = 0.10) for SO_2, with an average HOV R^2 = 0.40 (SD = 0.12); 0.71 (SD = 0.13) for NOx, with an average HOV R^2 = 0.35 (SD = 0.12) and 0.61 (SD = 0.13) for SPM, with an average HOV R^2 = 0.31 (SD = 0.12). Five-fold HOVs showed the highest stability for SO_2>SPM > NOx. The lowest stability was observed in particular where a low number of stations was used. In instance where there was insufficient monitoring data available for particular month, we were unable to construct a valid model. In such cases, we utilized a model from the following year and the same month. Such model was adjusted to the meteorological conditions and monitoring data specific to that month and year. This approach was not required for SO_2; however, it was necessary for NOx for 15 months, 12 of which were in 1991; and for SPM for 19 months, 12 of which were also in 1991. Notably in 1991, NOx was only monitored on average in 50% of monitoring stations compared to other years and in the case of SPM it was 70%. This recalculation consistent with other studies such as the SAVIAH (Small-Area Variations In Air Quality and Health) study, where models were even transformed from different sites (Briggs et al., 2000), or with other European LUR models (Beelen et al., 2013; Eeftens et al., 2012). Models with low station numbers were generally less robust, as can be seen from the HOV results.

All model performances are in detail described in Table S3; here yearly summary descriptive can be found in Table 2.

RMSE values were on average 10.22 μg/m^3 (SD = 6.18) for SO_2; 10.15 μg/m^3 (SD = 4.69) for NOx and 9.29 μg/m^3 (SD = 4.46) for SPM. The Moran's I spatial autocorrelation in residues was generally low and insignificant with just a few exceptions for several SO_2 models and one for NOx and SPM. Multicollinearity was observed within the modelling process and some of the models were recalculated based on high VIF. After recalculations, all final models had VIF<3.

The number of predictor variables in these models ranged from 2 to 9. Of the total number of the variables used in all models (n_{total} = 1022) roads and greenness variables were most frequent (27% and 25.8% respectively), followed by terrain and coordinates (13.5 and 10.3% respectively), and population characteristics (8.7%). In terms of buffer size, road length in 300 m was most frequently applied variable (33 times for 2nd class roads and 25 times for 1st class roads); for greenness, the 5000-m buffer (91 times urban green and 41 times the combination of urban green with nature areas); for terrain, the range of elevations in a radii of 3000 m (42 times); and in 100 m buffer (24 times) for urban fabric land use variable. Some land use variables like agriculture or mines and dumpsites were rarely present as a significant variable in any models. The distribution of grouped variables is summarized below in Fig. 3.

3.3. Estimated air concentrations in study area

The models for the study area were constructed in a 25 m grid. Brno and its surroundings cover an area of about 453 km^2, which translates into around 730,000 grid cells. Grid centroid was used for characterisation of LUR parameters. The results are displayed here as annual averages for all pollutants (Fig. 4).

Air concentrations corresponding to all individual addresses in the pre- and post-natal life periods were spatially joined with calculated 25 m grid concentration maps. Air concentrations for the residences of ELSPAC study participants are summarized in Table S4 and displayed in Fig. S3. Concentrations and their range were highest in winter months. The results also show that the distribution of concentrations in the study area varied depending on the month.

4. Discussion

We were able to construct monthly average LUR models for SO_2, NOx and SPM for the years 1990–1994 in the Czech Republic to link to a birth cohort study. The final models on average showed fairly good performance (adjusted R^2 = 0.59 and HOV R^2 = 0.40 for SO_2; adjusted R^2 = 0.70 with HOV R^2 = 0.35 for NOx; and adjusted R^2 = 0.61 with HOV R^2 = 0.31 for SPM.

We found a sizable drop in HOV performance compared to training model R^2. In a previous systematic comparison of the difference between model and HOV R^2 (Wang et al., 2012), this difference was strongly influenced by the number of sites. Specifically, for 36, 48 and 72 sites the drop in R^2 was 19, 18 and 13% respectively. Consistently, the drop in R^2 in our study was lower for SO_2 for which more sites were available than for NOx and SPM. The drop in R^2 in our study was modestly higher than the difference reported by Wang et al. (2012), where the NO_2 concentrations were based on four weekly samples in the year 2007. The larger drop may be related to the period and development of monthly models. We did not find any papers on monthly models in the early 1990s to test that hypothesis.

