Particle concentrations, dispersion modelling and evaluation in southern Sweden

Ralf Rittner¹ · Susanna Gustafsson² · Mårten Spanne² · Ebba Malmqvist¹

Received: 23 December 2019 / Accepted: 15 April 2020 / Published online: 4 May 2020

© The Author(s) 2020

Abstract

Health impact assessments of differential air pollution rely on epidemiologically established relationships between concentration levels where people are exposed and adverse health outcomes. To assess air pollution concentrations, land use regression is commonly used. However, an alternative tool is dispersion modelling, where a detailed inventory of pollution sources together with meteorological data drives calculations of compound dispersion. With this, both spatial and temporal variation can be assessed. In this study, we evaluated results of a Gaussian dispersion model applied to an emissions inventory for Scania, the southernmost county in Sweden. The dispersion considered was particulate matter of aerodynamic diameter $< 10 \mu\text{m}$ (PM_{10}), particulate matter of aerodynamic diameter $< 2.5 \mu\text{m}$ ($\text{PM}_{2.5}$) and black carbon (BC) during an 11-year period (2000–2011). Mean concentrations and 95th percentiles expressed in $\mu\text{g}/\text{m}^3$ ranged from 10.1 to 12.6 and 16.6 to 20.7 for $\text{PM}_{2.5}$ and from 14.0 to 18.8 and 22.6 to 27.0 for PM_{10} , respectively. Seven monitoring stations were used for evaluation. Correlations (R^2) ranged from 0.44 to 0.86 for $\text{PM}_{2.5}$ (mean bias from -9.0 to $0.1 \mu\text{g}/\text{m}^3$) and from 0.46 to 0.83 for PM_{10} (mean bias -6.1 to $3.5 \mu\text{g}/\text{m}^3$). An evaluated database of PM and BC concentrations for Scania is now available for future exposure assessment projects. Calculations were based on a well-known dispersion model with detailed emission data as input. The evaluation showed correlation coefficients for PM in line with previous literature. The data on PM_{10} , $\text{PM}_{2.5}$ and BC concentrations will, therefore, be used in subsequent studies, epidemiological as well as health impact assessments.

Keywords Dispersion modelling · Air pollution · Particulate matter · Emission database · Dispersion model evaluation

1 Introduction

Air pollution is the major environmental health threat globally [1]. Air pollution consists of a mixture of gases and particles suspended in the air. Particles, called particulate matter (PM), are divided into different sizes in aerodynamic diameter: particulate matter $< 10 \mu\text{m}$ (PM_{10}) and particulate matter $< 2.5 \mu\text{m}$ ($\text{PM}_{2.5}$). When humans inhale polluted air, particles can reach different parts of the respiratory system depending on their size. Indeed, these pollutants

account for much of the negative health consequences reported in international health impact assessments, such as the Global Burden of Disease [2]. In addition, black carbon (BC) emitted from incomplete combustion has been suggested to be a better indicator of harmful substances than undifferentiated PM [3]. Adverse health effects could be caused by BC itself or the chemicals it carries to and beyond the lungs [4].

In order to assess the health-related impacts from air pollution, we need to first assess levels of exposure at the

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s42452-020-2769-1>) contains supplementary material, which is available to authorized users.

✉ Ralf Rittner, Ralf.Rittner@med.lu.se | ¹Division of Occupational and Environmental Medicine, Lund University, Medicon Village, Scheelevägen 8, 223 81 Lund, Sweden. ²Environmental Department of the City of Malmö, Malmö, Sweden.

SN Applied Sciences (2020) 2:1013 | <https://doi.org/10.1007/s42452-020-2769-1>

places where people are situated. This can be done with direct or indirect methods. For the former, the individual's exposure is traced through personal portable monitors or assessed through the collection of relevant exposure biomarker samples. In a large cohort, however, this would be both costly and labour-intensive. Due to the limitations of these direct methods, indirect approaches are commonly used. Here, the aspiration is to determine where a person spends time and to combine this with information about pollutants at that particular place, which would require monitoring individuals' geographical movements. This is also not particularly feasible with a larger study population. Therefore, air pollution concentrations at an individuals' residential position have historically been ascribed to concentrations recorded at the nearest air quality monitoring station. An underlying assumption here is that all individuals residing in the area surrounding the monitoring station experience the same pollution concentrations. Consequently, this method is limited because it disregards air pollutants' spatial variation, which is especially important when studying long-term health effects [5]. The extent of this problem can vary depending upon the study design chosen, the pollutants being investigated and the density of existing monitoring stations. With the development of Geographical Information Systems (GIS), more advanced models attempting to correct this shortcoming have been developed in recent decades. For instance, land use regression (LUR) models and dispersion models (DM), or a combination of the two, are popular, commonly used models. Understandably, there are advantages and disadvantages to both, which are further illustrated in a review by Jerrett et al. [5].

Dispersion models are based on detailed knowledge of dynamic atmospheric processes and incorporate information on emissions as well as the characteristics of their sources. In addition, meteorology data, including wind speed and direction, temperature and solar radiation, are used to predict the ground level spatial distribution of pollutants' concentrations [6]. Simplified atmospheric chemistry, diffusion and transport are also integrated into the model. Gaussian dispersion models are based on a Gaussian plume equation to estimate the spatial distribution of pollutant concentrations [5, 7]. The Gaussian dispersion model was originally designed as an air quality management tool but has been widely used for estimating long-term exposures [7]. In order to calibrate a dispersion model, data on pollution concentrations measured at urban background monitoring stations in the studied area are typically used. Further, emissions are often divided into stationary and mobile (line) sources [5]. The dispersion model has the advantage of incorporating both spatial and temporal variation without the need of extensive monitoring [5]. Despite this, it is costly and requires highly

trained personnel in both computer programming and GIS. Results from dispersion modelling, which are simulations of how the studied compound will be dispersed, will be highly depending on the quality of the emission data inserted into the model [8]. In Nordic countries with high-quality data on emissions, dispersion models have been utilized regularly [9–13].

