



Activity-based air pollution exposure assessment: Differences between homemakers and cycling commuters

Meng Lu^{a,*}, Oliver Schmitz^a, Ilonca Vaartjes^b, Derek Karssenbergh^b

^a Department of Physical Geography, Faculty of Geoscience, Utrecht University, Princetonlaan 8a, 3584, CB, Utrecht, the Netherlands

^b Department of Epidemiology, UMC Utrecht, Div. Julius Centrum, UMC, Huispost Str., 6.131, 3508, GA, Utrecht, the Netherlands

ARTICLE INFO

Keywords:

Exposure
Activity
Air pollution
Agent-based modelling
Space-time path
Data poor

ABSTRACT

Long-term air pollution exposure may lead to an increase in incidences and mortality rates of chronic diseases and adversely affect human health. The effects of long-term air pollution exposure have not been comprehensively studied due to the lack of human mobility data collected over a long period. In this study, we develop and apply a personal mobility model to long-term hourly air pollution concentration predictions to quantify personal long-term air pollution exposure for all individuals. We implement our model assuming mobility patterns for commuters and homemakers, and separate between weekdays and weekend. Our results show that NO₂ exposure of commuters are on average slightly higher and vary less spatially as they are exposed to NO₂ at multiple locations.

1. Introduction

Exposure to air pollution has been shown to cause a higher incidence of chronic diseases (WHO; Chen and Goldberg, 2009) including lung (Johansson et al., 2015; Gehring et al., 2013) and cardiovascular disease (Chen et al., 2010). Epidemiological studies often use personal air pollution exposure to identify and quantify the impacts of air pollution on health (Zou et al., 2009; Zhou et al., 2001). These studies mostly require long-term personal exposures (Kan et al., 2012), that is, personal exposure aggregated over a long time span representing a considerable part of a person's life (preferably multiple years), as the effect of air pollution on chronic diseases accumulates over time. Assessment of personal exposures is a challenge as the spatiotemporal variation in air pollution is high. This implies that exposure assessment preferably needs to consider space-time activity patterns of individuals, to enable the integration of air pollution along space-time tracks of individuals.

Although considerable progress has been made (Dias and Tchepel, 2018), we argue that there is still a need for long-term, population-wide, personal exposure assessment techniques that can be applied over large numbers of individuals. These techniques can be used in public health studies to assess personal air pollution across a population, to study differences between population subgroups with different activity schemes, for instance, socio-economic groups, and in studies that need health cohort data enriched with air pollution exposures to study effects

of air pollution on health. These application cases are inherently data poor in particular regarding personal data as it is thus far not feasible to track all persons in such large populations. For instance, the geographically referenced information in many health cohorts is mostly restricted to the home location of individuals, possibly including the occupation of the persons, which is insufficient information to reconstruct their space-time tracks. Therefore, population-wide personal exposure assessment techniques need to rely on methods that are capable of representing space-time tracks using sparse geographical referenced information on individuals. Current approaches that rely on time averaged air pollution at the home location (Strak et al., 2017) are restricted; many studies have shown over- or under-estimation of air pollution exposure as a result of the insufficient representation of personal activity (Gurram et al., 2015; Baxter et al., 2013; Dias and Tchepel, 2018; Tang et al., 2018; Park and Kwan, 2017; Yoo et al., 2015; Dons et al., 2011).

Existing personal trajectory based exposure assessment techniques range from data rich approaches that use time continuous individual measurements, to intermediate data availability approaches that use diary survey data to constrain activity models, to data poor approaches that mainly rely on mechanistic modelling of activity patterns constrained by very limited data at the individual level.

Data rich approaches use continuous time measurements at the individual level. Air pollution exposure can be directly measured using mobile sensors placed in or close to the breath zone of individuals

* Corresponding author.

E-mail address: m.lu@uu.nl (M. Lu).

(Steinle et al., 2013). Alternatively, space-time tracks of individuals can be measured with mobile devices equipped with a GPS receiver (Minet et al., 2018). These space-time tracks are combined with air pollution maps or air pollution calculated at locations visited to integrate air pollution along the tracks to calculate the personal exposure. These approaches enable a detailed assessment of exposure but come with two shortcomings for population-wide long-term exposure assessment. One is that these approaches quantify exposures over the time span for which personal data are available only, typically a few weeks (Dons et al., 2011). Also, these approaches only provide exposures for individuals equipped with a sensor who also need to be willing to share their personal data.