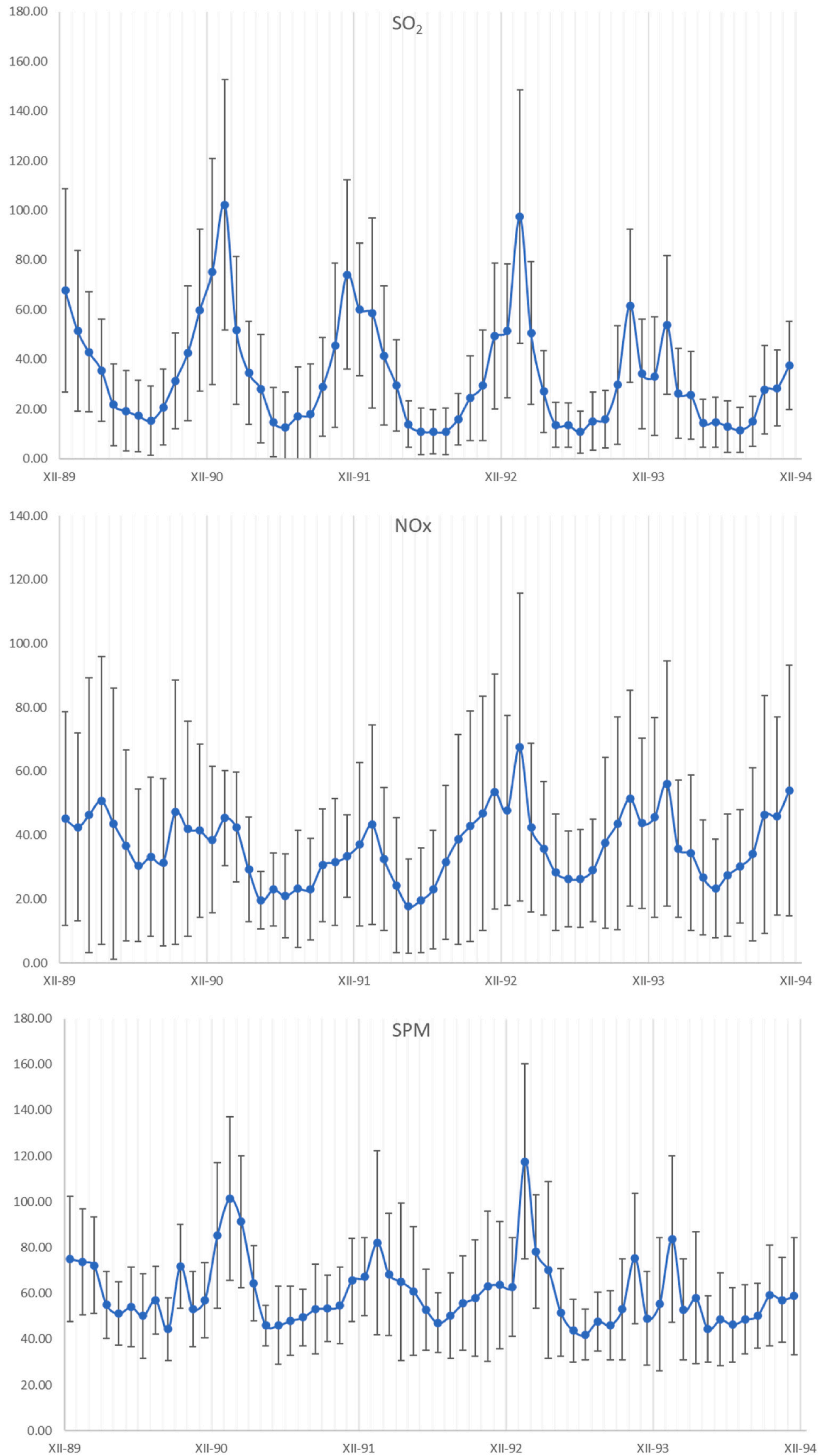


Fig. 2. Monthly mean from monitoring stations in μg/m^3 and standard deviations for the whole study period. On the x-axis only Decembers (XII) are labelled for the sake of clarity.

Table 2
Summary description of model performances.

SO_2	mean adjusted R^2	mean HOV R^2	range of R^2	range of HOV R^2	mean N	range N
1990	0.598	0.334	0.397–0.727	0.198–0.525	72	31–91
1991	0.618	0.382	0.405–0.811	0.148–0.636	65	33–76
1992	0.687	0.458	0.454–0.831	0.245–0.598	72	60–88
1993	0.628	0.452	0.405–0.735	0.202–0.614	81	66–119
1994	0.591	0.402	0.400–0.705	0.205–0.529	79	66–95
NOx						
1990	0.697	0.328	0.368–0.843	0.113–0.571	36	24–56
1991[a]	0.808	0.309	0.516–0.951	0.187–0.567	29	20–43
1992	0.808	0.309	0.516–0.951	0.187–0.567	29	20–43
1993	0.731	0.362	0.539–0.919	0.210–0.737	43	29–69
1994	0.719	0.422	0.577–0.816	0.289–0.589	53	42–80
SPM						
1990	0.634	0.243	0.390–0.687	0.104–0.624	39	27–51
1991[b]	0.637	0.311	0.390–0.715	0.111–0.624	43	35–52
1992	0.605	0.303	0.278–0.715	0.111–0.624	45	35–64
1993	0.716	0.367	0.278–0.949	0.160–0.526	36	31–64
1994	0.727	0.315	0.478–0.882	0.164–0.395	40	27–58

Mean adjusted R^2-mean squared Pearson correlation of the adjusted model, HOV- fivefold hold-out validation, N- number of monitoring stations used for modelling.
[a] models for NOx 1991 are taken from 1992.
[b] models for SPM 1991 are taken from 1992 and 1990 for March.

4.1. Comparison with previous studies

The performance of our models is broadly comparable to other studies. Performance statistics are difficult to compare across studies as the validation method differs (e.g. internal or external). Furthermore, our models are monthly models and most previous models were annual average models. Finally, our models were developed for a period for which it was challenging to obtain relevant predictors and monitoring data. LUR models have not been widely developed for SO_2. In the published studies R^2 was between 0.24 and 0.88 (Amini et al., 2014; Gulliver et al., 2011; Huang et al., 2017; Poirier et al., 2015), and with validation R^2 0.26–0.71 where described. While our R^2 ranged between 0.42 and 0.84 (mean 0.59) with mean HOV R^2 0.40. Of these models, only the Gulliver et al. (2011b) model dealt with a similar period, namely the annual model for 1991, where the adjusted R^2 was 0.24 and the explanatory parameters were coordinates and variables describing the degree of urbanization. NOx is modelled less often than NO_2, with R^2 between 0.63 and 0.91 (Giorgis-Allemand et al., 2017; Jason G. Su et al., 2009; Wolf et al., 2017), and with validation R^2 0.31–0.88 where described in the study; in our case R^2 was 0.44–0.96 with a mean value of 0.71 and a mean HOV R^2 of 0.35. PM_{10} is a fraction of particle matter in aerodynamic diameter smaller than 10 μm, for which some LUR models were also developed with an R^2 ranging between 0.42 and 0.90 (Amini et al., 2014; Dadvand et al., 2013; Giorgis-Allemand et al., 2017). In our study, R^2 for SPM, which also includes bigger particles than PM_{10}, ranged from 0.31 to 0.96 with a mean value of 0.61 and a mean HOV R^2 of 0.31. To our knowledge, only two studies developed an SPM LUR models. First study with R^2 0.11 and the correlation between predicted and observed SPM concentrations of 0.25 (Kashima et al., 2009); in a second study, R^2 was 0.55 and LOOCV R^2 0.44 (Fukuo et al., 2018).