1.1 Objective

The objective was to model and subsequently evaluate the modelled concentrations of PM_{10} , $PM_{2.5}$ and BC in Scania, Sweden, from 2000 to 2011 using a dispersion model and a detailed emissions inventory. The applied spatial resolution was 100 m * 100 m, and the temporal resolution was calendar months. The evaluation performed was based on available measurements from monitoring stations in Scania. The modelled levels of pollutants using the method described in this paper will be linked to individuals' exposures and used in future epidemiological and health impact studies.

2 Method

2.1 Study area

Scania (Skåne) is the southernmost county in Sweden, covering an area of around 11,350 km², or about 2% of the country's total area (see Fig. 1). It is one of the most densely populated regions in the country, with about 1.3 million inhabitants (approximately 13% of the total Swedish population). Within Scania, most people live in the west where levels of pollutants are higher than other areas of Sweden. This greater exposure is due to road transport to and from the European continent as well as a considerable amount of cargo shipping and ferry transport along the coast.

2.2 Study design

A schematic overview of the modelling process is presented in Fig. 2.

2.3 Emission database

The emission inventory (emission database) is maintained by the Environmental Department of the City of Malmö but covers all of Scania. It was first established in 2006 with emission data from 2001 and has since been regularly updated, so that versions reflecting the current situation at different times are available. Primarily designed for GIS applications, the emission database

Fig. 1 Map of study area with inserted picture containing major towns and road network. Image by Emilie Stroh

consists of geo-coded emission sources, such as point (i.e. chimney), line (i.e. road segment), polygon, and grid sources. Polygon and grid sources differ in that emissions are equally dispersed over the area for polygon sources, while emissions can vary in space for a grid source [14]. The number of emission sources (meaning each individual road segment or chimney) in the database amounts to several hundred thousand. Modelling in this project was based on the database versions best

reflecting emissions for the years starting and ending the studied period (2000 and 2011), respectively.

All emission sources can be divided into eight categories [15]:

1. **Road traffic** Road traffic descriptions include data pertaining to the vehicles driving on each road segment, such as the number of passenger cars and different types of heavy vehicles (buses and trucks with or without trailers). The vehicles are also differentiated

Fig. 2 Schematic representation of the modelling process

based on fuel consumption: petrol and diesel. In some cases, compressed natural gas (CNG) vehicles comprise a significant portion of the traffic composition (e.g. in Malmö) and are, thus, also included. Additionally, the road segment contains information on speed limit, traffic flow and classification into one of 36 distinct types of roads in an urban or rural environment. All data used in this classification work originate from the Swedish Road Administration as well as local municipalities.

2. *Shipping* As the ships nearing Scania's coast vary in type and route, temporal and spatial patterns as well as the quantities of air pollutants emitted from shipping are divided into five categories in the database. Examples include ferries that depart from/arrive at harbours in Scania versus cargo ships passing by. These figures were manually collected for each harbour by Gustafsson [15] for the year 2000. Emission data from shipping has been improved by the project Shipair [16], from which data for the year 2011 was gathered and used for this project. Ships were treated as mobile point sources by utilizing actual positioning and identification data from the ships' own Automatic Identification System (AIS). This information was then reduced to a gridded source. Within the greater harbours (Malmö, Helsingborg and Trelleborg), a resolution of 50 m * 50 m was used, while 500 m * 500 m was applied outside harbours. Even leisure boating was accounted for as an overall grid source.
3. *Aviation* We included only aviation emissions below 3000 feet (912 m) in the emission database. This data

originates from yearly environmental reports produced by Scandinavian airports.

4. *Railroads* Emissions from railroads are minor compared to other sources, mainly because nearly all of Sweden's railroads are electric. Thus, emissions included in the database arise from the small proportion of diesel engine trains, mainly used in cargo transport.
5. *Industries and major energy and heat producers* Air pollution data were collected from the national database EMIR (EMissionsRegister), which is managed by Sweden's county level authorities. It contains information regarding chemical composition of discharge, temperature of exhaust gas, height of chimney, etc., for all of the larger industrial facilities in the area.
6. *Small-scale heating* Information regarding the area's 150,000 in-house heating installations has been collected: 70,000 are oil-burning heaters, 13,000 are wood- or pellet-fueled heaters and 67,000 are small fireplaces. The frequency and use of the stoves can be roughly estimated based on how often the chimney is swept, which was derived from the National Rescue Agency's chimney register (2001). Coordinates for each chimney were not available for the year 2000. For this reason, they have been distributed in residential areas within municipalities according to a model further described Gustafsson [15]. Concerning 2011, the chimney sweep registers from 2013 were considered to be the most appropriate information source, containing data from 32 of the 33 municipalities. Data from the small municipality of Båstad was not digitalized and was, thus, estimated using a more

simplistic method. For this year (2011), information was improved with respect to location, as residential geographical coordinates were linked to each chimney [17]. We identified 206,321 heating objects, out of which 45% were stoves, 25% tiled stoves and open stoves, and 24% boilers. Nearly, all installations were heated by wood except for boilers, which were heated by either gas (28%), oil (25%), wood pellets (23%) or wood (15% in a non-environmental certified boiler and 8% in an environmental certified boiler). Percentages are based number of installations and thus not reflecting amount of fuel consumed [17].