To overcome these limitations, human activities have been reconstructed from activity diaries or surveys (Tang et al., 2018; Møller et al., 2012). Activity models (Miller and Roorda, 2003; Shekarrizfard et al., 2017; Deffner et al., 2016; Gulliver and Briggs, 2005) have been developed that enable the calculation of personal exposures using less detailed data at the individual level, thus making it more feasible to estimate exposures over large populations. The input of activity models entails information from diary surveys consisting of locations visited and origin-destination times. The activity model, then, reconstructs a continuous-time mobility track of each individual. This is done by a simulation model that relies on theory of mobility patterns and space-time accessibility (Nguyen et al., 2011; Gonzalez et al., 2008; Yang et al., 2010; Yu, 2006; Alessandretti et al., 2017; Miller, 1991). Each route simulated can be contingent on, for instance, distance, safety, city infrastructure, and land use (Law et al., 2014). Just like with data rich approaches, personal exposures are calculated by integrating air pollution along space-time tracks. Beckx et al. (2009) applied an activity model (travel forecasting resource) to simulate human hourly activities, but the exposures during the commuting times are not calculated along the routes. Shekarrizfard et al. (2017) assigned the predictions of a travel demand model to a road network to predict a person's hourly trajectories. For each person, the model selects a path from all possible paths by comparing the assigned travel time and the survey travel time. These studies applying activity-based models have focused on simulating hourly activities but pay less attention to representing variation over longer time spans, such as separating between seasons and considering holidays. In addition, the activity models depending heavily on the representative survey sampling of the real activities, which in practice may be a challenge.

In data poor application cases, that is, studies where detailed activity or location data on individuals is lacking, which is the focus of this study, personal exposure assessment has to rely on mechanistic models of space-time activity, and they may need to include probabilistic rules to represent space-time activities that are not exactly known or observed. Limited geographically referenced data at the individual level, in particular, home address and possibly work address, if available, can be used as input to the model. In addition, individual level information that is informative for the space-time activity pattern of an individual, for instance, socio-economic class or age group, can be used to constrain the space-time activity simulation of the individual. This information is often available, for instance, in health cohorts. A rather limited number of data poor approaches have been described. Yang et al. (2018) proposed an agent-based modelling framework for the assessment of exposures to environmental stress. The framework models daily routines, for instance, the probability that a place is visited, of an individual and aggregates the environment an individual is exposed to. Park and Kwan (2017) assumed a daily activity schedule and assessed air pollution exposure as air pollution concentration along the path between home locations and randomly selected work locations. However, they only consider air pollution exposure of a single day, and thus not providing long-term exposure. Also, the framework described in Yang et al. (2010), as well as the model of Park and Kwan (2017), do not provide an estimate of the uncertainty from the process of randomly assuming a working location of the agent it is modelling.

Here, we build on these existing data poor approaches mainly by extending these to long-term personal exposure assessment, so that the exposure of multiple socio-economic groups each with a particular space-time activity pattern could be assessed. Our first research question is how space-time activity of each particular socio-economic group can be simulated in data poor situations. Using a new simulation model for data poor situations, we will then address the second research question: what is the distribution of personal exposures of an urban population, and how do personal exposures vary between socio-economic or age groups and as a function of place of living?

Our approach combines agent-based modelling of human space-time activity with an hourly land use regression (LUR) model predicting air pollution climate, which is the long-term average air pollution for each hour of the day, separately for weekdays and weekends. Personal exposures to these modelled air pollution values are simulated hourly. The agent-based model uses pre-assumed activity schedules and routing information to model human space-time paths and Monte Carlo simulation to account for uncertainties of unknown working locations.

Different gaseous and particulate air pollutants have different spatiotemporal gradients. Hankey and Marshall (2015) found the spatiotemporal variation increases when the particle size decreases and the correlation between particulate pollutants differs. Li et al. (2019) and Luengo-Oroz and Reis (2019) have shown that the Ultra Fine particles (UFP) are highly spatially and temporally dynamic and Luengo-Oroz and Reis (2019) showed alternative commuting routes of bicyclers could lead to substantially different exposure. Pollutants with very high spatiotemporal variation may be infeasible to be mapped at a sufficiently high spatiotemporal resolution for the proposed data poor approach. The NO₂ has a much lower spatiotemporal dynamic compared to the UFP and is commonly measured every 15 min, which makes it a suitable pollutant for the proposed data poor approach and we thus selected it for our study case. In addition, NO₂ is highly traffic related, which makes it suitable to study exposure as a result of different commuting scenarios with the proposed approach. Besides NO₂, other pollutants that have relatively limited small scale spatiotemporal dynamics, for instance O₃ and PM_{2.5-10} (Van den Bossche et al., 2015), can also be modelled with the proposed data poor approach.