4.2. Air pollution changes in studied period

One of our challenges was to assess exposure levels in a transition period in Czech Republic with fewer data on concentrations and key sources than nowadays. The air pollution levels experienced in the Czech Republic in the 1990s were far higher than at present. Mean SO_2 concentrations ranged between 10.73 and 102.25 μg/m^3, while the highest yearly mean concentration in 2020 from a station located in a heavy industry area was 8.8 μg/m^3. Mean NOx concentrations ranged between 17.79 and 67.60 μg/m^3, whereas nowadays monthly mean concentrations range from 2 μg/m^3 in rural areas to 90.2 μg/m^3 of NOx at a hot spot traffic station. The SPM monthly mean distribution was 41.90–117.52 μg/m^3. SPM is the fraction of aerosols sampled with high-volume samplers and it has a diameter of approximately <50–100 μm and thus cannot be directly compared with current measurements of PM_{10} (EEA, 1996). The range of the PM_{10} mean concentrations in the Czech Republic in 2020 was 7–33 μg/m^3, where the highest mean values were measured on the industrial stations, which were not included in our modelling (CHMI, 2021).

The two most important anthropogenic emission sources of SO_2, NOx and SPM are typically transport and fossil fuel combustion (EEA, 2018). Although some efforts to reduce vehicle emissions in Europe started as early as 1970 with Directive 70/220/EC, a more coordinated emissions reductions started in 1993. Gradually, "Euro" emissions standards for NOx and particles, and other compounds were tightened. When sulphur-containing fuels are burned, sulphur, as a natural component of oil and coal, is released as SO2. However, the Directive on the sulphur content of certain fuels only came into force in 2001 (EEA, 2010). As shown in Fig. S1., there was a dramatic change in air pollutant emissions in the Czech Republic during the 1990s. Firstly, this was due to the economic transition towards a market economy, concurrent with a decline in heavy industry (Kunc and Tonev, 2022). Secondly, the law on the regulation of emissions from large sources was adopted in 1991 and has been in force since 1998. Thus, the 1990s had a gradual reduction in emissions from larger emission sources (CHMI, 2021). All participating children in ELSPAC study were born in 1991 and 1992.

Because of these temporal changes in sources, we could not use simple back-extrapolated models. Back-extrapolated models assume a broadly stable spatial pattern with only the absolute level changing (Chen et al., 2010b; Knibbs et al., 2018). Most of the models contained greenness and road variables, as expected based on the similarities to other LUR models developed previously (Gulliver et al., 2013; Hoek et al., 2008; Saucy et al., 2018), but differed among years and months for each pollutant. These differences may reflect changes in the spatial patterns of pollutant levels and source contributions over time. This highlights the need to build these models for specific months and years to improve the accuracy of exposure estimates. A simple historical back-extrapolation may not be adequate, especially in periods of broad socioeconomic changes.

4.3. Predictor variables

Predictor variables were generally consistent with previous LUR models (Hoek et al., 2008). One or more road variables occurred in almost all models constructed for NOx and SPM, but somewhat less so in the models built for SO_2. Transportation was a significant source of emissions in the 1990s, but to a somewhat lesser extent than industrial