7. *Non-road vehicles* These vehicles, such as cargo trucks, construction and farming machinery, and others, are divided into 11 groups based on the industry to which they belong: agriculture, forestry, harbour, building, domestic, mining, the iron industry, the forest industry, railroads, aviation and military defence. Only those considered resulting in emissions in Scania were included; thus, mining, the iron industry and military defence were excluded. Most information originates from a report published by the 'Swedish Environmental Research Institute' (IVL) [18].
8. *Emissions from Zealand, Denmark* Since emissions throughout Zealand, Denmark, are quite high and westerly winds are dominant, air pollution produced here must be included. These emissions have been shown to reach the western part of Scania, which is where dispersion modelling was performed. Emission data for Denmark was supplied by the Danish National Environmental Research Institute (DMU).

2.3.1 Emission factors for PM₁₀, PM_{2.5} and black carbon

The database was built for emissions of nitrogen oxides (NO_x), sulphur dioxide (SO₂), carbon monoxide (CO), PM₁₀ and volatile organic compounds (VOC). A number of previous studies on air pollution in Scania have been successfully based on the NO_x emissions from this database [10, 13, 15, 19–21]. Emission factors from road traffic exhaust and the suspension of road dust particles were built on the emission model described in Handbook Emission Factors for Road Transport HBEFA 3.2 [22]. The suspension of road dust particles during winter months was also based on the proportion of studded tires (35% of cars in 2000 and 25% in 2011). For this project, emission factors for PM_{2.5} and BC were added. Concerning primary exhaust particles from vehicle internal combustion engines, we can assume that most PM₁₀ particles are actually less than 2.5 µm in diameter: PM_{2.5} equals PM₁₀. For resuspension of particles from road dust, we estimated that only 20% of PM₁₀ particles

are less than 2.5 µm in diameter: PM_{2.5} equals 0.2 of PM₁₀. Traffic emission factors developed by the project *Transphorm* were utilized for BC.

2.4 Dispersion model

The dispersion model used in this project is part of the software suite ENVIMAN, which has been developed by the company Opsis ab and was released in 2006. Within ENVIMAN, the dispersion model AERMOD was utilized [23]. AERMOD is a Gaussian dispersion model provided by the United States Environmental Protection Agency. OPSIS has modified this model to fit the structures of emission and meteorology databases. The Gaussian model is a flat, two-dimensional model. This means that topography and buildings are not taken into consideration, yet the height of both the emission sources and the modelling target are incorporated. Also available in ENVIMAN is a street-canyon model called Operational Street Pollution Model (OSPM) developed by DMU [24]. This model was not used in our project, however, because collecting information on heights of all buildings necessary for using this model (OSPM), was not considered feasible. Further, the cities within the study area contain few narrow streets; thus, a street canyon model would not likely yield any considerable additional contribution.

2.4.1 Meteorology

The ENVIMAN software can calculate the dispersion of emissions over an area with hourly resolution using actual meteorological data for the time period of interest but only monitored at one geographical point. We used the Heleneholm meteorology monitoring station in Malmö, which was the station considered to be of the best quality in our study setting. See Fig. 3.

The following meteorological parameters are used by the model:

1. Temperature at a height of 2 m above ground.
2. Wind speed.
3. Wind direction at 10 and 24 m above ground.
4. Global solar radiation

2.4.2 Resolution

Modelling in ENVIMAN is performed using a grid with adjustable resolution. For instance, the software allows one to not only account for emission sources inside the resulting grid but also in the surrounding area, using two levels of coarser surrounding grids. In this study, we used a spatial resolution of 100 m * 100 m grid cells with an hourly resolution for each as the primary output. These results

Fig. 3 Heleneholm meteorological monitoring station in Malmö

were then aggregated into monthly means based on calendar months. Due to the technical properties (maximum file size) of the software, each simulation run had to be reduced in time (half year) and spatial extent (250×250 squares of 100×100 m). As a result, 60 simulations had to be performed per pollutant and year modelled.

2.5 Background levels

Background level concentrations are those not emitted by sources identified in the emission database, have been transported long-range or are of biogenic origin. These were calculated as the difference between observed concentrations at a background monitoring station and modelled concentrations at that same point. In order to follow established air quality guidelines, authorities monitor air pollution levels throughout the country. In Sweden, these data are held at the Swedish Meteorological and Hydrological Institute (SMHI) where it can be downloaded for free. Some additional data are stored at the municipal level. We retrieved this data in order to compute background levels using the most complete dataset from available monitoring stations, as decided prior to the study beginning. An urban background monitoring station in Malmö city was

chosen, and, for reference, we also computed background pollution levels using a rural background monitoring station with slightly fewer available data points (see Supplementary material).

2.6 Interpolation between years

As modelling each 100-m grid-square over the whole modelling area and for the entire study period is both time- and resource-intensive, the dispersion calculation was only performed for the start and end years of the study period (2000 and 2011). Data for the years falling between these two were interpolated using a linear model. To compensate for seasonal variation of dispersion due to meteorological conditions, an atmospheric ventilation index was calculated and applied to the interpolated concentration data. We used the following method: first, a fixed point that (1) was an urban background station close to the meteorological station (Heleneholm in Malmö) and (2) had a long time series of measurements was chosen. Fitting these parameters, the urban background monitoring station on the roof of Malmö's city hall was selected. We then modelled concentrations for every hour at this point for the complete 11-year study period (105,192 h), using the exact same emission database throughout, and further aggregated the data into monthly means. Using the output of this modelling, the atmospheric ventilation index was calculated, which quantifies the meteorological variation for each month in the modelled concentrations from 2000 to 2011. This index could be applied to all grid cells by multiplying each with the meteorological factor given by the index. When employing the atmospheric ventilation index, we could also incorporate all changes occurring in the emission database versions reflecting the 2 years for which the whole county was modelled originally. This method was developed by the Swedish Clean Air and Climate research programme (SCAC) to obtain yearly means and was further developed by our research team in order to capture monthly variations.