This manuscript is structured as follows: Section 1 reviews air pollution exposure models and their limitations to study the health effects of long-term and all-individual air pollution exposure. Section 2 describes our model framework to model long-term air pollution exposure of all-individuals with unknown individual working locations, model implementations of different activity patterns, and sensitivity analysis. Section 3 shows the results of our model and the sensitivity of the model variables. Section 4 discusses the results, advantages and limitations of our model, and gives an outlook. Section 5 finishes with a conclusion.

2. Methods and study area

2.1. Concepts

We argue that, even in data poor situations, it is preferable to incorporate mobility of individuals into the exposure assessment, as this enables accounting for spatiotemporal variation in air pollution along a persons' space-time track. We thus propose an agent-based modelling approach that simulates space-time tracks of each person considered assuming the home address of each person is known as well as information on the type of space-time activity pattern of the person (e.g. a commuter, a homemaker). Uncertainties in the simulation due to sparse data are represented using probabilistic model inputs and parameters. The general framework of our approach is as follows:

- The agent-based simulation of the activity of a person is configured according to the socio-economic group the person belongs to. A socio-economic group is defined here as consisting of people with a

Table 1

Activity calendar of the scenario we studied. Commuters arrive at work at 9 a.m. and leave from work at 5 p.m., the duration of the commute trip is calculated by the model.

	weekdays		weekend	
	time	location	time	location
homemaker	all day	home	8 a.m. - 11 p.m.	1 h outside, other times at home
commuter	12 a.m. - 9 a.m.	home, on road	11 p.m. - 8 a.m.	home
	9 a.m. - 5 p.m.	work		
	5 p.m. - 12 a.m.	on road, home		

similar space-time activity, for instance, people who stay at home and commute to work by bike. The calculated personal exposures are aggregated over multiple years (approximately 5 years), and within this time span, it is assumed that the person stays in the same socio-economic group.

- In the agent-based simulation, the year is subdivided into different day types (e.g. weekday, weekend) that have a distinct activity pattern for the socio-economic group considered. For each day type, the daily mobility of the individuals is simulated, assuming the same mobility pattern applies for all days within a day type. Different seasons that may affect human activity pattern are not taken into account in our current study but could be included in future studies following a similar approach.
- The agent-based simulation is used to calculate the space-time track of individuals over a day (for a particular day type). An individual visits multiple locations. Routes between locations are calculated as the shortest path over the infrastructure network for a particular mean of transport. Depending on the socio-economic group a person belongs to, a time calendar is used as input to the simulation providing departure and arrival times at locations where the individual undertakes a certain activity (e.g. work location, home location) and the geographical position of these locations. Positions that are exactly known, for instance, place of living of a person, are deterministic inputs to the simulation, while positions that are not exactly known, for instance, work location, are defined as stochastic inputs, where locations have a particular probability that they are visited. The agent-based simulation uses a time step of 1 h.
- As in data poor situations, only the home address is known. All uncertainties in the remaining inputs and parameters are analysed using Monte Carlo simulation, in particular, the geographical location of locations visited, the daily time calendar, and parameters such as travel speed and building infiltration factor.
- Long-term (5 years) average air pollution is mapped for each hour of the day, for each day type (weekday, weekend), for each pixel, using a temporal land use regression model. Personal exposures of individuals, then, are calculated by integrating air pollution over the space-time tracks retrieved using agent-based simulation, aggregating exposures for each day type in a year, resulting in long-term average personal air pollution exposures.

2.2. Implementation for homemakers and bike commuters

2.2.1. Study area

We apply our model to the Dutch municipality of Utrecht (population ca. 345,000, area ca. 99.2 km²) and to the air pollutant NO₂, to assess 5-year average air pollution exposure from 2011 to 2016. Utrecht is the fourth largest city of the Netherlands with an infrastructure typical for Dutch and many other European cities, having a dense network of cycling lanes.

2.2.2. Agent-based simulation of individual space-time path

We use the general framework defined in concepts to identify the space-time path for two social-economic groups that have distinct mobility patterns, namely homemakers and bike commuters. To provide

simple and contrasting model framework implementation cases, we further assume homemakers only stay at home and bike commuters cycle to a working place during weekdays. The space-time path of a person is calculated according to a time schedule that is defined in our study based on these assumptions callendar [Table 1](#). For weekdays, it is assumed that cycling commuters commute to work every day while homemakers stay at home. For weekends, both socioeconomic groups are assumed to be outside for 1 h and the remaining time of the day inside.

When a person is at home or work, we assume the person is indoor. The space-time mobility of an indoor environment is represented as a 60 × 60m² spatial window centred at the front door location. An area of 60 × 60m² is used as an approximation of the size of a building and the NO₂ is assumed to disperse at this scale. This square is called the indoor window in this study. Within the indoor window, air pollution is assumed to be the same.