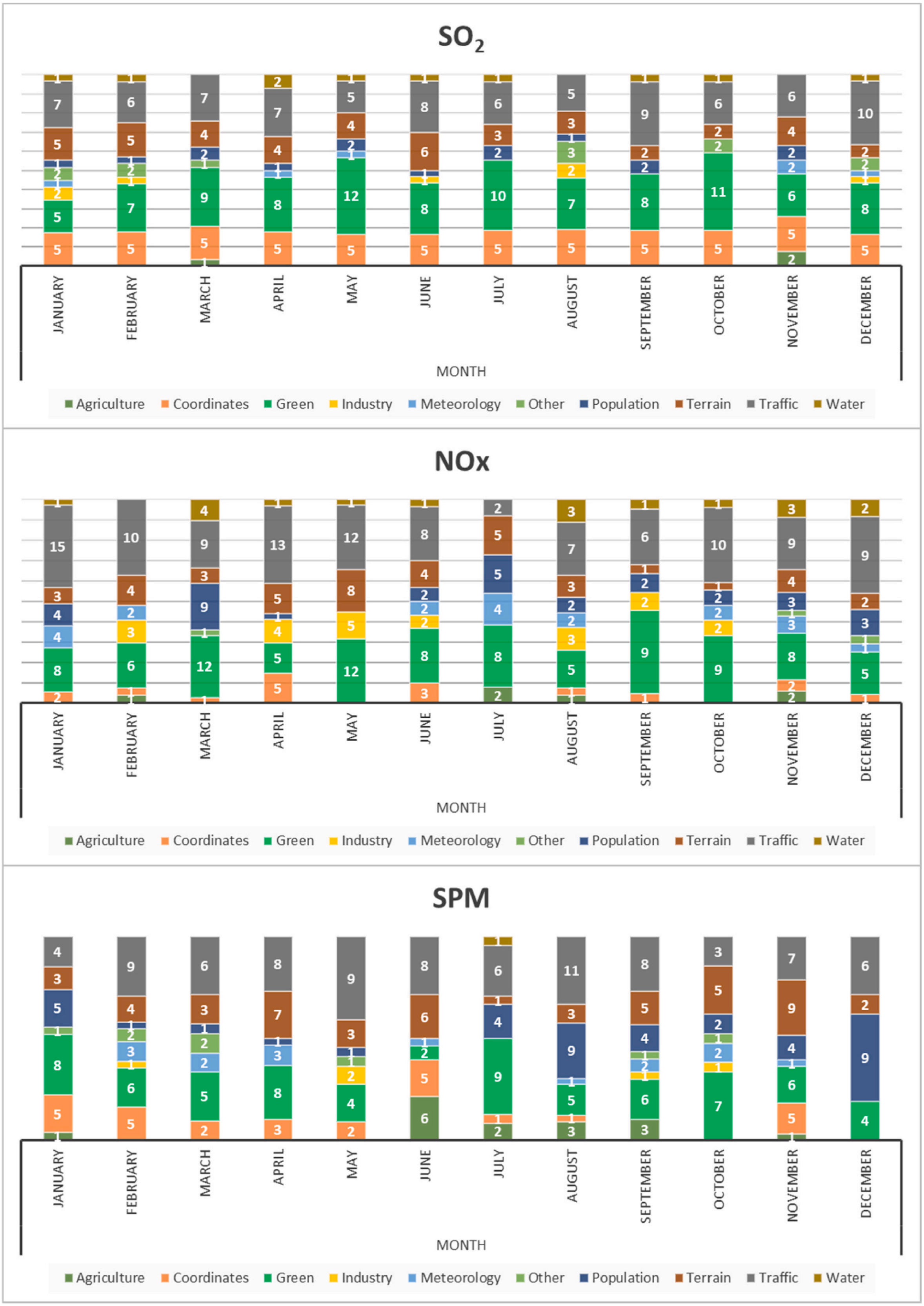


Fig. 3. Number of occurrences of each category of variables in the LUR models for each pollutant and month.

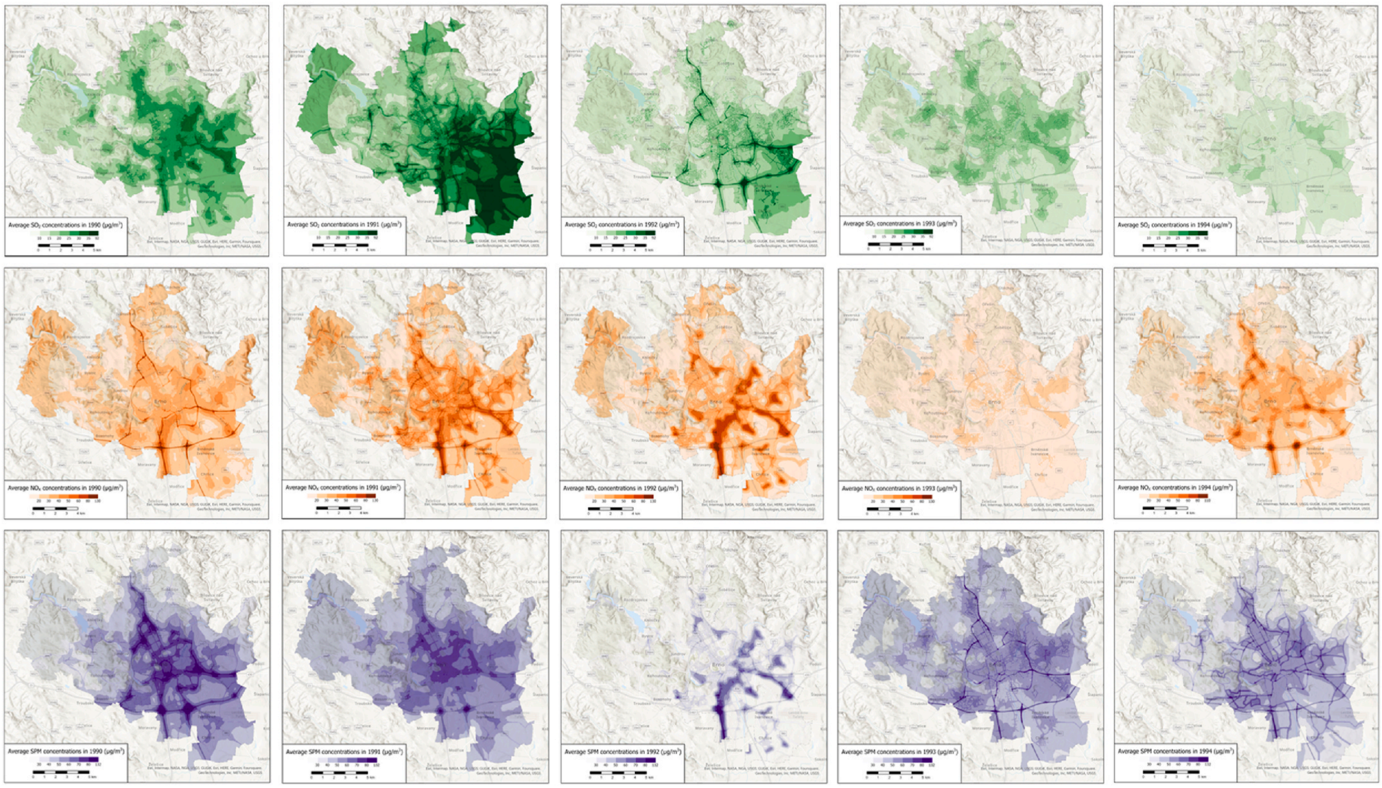


Fig. 4. Final maps in 25 × 25 m resolution for annual average concentrations modelled for all pollutants. From top to bottom SO_2 (in green), NOx (in orange) and SPM (in blue). From left to right 1990–1994 (μg/m^3).

and residential sources.

Greenness in the cities can serve to dilute pollution or simply represents a lack of anthropogenic emission sources. All greenness-related variables were used in many models quite uniformly distributed between pollutants and months, although slightly more often in the growing season. Variables representing green spaces were most common in the 5000-m buffers, likely reflecting the broad influence of greenery on the surrounding area.

Latitude, with predicted concentrations declining from north to south, was included in 99 models. We attribute this to the fact that the areas with the highest emissions were in the northern parts of the Czech Republic. Monitoring stations within highly industrialized localities were excluded from our analysis, but the north/south distribution was still observed in our data. Latitude appeared in all SO_2 models. SO_2 main emissions are primarily from coal-burning facilities, typically situated close to coal mines. Latitude does not have a significant effect on the modelled concentrations in our study area (Brno) but is important for the potential application of our models in other urban areas in the Central European region.