2.7 Evaluation

Evaluation of the model was completed by comparing modelled concentrations to measurement data (aggregated to monthly averages) at the same point and time period. Importantly, measurement data used for evaluation purposes were derived from other available monitoring stations not previously used for the calibration or calculation of background levels. The location of those monitoring stations is shown in Fig. 4. The evaluation included an assessment of mean bias (MB), calculated as the mean difference between modelled and measured concentrations; Pearson correlations (r_p and R^2); and root

Fig. 4 Locations of measurement monitors used

mean square errors (RMSE). The amount of obtainable monitor data for the 2000–2011 period varied by pollutant. PM_{10} had the greatest amount of measured concentration data available, while fewer measurements existed for $PM_{2.5}$ and even less for BC. We used monthly averages (aggregated from hourly or daily means) from seven stations with available PM_{10} data and three stations with accessible $PM_{2.5}$ measurements. For monitor measurements to be accepted, monthly averages were required to be based on non-missing data from at least 8 days in that month. Many of these locations lacked continuous measurements, and concentrations, especially $PM_{2.5}$ concentrations, were frequently measured as part of organized campaigns that often occurred in the wintertime. Despite the risk of an overrepresentation of winter data, all available measurements were used in such cases (three stations). We could unfortunately not find any existing BC measurements for the study period. Calculations were performed using MS Excel® in R 3.4.3 [25].

3 Results

3.1 Concentration of and distribution of pollutants

Figures 5 and 6 illustrate the modelled concentrations of locally emitted pollutants by monitoring station location. The concentrations in $\mu\text{g}/\text{m}^3$ were 0.3–15.3 for PM_{10} , 0.1–9.4 for $PM_{2.5}$ and 0.05–1.48 for BC, with higher levels in Malmö and Trelleborg and the lowest levels in the more north-eastern towns of Hässleholm and Kristianstad. See Table 1.

In addition to the modelled concentrations from locally emitted sources, we added background levels, or long-range transported pollutants, by using measured data with the locally modelled concentrations

subtracted. Locally modelled and measured concentrations of PM_{10} and $PM_{2.5}$ both were from the Malmö City Hall urban background station; no such measured data was available for BC. This resulted in total modelled concentrations that were 10.9–34.6 $\mu\text{g}/\text{m}^3$ for PM_{10} and 5.6–34.4 $\mu\text{g}/\text{m}^3$ for $PM_{2.5}$ (see Table 1). The resulting concentrations were imported to GIS software and visualized as raster maps over Scania (see Figs. 7, 8, 9, 10). For illustration purposes, the second and second to last years (2002 and 2009) of the interpolated data were chosen to supplement the information in diagrams at the monitoring stations shown in Figs. 5 and 6.

Measured and modelled total concentrations can be found in Table 2, with the exception of modelled concentrations for months with missing measured data. More data were available for PM_{10} than $PM_{2.5}$, while BC concentrations were missing. With this, air pollution concentrations ranged between 6.2–75.5 (measured) and 6.3–33.6 (modelled) $\mu\text{g}/\text{m}^3$ for PM_{10} . Regarding $PM_{2.5}$, the ranges found were 5.2–34.8 (measured) and 7.7–34.3 (modelled) $\mu\text{g}/\text{m}^3$.

As a sensitivity analysis when determining how to assess background levels, we used measured concentrations from a regional background site, Vavihill, to calculate total levels in Trelleborg and Kristianstad (see Supplementary materials, Tables S1–S3). This measurement station had more missing values than the data collected from Malmö City Hall. Using the Vavihill data did not improve R^2 correlations for PM_{10} (Supplementary materials, Table S3); however, stronger correlations were seen for $PM_{2.5}$, but this was based on few (10) measurements.

3.2 Model performance (evaluation)

Table 3 displays the statistical comparisons between modelled and measured air pollution levels shown in Table 2. The modelled concentrations demonstrated good overall agreement with measured levels based on the R^2 Pearson correlation, especially in the south-western part of Scania. These R^2 values ranged from 0.46 to 0.83 for PM_{10} and from 0.44 to 0.86 for $PM_{2.5}$. Mean bias was both negative and positive spanning from –6.13 to 3.49 $\mu\text{g}/\text{m}^3$ for PM_{10} and –8.99 to 4.59 $\mu\text{g}/\text{m}^3$ for $PM_{2.5}$. MB was larger in absolute value for monitoring stations with few observations but remained below the fifth percentile of concentration levels. Again, this was not possible to perform for BC due to lack of measurement data during the study period.