Topography has been identified as a significant predictor in some LUR models, (Giorgis-Allemand et al., 2017; Schembari et al., 2015), and was also found to be important for our models as well. Terrain variables were most often used with the largest 3000-m buffers for all pollutants, both for average elevation and for the range of altitudes, which better defines the overall hilliness of the surrounding terrain.

Housing and population densities are important parameters in many LUR models (Giorgis-Allemand et al., 2017; Wolf et al., 2017). Population-related variables were most frequently present in SPM (41 cases) followed by NOx (33) and relevant for SO_2 in only 15 cases. Urban land use as a variable was most common population description variable. We expected the variables describing population to be more strongly associated with the heating season, but this was not the case in all months. However, it is also possible that more populated areas, while needing more heating, also have a more frequent central heat supply, which is the case, for example, in the study area of Brno.

Industry can be an important source of air pollution and it has been used as a LUR predictor variable in several studies (Amini et al., 2014; Giorgis-Allemand et al., 2017; Gulliver et al., 2011). However, our modelling found that it was only effective predictor in 3.2% of cases. Industrial emissions typically originate from sources elevated above the ground, such as smokestacks. Therefore air dispersion for these emission sources have a great influence on air quality values measured at ground monitoring stations. More careful consideration of industrial variables (e.g. specific point emissions) may be required to better implement this variable into LUR modelling (Muttoo et al., 2018).

4.4. Strengths and limitations

Our study has a number of strengths, particularly the modelling of air pollution exposures in a period with sparse data on concentrations and predictor variables. We were able to develop LUR models for monthly averages, accounting for temporal variability. LUR models may not be easily transferable to different regions or countries, as the predictors and relationships between land-use and air pollution can vary across regions. In our case, we sought to create a model that would be applicable to the entire republic, but also to neighbouring countries, such as Slovakia, with which the Czech Republic at that time formed one state.

Our study has clear limitations as well. The number of stations and the number of months with valid data varied considerably during this dynamic period. Using stations that had all valid data for all months, would have resulted to an inability to construct valid models. Only five stations had 100% values for SO_2 and none for the other two pollutants (Table S1).

Some of the predictor variables were available only from global databases and did not always refer to the period of interest. For the road data we used paper maps that were adapted to the recent

OpenStreetMap data. We did not have access to traffic volume data which likely contributed to the moderate performance of especially our NO_x and SPM models. Finally, we had no information on emissions of key industrial point sources in the study period. We instead worked with industrial land use from the Corine database.

We developed separate monthly models for the five years of the study period for the sake of simplicity. We did not use the more complex geographically-temporally weighted models (Shen et al., 2022) or other spatio-temporal models. Shen et al. (2022) found little difference in performance between 20 annual-specific and a single spatio-temporal model. A single model is computationally more efficient and benefits from neighbouring months.

5. Conclusions

We were able to construct monthly average LUR models for SO_2, NOx and SPM for the years 1990–1994 in the Czech Republic for to be linked to a birth cohort study. Our models will be, with the same methodology, but time-specific variables, later extended also for other life periods. Our models will serve as the basis for exposure estimates in longitudinal environmental epidemiological studies for cohorts in an area where no similar models have been applied, particularly during the socio-economic transition of the region.

Our results showed significant differences in exposures to air pollutants during different trimesters and early-life vulnerable windows in children, as we were able to estimate monthly and spatially-resolved air concentrations in detail. These findings will be of great use for studying the impacts of air pollution on birth outcomes, respiratory diseases, and neurodevelopmental disorders in the prospective longitudinal cohort study ELSPAC.

Furthermore, the models can be utilized in other Central European cities during the same period.

CRediT authorship contribution statement

Ondřej Mikeš: Conceptualization, Formal analysis, Investigation, Data curation, Writing – original draft. **Ondřej Sáňka:** Data curation, Formal analysis, Writing – review & editing. **Aneta Rafajová:** Data curation, Formal analysis. **Jelle Vlaanderen:** Methodology, Writing – review & editing. **Jie Chen:** Methodology, Model development, Writing – review & editing. **Gerard Hoek:** Supervision, Methodology, Model development, Writing – review & editing. **Jana Klánová:** Funding acquisition, Project administration, Resources, Supervision. **Pavel Čupr:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

Acknowledgement

This project received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 857340, grant agreement No 874627 and grant agreement No 857560. The authors thank Research Infrastructure RECETOX RI (No LM2018121) financed by the Ministry of Education, Youth and Sports, and Operational Programme Research, Development and Innovation – project CETOCOEN EXCELLENCE (No CZ.02.1.01/0.0/0.0/17_043/0009632) for supportive background. This work was carried out in the framework of the European Partnership for the Assessment of Risks from Chemicals (PARC) and has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101057014. This publication reflects only the author's view, and the European Commission is not responsible for any use that may be made of the information it contains.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2023.119688.

References

Ahmed, S.M., Mishra, G.D., Moss, K.M., Yang, I.A., Lycett, K., Knibbs, L.D., 2022. Maternal and childhood ambient air pollution exposure and mental health symptoms and psychomotor development in children: an Australian population-based longitudinal study. Environ. Int. 158, 107003 https://doi.org/10.1016/J.ENVINT.2021.107003.

Amini, H., Taghavi-Shahri, S.M., Henderson, S.B., Naddafi, K., Nabizadeh, R., Yunesian, M., 2014. Land use regression models to estimate the annual and seasonal spatial variability of sulfur dioxide and particulate matter in Tehran, Iran. Sci. Total Environ. 488–489, 343–353. https://doi.org/10.1016/J.SCITOTENV.2014.04.106.

Arcdata, Z.Ú.Č.S.Ú., 2016. Digitální Geografická Databáze 1 : 500 000 ArcČR 500 [WWW Document]. URL. https://www.arcdata.cz/produkty/geograficka-data/arccr-4-0. (Accessed 13 June 2022).

Baïz, N., Slama, R., Béné, M.C., Charles, M.A., Kolopp-Sarda, M.N., Magnan, A., Thiebaugeorges, O., Faure, G., Annesi-Maesano, I., 2011. Maternal exposure to air pollution before and during pregnancy related to changes in newborn's cord blood lymphocyte subpopulations. The EDEN study cohort. BMC Pregnancy Childbirth 11, 87. https://doi.org/10.1186/1471-2393-11-87.

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K.T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., von Klot, S., Nádor, G., Varró, M.J., Dedele, A., Gražulevičiene, R., Mölter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Krämer, U., Kuhlbusch, T., Cirach, M., de Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Strömgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B., de Hoogh, K., 2013. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - the ESCAPE project. Atmos. Environ. 72, 10–23. https://doi.org/10.1016/j.atmosenv.2013.02.037.

Bettiol, A., Gelain, E., Milanesio, E., Asta, F., Rusconi, F., 2021. The first 1000 days of life: traffic-related air pollution and development of wheezing and asthma in childhood. A systematic review of birth cohort studies. Environ. Heal. A Glob. Access Sci. Source 20, 1–10. https://doi.org/10.1186/S12940-021-00728-9/TABLES/4.

Briggs, D.J., De Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S., Smallbone, K., 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. Sci. Total Environ. 253, 151–167. https://doi.org/10.1016/S0048-9697(00)00429-0.

Chen, H., Goldberg, M.S., Crouse, D.L., Burnett, R.T., Jerrett, M., Villeneuve, P.J., Wheeler, A.J., Labrèche, F., Ross, N.A., 2010a. Back-extrapolation of estimates of exposure from current land-use regression models. Atmos. Environ. 44, 4346–4354. https://doi.org/10.1016/j.atmosenv.2010.07.061.

Chen, H., Goldberg, M.S., Crouse, D.L., Burnett, R.T., Jerrett, M., Villeneuve, P.J., Wheeler, A.J., Labrèche, F., Ross, N.A., 2010b. Back-extrapolation of estimates of exposure from current land-use regression models. Atmos. Environ. 44, 4346–4354. https://doi.org/10.1016/J.ATMOSENV.2010.07.061.

Chen, J., de Hoogh, K., Gulliver, J., Hoffmann, B., Hertel, O., Ketzel, M., Bauwelinck, M., van Donkelaar, A., Hvidtfeldt, U.A., Katsouyanni, K., Janssen, N.A.H., Martin, R.V., Samoli, E., Schwartz, P.E., Stafoggia, M., Bellander, T., Strak, M., Wolf, K., Vienneau, D., Vermeulen, R., Brunekreef, B., Hoek, G., 2019. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. Environ. Int. 130, 104934 https://doi.org/10.1016/j.envint.2019.104934.

CHMI, 2021. Znečištění Ovzduší Na Území České Republiky V Roce 2020 (Air Pollution in Czech Republic in 2020) (Praha).

Dadvand, P., Figueras, F., Basagaña, X., Beelen, R., Martinez, D., Cirach, M., Schembari, A., Hoek, G., Brunekreef, B., Nieuwenhuijsen, M.J., 2013. Ambient air pollution and preeclampsia: a spatiotemporal analysis. Environ. Health Perspect. 121, 1365–1371. https://doi.org/10.1289/EHP.1206430.

Dalecká, A., Hermanová, B., Janoš, T., Riedlová, P., Jiřík, V., 2020. Epidemiological evidence of ambient air pollution health effects. Hygiena 65, 106–114. https://doi.org/10.21101/hygiena.a1742.

@ Darling, J.C., Bamidis, P.D., Burberry, J., Rudolf, M.C.J., Darling, J, C., 2020. The First Thousand Days: early, integrated and evidence-based approaches to improving child health: coming to a population near you?. Arch. Dis. Child. 105, 837–841. https://doi.org/10.1136/archdischild-2019-316929.

EEA, 2018. European Union Emission Inventory Report 1990–2016 under the UNECE Convention on Long-Range Transboundary Air Pollution (LRTAP). Rep. No 6/18. Copenhagen.

EEA, 2010. Sulphur Dioxide (SO2) Emissions. European Environment Agency.

EEA (European Environment Agency), 1996. AIR QUALITY IN EUROPE, 1993 A PILOT REPORT (Copenhagen).

Eeftens, M., Beelen, R., De Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dedele, A., Dons, E., De Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Gražulevičiene, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Nádor, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Klot, S., Von Wolf, K., Brunekreef, B., Hoek, G., 2012. Development of land use regression models for PM2.5, PM 2.5 absorbance, PM10 and PMcoarse in 20 European study areas; Results of the ESCAPE project. Environ. Sci. Technol. 46, 11195–11205. https://doi.org/10.1021/es301948k.

Eeftens, M., Beelen, R., Fischer, P., Brunekreef, B., Meliefste, K., Hoek, G., 2011. Stability of measured and modelled spatial contrasts in NO_2 over time. Occup. Environ. Med. 68, 765. https://doi.org/10.1136/oem.2010.061135. LP – 770.

Etzel, R.A., 2020. The special vulnerability of children. Int. J. Hyg Environ. Health 227, 113516. https://doi.org/10.1016/J.IJHEH.2020.113516.

European Environment Agency, 2018a. Unequal Exposure and Unequal Impacts 97.

European Environment Agency, 2018b. European union, copernicus land monitoring service (1990) [WWW Document]. URL. https://land.copernicus.eu/pan-european/corine-land-cover/clc-1990.

Florczyk, A.J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., Sabo, F., Zanchetta, L., 2019. GHSL Data Package. Publications Office of the European Union, Luxembourg, ISBN 978-92-76-13186-1. https://doi.org/10.2760/290498. JRC 117104.

Fukuo, A., Hanaoka, K., Arai, T., Nakai, S., 2018. Development of land use regression models for assessing exposure to NO_2, NO_x and SPM, and their applicability to air pollution epidemiologic studies in Japan. J. Japan Soc. Atmos. Environ./Taiki Kankyo Gakkaishi 53, 153–164. https://doi.org/10.11298/taiki.53.153.