Fig. 5 Modelled local concentrations at 5 locations within Malmö for **a** PM_{10} , **b** $PM_{2.5}$ and **c** BC (black carbon)

4 Discussion

Particles of different sizes and from different sources were modelled in the Scania region in Sweden. Modelled

concentrations showed good correlations with measured concentrations, as reported R^2 values ranged from 44 to 86, with higher correlations seen in the south-west part of Scania. The mean bias (MB) calculated did not indicate

Fig. 6 Modelled local concentrations at locations outside Malmö for **a** PM_{10} , **b** $PM_{2.5}$ and **c** BC (black carbon)

Table 1 Summary of modelled concentrations of particulate matter (PM) and black carbon (BC) in monthly averages in $\mu\text{g}/\text{m}^3$ over the study period, total levels with included background and levels without background in brackets

Station	Malmö Dalaplan	Trelleborg	Kristianstad	Landskrona	Lund	Hässleholm	Burlöv
PM_{2.5}							
N observations	123 (132)	123 (132)	123 (132)	123 (132)	123 (132)	123 (132)	123 (132)
Mean	12.9 (2.8)	13.5 (3.4)	11.0 (0.9)	11.3 (1.1)	11.8 (1.7)	11.0 (0.9)	11.9 (1.7)
Minimum	7.7 (0.4)	8.0 (0.4)	5.6 (0.1)	5.8 (0.2)	6.3 (0.4)	5.6 (0.1)	6.3 (0.2)
5th percentile	8.6 (2.3)	8.9 (1.4)	7.3 (0.3)	7.5 (0.5)	8.0 (0.8)	7.3 (0.3)	7.9 (0.8)
Median	12.1 (2.7)	12.6 (3.3)	10.3 (0.9)	10.6 (1.1)	11.2 (1.6)	10.3 (0.8)	11.2 (1.7)
95th percentile	20.0 (4.7)	20.7 (5.7)	17.7 (1.7)	17.8 (1.9)	18.4 (2.9)	17.7 (1.6)	18.7 (2.9)
Maximum	34.3 (6.7)	34.4 (9.4)	32.2 (2.7)	32.7 (3.3)	33.1 (4.4)	32.2 (2.7)	33.2 (4.8)
PM₁₀							
N observations	132 (132)	132 (132)	132 (132)	123 (132)	132 (132)	132 (132)	132 (132)
Mean	19.1 (6.2)	17.3 (4.4)	14.7 (1.9)	14.8 (1.8)	16.0 (3.1)	14.8 (1.9)	16.5 (3.6)
Minimum	10.9 (1.1)	8.4 (0.6)	6.3 (0.3)	6.5 (0.4)	7.5 (0.6)	6.3 (0.3)	8.1 (0.5)
5th percentile	12.9 (2.1)	11.8 (1.7)	9.5 (0.6)	9.3 (0.8)	10.6 (1.1)	9.5 (0.5)	11.0 (1.1)
Median	18.3 (5.7)	16.5 (4.1)	14.0 (1.8)	14.0 (1.8)	15.3 (2.9)	14.0 (1.8)	15.8 (3.2)
95th percentile	27.0 (11.9)	24.7 (7.3)	22.4 (3.5)	22.6 (3.2)	23.6 (5.8)	22.5 (3.4)	23.8 (6.7)
Maximum	39.2 (15.3)	35.6 (2.1)	31.8 (5.5)	32.0 (4.8)	34.3 (8.6)	31.8 (5.5)	34.6 (11.0)
BC							
N observations	144	144	144	144	144	144	144
Mean	(0.64)	(0.89)	(0.22)	(0.30)	(0.42)	(0.35)	(0.43)
Minimum	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
5th percentile	(0.42)	(0.37)	(0.13)	(0.17)	(0.25)	(0.18)	(0.26)
Median	(0.64)	(0.89)	(0.22)	(0.29)	(0.41)	(0.34)	(0.43)
95th percentile	(0.90)	(1.31)	(0.32)	(0.51)	(0.61)	(0.51)	(0.62)
Max	(1.11)	(1.48)	(0.39)	(0.59)	(0.73)	(0.70)	(0.77)

the presence of systematic modelling error in any specific direction (i.e. over- or underestimation).

The modelled levels can and will be used in further epidemiological studies and for conducting health impact assessments of air pollution. Previously, such studies performed in this study area have only had modelled levels of nitrogen oxides (NO_x) to rely on when estimating health impacts. The availability of modelled levels of PM_{10} , $\text{PM}_{2.5}$ and BC will allow this research group to conduct further research on health impacts using well-established exposure–response functions presented by WHO [26].

Our study is in agreement with other studies in Denmark and Sweden that report similar concentrations and correlations [12, 27]. To put our results into an international perspective, Gulliver et al. [28] considered an R^2 of 0.47 to be an acceptable level when they modelled PM_{10} using LUR. Ryan and LeMasters [29] found R^2 to range between 0.54 and 0.81 in their review of 12 LUR models for primary pollutants. In another review, Hoek et al. [30] documented most R^2 to be around 0.6–0.7 for nitrogen dioxide (NO_2). However, it should be mentioned that the LUR validation method, leave one out cross-validation (LOOCV), has been criticized for overestimating the predictive ability of

LUR models at independent sites [31]. Dispersion modelling validations, on the other hand, are performed using independent sites. Nonetheless, comparisons have been made between LUR and dispersion models in Europe, and results are largely affected by the pollutants investigated, type of dispersion model selected and resolution used [7]. In general, Gaussian dispersion models with high spatial resolution predict air pollution concentrations better than Eulerian/CFD (computational fluid dynamics) with coarser spatial resolution [7]. The comparison might be subject to the above-mentioned issue of overestimating the predictability in LUR [31]. Furthermore, both LUR and dispersion models seem to better predict NO_2 than PM_{10} and $\text{PM}_{2.5}$ [7]. One possible explanation is that a higher proportion of regional background sources and a smaller contribution of local sources comprise fine particle concentrations [32]. This is also true for our model, especially in areas further away from the monitoring station used for estimating background levels. To address this uncertainty, we also used a regional background site, Vavihill, which was more centrally located within the study area. There were, however, no large differences in levels between Vavihill and Malmö City Hall, and correlations with measured levels

Fig. 7 Concentrations of PM_{10} , $PM_{2.5}$ and BC in Scania during January 2002 with background levels added where possible

were not clearly improved. Our model has previously been used for modelling NO_x and NO_2 throughout Scania with well-accepted results that have gained great notice [10, 19–21]. There are other more sophisticated models available for the calculation of regional contributions, one being WRF-CHEM (Weather Research and Forecasting (WRF)

Fig. 8 Concentrations of PM_{10} , $PM_{2.5}$ and BC in Scania during July 2002 with background levels added where possible

model coupled with Chemistry) developed by the National Center for Atmospheric Research; Unfortunately, we could not use this model within the scope of this project.

An obvious limitation is the lack of measurement data on BC for the study period. Because of this, it was only

Fig. 9 Concentrations of PM_{10} , $PM_{2.5}$ and BC in Scania during January 2009 with background levels added where possible

possible to model concentrations of local emissions and thus no background levels describing longer distance transportation of BC and, consequently, no total BC levels could be estimated either. Nevertheless, the model has previously been used successfully for NO_x and has proven

Fig. 10 Concentrations of PM_{10} , $PM_{2.5}$ and BC in Scania with during July 2009 background levels added where possible

to be reliable for PM in this study, which is why we could confidently include the results for local BC as well.

Another notable limitation is that meteorological conditions were only recorded at a single geographical point and subsequently used for the rather large area that our modelling covers. This likely contributes to the stronger correlations seen in the areas closer to the chosen

Table 2 Summary of measured (a) and modelled (b) total concentrations of particulate matter (PM) as monthly averages in $\mu\text{g}/\text{m}^3$

Station	Malmö Dalaplan		Trelleborg		Kristianstad		Landskrona		Lund		Hässleholm		Burlöv	
	a	b	a	b	a	b	a	b	a	b	a	b	a	b
PM _{2.5}														
N observations	68		10		15		7							
Mean	13.3	13.3	7.9	12.5	20.7	11.9	10.2	12.2						
Minimum	7.3	7.7	5.2	8.9	8.4	8.1	5.8	7.5						
5th percentile	9.0	8.5	5.3	9.3	11.5	8.3	6.5	8.3						
Median	12.2	12.7	6.7	12.5	21.1	10.8	10.5	12.3						
95th percentile	20.0	19.8	13.0	16.6	30.9	8.3	13.1	16.3						
Max	33.4	34.3	13.8	18.1	34.8	17.6	13.2	17.8						
PM ₁₀														
N observations	72		84		79		64		76		12		58	
Mean	21.5	18.7	19.5	16.6	26.2	13.9	20.7	14.3	15.2	16.1	16.8	12.4	17.9	15.5
Minimum	12.0	10.9	10.8	8.4	11.6	6.3	10.7	7.7	6.2	7.5	12.3	10.3	10.1	8.1
5th percentile	16.1	11.6	12.7	11.2	13.0	9.1	14.3	9.3	10.1	11.0	13.0	10.3	12.0	10.0
Median	19.6	12.2	18.8	15.4	21.9	13.0	19.5	13.4	13.8	15.5	16.1	11.6	17.0	14.5
95th percentile	31.1	26.8	29.0	24.6	47.5	20.9	30.0	23.0	25.9	23.5	21.1	16.3	25.0	22.4
Maximum	43.6	36.7	40.0	35.6	75.5	31.8	42.3	31.5	38.0	33.1	22.6	17.0	34.8	33.6

Table 3 Statistical comparison between measured and modelled concentrations of particulate matter (PM) as monthly averages

Station	Malmö Dalaplan	Trelleborg	Kristianstad	Landskrona	Lund	Hässleholm	Burlöv
PM_{2.5}							
N observations	68	10		7			
Time	(2006–2011)	(2009)		(2009–2010)			
MB ^a	0.10	4.59		−8.99			
r_p^b	0.93	0.80		0.65			
R^2^c	0.86	0.64		0.44			
RMSE ^d	1.50	4.90		9.33			
PM₁₀							
N observations	72	114	115	64	73	12	60
Time	(2006–2011)	(2001–2011)	(2001–2011)	(Winters 2001–2011)	(2002–2011)	(Winter 2004, 2005, 2006)	(2007–2011)
MB ^a	−2.83	3.49	−3.80	−6.13	1.06	−4.36	−2.42
r_p^b	0.91	0.86	0.68	0.79	0.74	0.77	0.84
R^2^c	0.83	0.74	0.46	0.63	0.55	0.58	0.70
RMSE ^d	3.59	10.78	15.66	7.22	3.53	4.72	3.58

^aMean bias^bPearson correlation coefficient^cSquare Pearson correlation coefficient^dRoot-mean-squared error

meteorological station, namely the south-western region of Scania. The incorporation of more advanced meteorological models, such as the WRF Model mentioned earlier, would have been ideal. However, such an addition was, unfortunately, out of scope for this project based on available external software. Precipitation is another meteorological factor that can play a large role in the

concentrations of PM in the air [33]. Precipitation was not taken into account during the acquisition and implementation of modelling software because programs/software available for this project lacked such capabilities.

Also potentially limiting is relying on the adjusted interpolation between years 2000 and 2011 as opposed to running the dispersion for each month of the studied period.