Giorgis-Allemand, L., Pedersen, M., Bernard, C., Aguilera, I., Beelen, R.M.J., Chatzi, L., Cirach, M., Danileviciute, A., Dedele, A., Van Eijsden, M., Estarlich, M., Fernández-Somoano, A., Fernández, M.F., Forastiere, F., Gehring, U., Grazuleviciene, R., Gruzieva, O., Heude, B., Hoek, G., De Hoogh, K., Van Den Hooven, E.H., Håberg, S. E., Iñiguez, C., Jaddoe, V.W.V., Korek, M., Lertxundi, A., Lepeule, J., Nystad, W., Patelarou, E., Porta, D., Postma, D., Raaschou-Nielsen, O., Rudnai, P., Siroux, V., Sunyer, J., Stephanou, E., Sørensen, M., Eriksen, K.T., Tuffnell, D., Varró, M.J., Vrijkotte, T.G.M., Wijga, A., Wright, J., Nieuwenhuijsen, M.J., Pershagen, G., Brunekreef, B., Kogevinas, M., Slama, R., 2017. Original Contribution the Influence of Meteorological Factors and Atmospheric Pollutants on the Risk of Preterm Birth. https://doi.org/10.1093/aje/kww141.

Grulichová, M., Zlámal, F., Andrýsková, L., Švancara, J., Pikhart, H., Bienertová-Vašků, J., 2020. Dietary pattern longitudinally during 8 years in children: results from the European longitudinal study of pregnancy and childhood (ELSPAC-CZ). Publ. Health Nutr. 1–7. https://doi.org/10.1017/S1368980020001056.

Gulliver, J., De Hoogh, K., Hansell, A., Vienneau, D., 2013. Development and back-extrapolation of NO2 land use regression models for historic exposure assessment in Great Britain. Environ. Sci. Technol. 47, 7804–7811. https://doi.org/10.1021/es4008849.

Gulliver, J., Elliott, P., Henderson, J., Hansell, A.L., Vienneau, D., Cai, Y., McCrea, A., Garwood, K., Boyd, A., Neal, L., Agnew, P., Fecht, D., Briggs, D., de Hoogh, K., 2018. Local- and regional-scale air pollution modelling (PM10) and exposure assessment for pregnancy trimesters, infancy, and childhood to age 15 years: avon Longitudinal Study of Parents and Children (ALSPAC). Environ. Int. 113, 10–19. https://doi.org/10.1016/j.envint.2018.01.017.

Gulliver, J., Morris, C., Lee, K., Vienneau, D., Briggs, D., Hansell, A., 2011. Land use regression modeling to estimate historic (1962-1991) concentrations of black smoke and sulfur dioxide for Great Britain. Environ. Sci. Technol. 45, 3526–3532. https://doi.org/10.1021/es103821y.

Hoek, G., Beelen, R., Brunekreef, B., 2015. Land use regression models for outdoor air pollution. In: Nieuwenhuijsen, M.J. (Ed.), Exposure Assesment in Environmental Epidemiology. Oxford University Press, New York, pp. 271–294.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. https://doi.org/10.1016/j.atmosenv.2008.05.057.

Hogg, K., Price, E.M., Hanna, C.W., Robinson, W.P., 2012. Prenatal and perinatal environmental influences on the human fetal and placental epigenome. Clin. Pharmacol. Ther. 92, 716–726. https://doi.org/10.1038/clpt.2012.141.

Huang, L., Zhang, C., Bi, J., 2017. Development of land use regression models for PM2.5, SO2, NO2 and O3 in Nanjing, China. Environ. Res. 158, 542–552. https://doi.org/10.1016/J.ENVRES.2017.07.010.

Kajekar, R., 2007. Environmental factors and developmental outcomes in the lung. Pharmacol. Ther. 114, 129–145. https://doi.org/10.1016/j.pharmthera.2007.01.011.

Kashima, S., Yorifuji, T., Tsuda, T., Doi, H., 2009. Application of land use regression to regulatory air quality data in Japan. Sci. Total Environ. 407, 3055–3062. https://doi.org/10.1016/J.SCITOTENV.2008.12.038.

Klepac, P., Locatelli, I., Korošec, S., Künzli, N., Kukec, A., 2018. Ambient air pollution and pregnancy outcomes: a comprehensive review and identification of environmental public health challenges. Environ. Res. 167, 144–159. https://doi.org/10.1016/j.envres.2018.07.008.

Knibbs, L.D., Coorey, C.P., Bechle, M.J., Marshall, J.D., Hewson, M.G., Jalaludin, B., Morgan, G.G., Barnett, A.G., 2018. Long-term nitrogen dioxide exposure assessment using back-extrapolation of satellite-based land-use regression models for Australia. Environ. Res. 163, 16–25. https://doi.org/10.1016/J.ENVRES.2018.01.046.

Kunc, J., Tonev, P., 2022. Growth and Change in Post-socialist Cities of Central Europe, Routledge. Routledge, Abingdon, Oxon.

Levy, I., Levin, N., Yuval, Y., Schwartz, J.D., Kark, J.D., 2015. Back-extrapolating a land use regression model for estimating past exposures to traffic-related air pollution. Environ. Sci. Technol. 49, 3603–3610. https://doi.org/10.1021/es505707e.

Luo, D., Kuang, T., Chen, Y.X., Huang, Y.H., Zhang, H., Xia, Y.Y., 2021. Air pollution and pregnancy outcomes based on exposure evaluation using a land use regression model: a systematic review. Taiwan. J. Obstet. Gynecol. 60, 193–215. https://doi.org/10.1016/j.tjog.2021.01.004.

Manisalidis, I., Stavropoulos, E., Stavropoulos, A., Bezirtzoglou, E., 2020. Environmental and health impacts of air pollution: a review. Front. Public Health 14. https://doi.org/10.3389/fpubh.2020.00014.

Mikeš, O., Vrbová, M., Klánová, J., Čupr, P., Švancara, J., Pikhart, H., 2019. Early-life exposure to household chemicals and wheezing in children. Sci. Total Environ. 663, 418–425. https://doi.org/10.1016/j.scitotenv.2019.01.254.