While this could have influenced the results, no strong evidence against this method was found when evaluating the modelled results against the measured values from monitoring stations. However, the integration of more advanced meteorological models into the interpolation process could also have been beneficial. Meteorological concerns such as mentioned above also have implications on the construction of the ventilation index previously described. Relying on a singular monitor for the entire study area may not accurately have captured the nuances of weather events across Scania. Again, the software utilized (ENVIMAN) only allows the input of one set of meteorological data at a time, which is what motivated this approach. On the other hand, we were also able to incorporate in the interpolation process variations due to changes of sources and their characteristics in the emission database versions for the start and endpoints of the studied period as well as variations due to meteorology.

We used a well-known Gaussian dispersion model and a detailed emission database, which can be considered the main strengths of this work. As stated in the methods section, the implementation of the Gaussian model does not account for topography as would other models, such as the CALPUFF in combination with digital elevation maps. Fortunately, Scania could be considered rather flat without any larger mountains and only moderately elevated ridges, making the Gaussian dispersion model highly suitable for our study setting. Furthermore, the emission database, which has been and will be continuously updated over the years, has now reached a considerable level of detail, which increases the reliability of results. This was especially evident for the south-west part of Scania, but also held true for the other monitoring stations as seen by the fair correlations demonstrated there. We, therefore, believe that more detailed emission source data, e.g. traffic intensity for road segments and emission factors for vehicles, further improves the accurate assessment of locally sourced emissions [34]. Such inclusions, as we have used here, should complement and can even exceed emission data available in programmes such as European Monitoring and Evaluation Programme (EMEP).

In conclusion, there is now a database available covering the Scania County with monthly concentrations of $PM_{2.5}$, PM_{10} and BC with a spatial resolution of 100-m grid squares covering an 11-year period. The evaluation conducted provides confidence in its use for exposure estimation and identifies where performance is best: the south-west, which is where the most detailed description of emissions is available. The data and results presented here can and will be the basis for future epidemiological studies as well as health impact assessments.

Acknowledgements Open access funding provided by Lund University. This work was funded by Swedish Environmental Protection

Agency (Naturårdsverket) (2251-15-007). We would like to thank Erin Flanagan for help with editing.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by SG, RR and MS. The first draft of the manuscript was written by RR and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Heroux ME, Braubach M, Korol N, Krzyzanowski M, Paunovic E, Zastenskaya I (2013) The main conclusions about the medical aspects of air pollution: the projects REVIHAAP and HRAPIE WHO/EC. *Gig Sanit* 6:9–14
2. Collaborators GBDMM (2016) Global, regional, and national levels of maternal mortality, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 388(10053):1775–1812. [https://doi.org/10.1016/S0140-6736\(16\)31470-2](https://doi.org/10.1016/S0140-6736(16)31470-2)
3. Janssen NA, Hoek G, Simic-Lawson M, Fischer P, van Bree L, ten Brink H, Keuken M, Atkinson RW, Anderson HR, Brunekreef B, Cassee FR (2011) Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM_{10} and $PM_{2.5}$. *Environ Health Perspect* 119(12):1691–1699. <https://doi.org/10.1289/ehp.1003369>
4. WHO (2012) Health effects from black carbon. World Health Organisation, European Chapter, Geneva
5. Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahuvaroglu T, Morrison J, Giovis C (2005) A review and evaluation of intraurban air pollution exposure models. *J Expo Anal Environ Epidemiol* 15(2):185–204. <https://doi.org/10.1038/sj.jea.7500388>
6. Holmes NS, Morawska L (2006) A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available. *Atmos Environ* 40(30):5902–5928. <https://doi.org/10.1016/j.atmosenv.2006.06.003>
7. de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, Badaloni C, Beelen R, Bolignano A, Cesaroni G, Pradas MC, Cyrus J, Douros J, Eeftens M, Forastiere F, Forsberg B, Fuks K, Gehring U, Gryparis A, Gulliver J, Hansell AL, Hoffmann B, Johansson C, Jonkers S, Kangas L, Katsouyanni K, Kunzli N, Lank T, Memmesheimer M, Moussiopoulos N, Modig L,

- Pershagen G, Probst-Hensch N, Schindler C, Schikowski T, Sugiri D, Teixeira O, Tsai MY, Yli-Tuomi T, Brunekreef B, Hoek G, Bellander T (2014) Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environ Int* 73:382–392. <https://doi.org/10.1016/j.envint.2014.08.011>
8. Berkowicz R, Winther M, Ketzel M (2006) Traffic pollution modelling and emission data. *Environ Modell Softw* 21(4):454–460. <https://doi.org/10.1016/j.envsoft.2004.06.013>
 9. Bellander T, Berglund N, Gustavsson P, Jonson T, Nyberg F, Pershagen G, Jarup L (2001) Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environ Health Perspect* 109(6):633–639
 10. Stroh E, Harrie L, Gustafsson S (2007) A study of spatial resolution in pollution exposure modelling. *Int J Health Geograph* 6:19. <https://doi.org/10.1186/1476-072X-6-19>
 11. Ketzel M, Berkowicz R, Hvidberg M, Jensen SS, Raaschou-Nielsen O (2011) Evaluation of AirGIS: a GIS-based air pollution and human exposure modelling system. *Int J Environ Pollut* 47(1–4):226–238
 12. Hvidtfeldt UA, Ketzel M, Sørensen M, Hertel O, Khan J, Brandt J, Raaschou-Nielsen O (2018) Evaluation of the Danish AirGIS air pollution modeling system against measured concentrations of PM_{2.5}, PM₁₀, and black carbon. *Environ Epidemiol* 2(2):e014. <https://doi.org/10.1097/ee9.0000000000000014>
 13. Stroh E, Rittner R, Oudin A, Ardo J, Jakobsson K, Bjork J, Tinnerberg H (2012) Measured and modeled personal and environmental NO₂ exposure. *Popul Health Metr* 10:10. <https://doi.org/10.1186/1478-7954-10-10>
 14. Stroh E (2006) The use of GIS in exposure-response studies. Lund University, Lund
 15. Gustafsson S (2007) Uppbyggnad och validering av emissionsdatabas avseende luftföroreningar i Skåne med basår 2001. Lund University, Lund
 16. SMHI (2016) Shipair. <https://www.smhi.se/airviro/modules/shipping/shipair-1.22691>. Accessed 13 Feb 2018
 17. Johansson J (2016) Småskalig uppvärmning- utsläpp och haltberäkningar för Skånes kommuner. Environmental Department, Malmö City, Malmö
 18. Persson KK (1999) Kartläggning av emissioner från arbetsfordon och arbetsreskap i kommuner. Gothenburg
 19. Malmqvist E, Jakobsson K, Tinnerberg H, Rignell-Hydbom A, Rylander L (2013) Gestational diabetes and preeclampsia in association with air pollution at levels below current air quality guidelines. *Environ Health Perspect* 121(4):488–493. <https://doi.org/10.1289/ehp.1205736>
 20. Malmqvist E, Larsson HE, Jonsson I, Rignell-Hydbom A, Ivarsson SA, Tinnerberg H, Stroh E, Rittner R, Jakobsson K, Swietlicki E, Rylander L (2015) Maternal exposure to air pollution and type 1 diabetes—accounting for genetic factors. *Environ Res* 140:268–274. <https://doi.org/10.1016/j.envres.2015.03.024>
 21. Stroh E (2011) The use of GIS in assessing exposure to airborne pollutants. Lund University, Lund
 22. Keller M, Wüthrich P (2014) Handbook emission factors for road transport 3.1/3.2. Bern, Switzerland
 23. Environmental Protection Agency US (2004) AERMOD: description of model formulation. U.S. Environmental Protection Agency
 24. Berkowicz R (2000) OSPM—a parameterised street pollution model. *Environ Monit Assess* 65(1–2):323–331. <https://doi.org/10.1023/a:1006448321977>
 25. Team RC (2017) R: a language and environment for statistical computing R foundation for statistical computing
 26. Heroux ME, Anderson HR, Atkinson R, Brunekreef B, Cohen A, Forastiere F, Hurley F, Katsouyanni K, Krewski D, Krzyzanowski M, Kunzli N, Mills I, Querol X, Ostro B, Walton H (2015) Quantifying the health impacts of ambient air pollutants: recommendations of a WHO/Europe project. *Int J Public Health* 60(5):619–627. <https://doi.org/10.1007/s00038-015-0690-y>
 27. Gidhagen L, Omstedt G, Pershagen G, Willers S, Bellander T (2013) High-resolution modeling of residential outdoor particulate levels in Sweden. *J Exposure Sci Environ Epidemiol* 23(3):306–314. <https://doi.org/10.1038/jes.2012.122>
 28. Gulliver J, de Hoogh K, Fecht D, Vienneau D, Briggs D (2011) Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. *Atmos Environ* 45(39):7072–7080. <https://doi.org/10.1016/j.atmosenv.2011.09.042>
 29. Ryan PH, LeMasters GK (2007) A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhal Toxicol* 19(Suppl 1):127–133. <https://doi.org/10.1080/08958370701495998>
 30. Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, Briggs D (2008) A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos Environ* 42(33):7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>
 31. Wang M, Beelen R, Basagana X, Becker T, Cesaroni G, de Hoogh K, Dedele A, Declercq C, Dimakopoulou K, Eeftens M, Forastiere F, Galassi C, Grazuleviciene R, Hoffmann B, Heinrich J, Iakovides M, Kunzli N, Korek M, Lindley S, Molter A, Mosler G, Madsen C, Nieuwenhuijsen M, Phuleria H, Pedeli X, Raaschou-Nielsen O, Ranzi A, Stehanou E, Sugiri D, Stempfelet M, Tsai MY, Lanki T, Udvardy O, Varro MJ, Wolf K, Weinmayr G, Yli-Tuomi T, Hoek G, Brunekreef B (2013) Evaluation of land use regression models for NO₂ and particulate matter in 20 European study areas: the ESCAPE project. *Environ Sci Technol* 47(9):4357–4364. <https://doi.org/10.1021/es305129t>
 32. Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, Declercq C, Dedele A, Dons E, de Nazelle A, Dimakopoulou K, Eriksen K, Falq G, Fischer P, Galassi C, Grazuleviciene R, Heinrich J, Hoffmann B, Jerrett M, Keidel D, Korek M, Lanki T, Lindley S, Madsen C, Molter A, Nador 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 MY, Yli-Tuomi T, Varro MJ, Vienneau D, von Klot S, Wolf K, Brunekreef B, Hoek G (2012) Development of Land use regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM_{coarse} in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* 46(20):11195–11205. <https://doi.org/10.1021/es301948k>
 33. Sparmacher H, Fulber K, Bonka H (1993) Below-cloud scavenging of aerosol-particles—particle-bound radionuclides experimental. *Atmos Environ A Gen* 27(4):605–618. [https://doi.org/10.1016/0960-1686\(93\)90218-N](https://doi.org/10.1016/0960-1686(93)90218-N)
 34. Hertel O, Hvidberg M, Ketzel M, Storm L, Stausgaard L (2008) A proper choice of route significantly reduces air pollution exposure—a study on bicycle and bus trips in urban streets. *Sci Total Environ* 389(1):58–70. <https://doi.org/10.1016/j.scitotenv.2007.08.058>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.