Muttoo, S., Ramsay, L., Brunekreef, B., Beelen, R., Meliefste, K., Naidoo, R.N., 2018. Land use regression modelling estimating nitrogen oxides exposure in industrial south Durban, South Africa. Sci. Total Environ. 610–611, 1439–1447. https://doi.org/10.1016/J.SCITOTENV.2017.07.278.

Novotny, E.V., Bechle, M.J., Millet, D.B., Marshall, J.D., 2011. National satellite-based land-use regression: NO 2 in the United States. Environ. Sci. Technol. 45, 4407–4414. https://doi.org/10.1021/es103578x.

Piler, P., Kandrnal, V., Kukla, L., Andrýsková, L., Švancara, J., Jarkovský, J., Dušek, L., Pikhart, H., Bobák, M., Klánová, J., Piler, Pavel, Kandrnal, Vít, Kukla, Lubomír, Andrýsková, Lenka, Jan, Švancara, Jarkovský, Jiří, Dušek, Ladislav, Pikhart, Hynek, Martin, Bobák, Jana, Klánová, 2016. Cohort profile: European longitudinal study of pregnancy and childhood (ELSPAC) in the Czech republic. Int. J. Epidemiol. https://doi.org/10.1093/ije/dyw091.

Poirier, A., Dodds, L., Dummer, T., Rainham, D., Maguire, B., Johnson, M., 2015. Maternal exposure to air pollution and adverse birth outcomes in Halifax, Nova Scotia. J. Occup. Environ. Med. 57, 1291–1298. https://doi.org/10.1097/JOM.0000000000000604.

Sanchez, M., Ambros, A., Milà, C., Salmon, M., Balakrishnan, K., Sambandam, S., Sreekanth, V., Marshall, J.D., Tonne, C., 2018. Development of land-use regression models for fine particles and black carbon in peri-urban South India. Sci. Total Environ. 634, 77–86. https://doi.org/10.1016/j.scitotenv.2018.03.308.

Saucy, A., Röösli, M., Künzli, N., Tsai, M.Y., Sieber, C., Olaniyan, T., Baatjies, R., Jeebhay, M., Davey, M., Flückiger, B., Naidoo, R.N., Dalvie, M.A., Badpa, M., de Hoogh, K., 2018. Land use regression modelling of outdoor NO2 and PM2.5 concentrations in three low income areas in the western cape province, South Africa. Int. J. Environ. Res. Publ. Health 15 (1452 15), 1452. https://doi.org/10.3390/IJERPH15071452, 2018.

Schembari, A., De Hoogh, K., Pedersen, M., Dadvand, P., Martinez, D., Hoek, G., Petherick, E.S., Wright, J., Nieuwenhuijsen, M.J., 2015. Ambient air pollution and newborn size and adiposity at birth: differences by maternal ethnicity (the born in bradford study cohort). Environ. Health Perspect. 123, 1208–1215. https://doi.org/10.1289/EHP.1408675.

Selevan, S.G., Kimmel, C.A., Mendola, P., 2000. Identifying critical windows of exposure for children's health. Environ. Health Perspect. 108, 451. https://doi.org/10.1289/EHP.00108S3451.

Shen, Y., de Hoogh, K., Schmitz, O., Clinton, N., Tuxen-Bettman, K., Brandt, J., Christensen, J.H., Frohn, L.M., Geels, C., Karssenberg, D., Vermeulen, R., Hoek, G., 2022. Europe-wide air pollution modeling from 2000 to 2019 using geographically weighted regression. Environ. Int. 168 https://doi.org/10.1016/J.ENVINT.2022.107485.

Smarr, M.M., Vadillo-Ortega, F., Castillo-Castrejon, M., O'Neill, M.S., 2013. The use of ultrasound measurements in environmental epidemiological studies of air pollution and fetal growth. Curr. Opin. Pediatr. https://doi.org/10.1097/MOP.0b013e32835e1e74.

Stepanikova, I., Thon, V., Mikes, O., Klanova, J., Irena Stepanikova, C., 2021. A model of perinatal stress and childhood wheezing: ELSPAC-CZ cohort. Pediatr. Pulmonol. 56, 1471–1483. https://doi.org/10.1002/ppul.25346.

Su, J.G., Jerrett, M., Beckerman, B., 2009. A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. Sci. Total Environ. 407, 3890–3898. https://doi.org/10.1016/j.scitotenv.2009.01.061.

Su, Jason G., Jerrett, M., Beckerman, B., Wilhelm, M., Ghosh, J.K., Ritz, B., 2009. Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. Environ. Res. 109, 657–670. https://doi.org/10.1016/J.ENVRES.2009.06.001.

Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G., Brunekreef, B., 2012. Systematic Evaluation of Land Use Regression Models for NO 2. https://doi.org/10.1021/es204183v.

Wang, R., Henderson, S.B., Sbihi, H., Allen, R.W., Brauer, M., 2013. Temporal stability of land use regression models for traffic-related air pollution. Atmos. Environ. 64, 312–319. https://doi.org/10.1016/J.ATMOSENV.2012.09.056.

WHO, 2013. Review of Evidence on Health Aspects of Air Pollution. REVIHAAP Project, Bonn.

Wolf, K., Cyrys, J., Harciníková, T., Gu, J., Kusch, T., Hampel, R., Schneider, A., Peters, A., 2017. Land use regression modeling of ultrafine particles, ozone, nitrogen oxides and markers of particulate matter pollution in Augsburg, Germany. Sci. Total Environ. 579, 1531–1540. https://doi.org/10.1016/J.SCITOTENV.2016.11.160.

Wright, R.O., 2017. Environment, susceptibility windows, development, and child health. Curr. Opin. Pediatr. https://doi.org/10.1097/MOP.0000000000000465.

Zhu, Y., Hinds, W.C., Kim, S., Shen, S., Sioutas, C., 2002. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. Atmos. Environ. 36, 4323–4335. https://doi.org/10.1016/S1352-2310(02)00354-0.