During work days, the bike commuters are assumed to arrive at work at a time ($T, -$) at 9 a.m. and work for 8 h, and cycle with a speed (S , km/h) of 16 km/h ([Woodcock et al., 2009](#)). As the specific working location of each commuter is unknown, a Monte Carlo simulation is used to randomly draw a working location from all the potential working locations in the city in each run. The potential working locations are known and are derived from all the functional buildings of Utrecht from the cadastral dataset ([Kadaster](#)). For each realisation of a working location, the commuting route from home to work is derived using the shortest distance route on roads or bicycle lanes. It is assumed that from work to home the same route is followed. The duration of commuting is calculated as the length of the commuting route divided by the S .

For weekends, it is assumed people do outdoor activities (includes shopping) for an hour between 8 a.m. and 11 p.m. The exposure during this time is represented as a 10 km radius circle centred at the front door home location. This circle is called an outdoor activity window in this study. Within this window, the visiting frequency of each location is assumed to be the same.

2.2.3. Spatiotemporal air pollution mapping

The NO₂ concentration of Utrecht is estimated using LUR (Land Use Regression) ([Hoek et al., 2013](#)). Here we provide the general approach, details, and validation of the results are given in [Soenario et al. \(2019\)](#). A number of 78 NO₂ measuring stations from the Dutch National Air Quality Monitoring network ([for Public Health and the Environment](#)), providing NO₂ measurements every 15 min, are used. The NO₂ is averaged for each hour, month, and weekend/weekday from 01 July 2006 up to 01 July 2011 (5 years). Land use variables used as candidate predictors for the land use regression include traffic, infrastructure, and population within 25, 50, 100, 300, 500, and 1000 m buffers. The NO₂ mapping consists of two steps: the selection of predictors and model fitting. The predictor variables are selected by sequentially applying the Lasso regression ([Tibshirani, 1996](#)) and best subset regression. Then, with selected predictors as independent variables, multiple linear regression models are built for every hour with hourly aggregated NO₂ measurements as the response variable. The LUR is run for each 5 m pixel and then averaged over 20 m grid cells for exposure assessment. The 20 m grid cell size is chosen as it would be difficult to simulate the

space-time activity track at a higher level of spatial detail. It would, for instance, require simulating on which lane of the street a person cycles. Also, the computation time is still feasible on a standard workstation used here, whereas computing times would significantly increase for simulations at higher resolutions. The estimated NO₂ concentration for all months is averaged to represent 5-year climate for each hour of the day separating between weekdays and weekend.

2.2.4. Personal exposure assessment

We implement our agent-based simulation model on a 20 m resolution grid, assuming spatially homogeneous air pollution within each grid cell. We use e to denote the exposure assessed in general and the exposure that is assessed in each Monte-Carlo run and E to denote the final exposure assessed and the exposure that is assessed from all the Monte-Carlo runs. In general, the air pollution exposure e ($\mu\text{g}/\text{m}^3$) of a person over a certain time period can be calculated from air pollution concentration along the space-time path of the person over the period considered (Hertel et al., 2001):

$$e = \frac{\sum_{i \in J} (C_i \cdot t_i)}{\sum_{i \in J} t_i} \quad (1)$$

In (1), C_i indicates air pollution concentration averaged over each microenvironment i for the time span the person visits the microenvironment, J indicates all microenvironments that form a person's space-time path over the time period considered, t_i indicates the time a person spends in a microenvironment i .

It is assumed that the indoor air pollution concentration is proportional to ambient air pollution with an indoor proportion R . The value of R varies with compounding factors such as traffic, cold and warm seasons, building ventilation, and infiltration (Rivas et al., 2015; Meier et al., 2015; Yang et al., 2004; Batty et al., 2003). In our study, we used a constant R of 0.7 based on the proportion identified by Yang et al. (2004); WHO; Rivas et al. (2015).

In our study, the 60×60 m indoor window forms the microenvironment when a person is inside. The C_i for a person at home (C_{home}) is calculated as the indoor air pollution concentration averaged over the time period that the person is at home. The C_i for a person at work (C_{work}) is calculated in a similar fashion using the indoor air pollution concentration at the work location and the time at work. When a person is commuting, each grid cell of the commuting route (r) is a microenvironment. The C_i at each route cell, denoted by $C_{i \in r}$, is calculated as the average ambient NO₂ concentration over the time span that a person passes the cell. As we assumed in this study a constant commuting speed over each route cell, the exposure e when a person is commuting is the mean of $C_{i \in r}$, denoted as \bar{C}_r .

The NO₂ exposure is calculated for weekdays and weekends separately, and the NO₂ exposure representative for a year, E_{NO_2} ($\mu\text{g}/\text{m}^3$), is calculated as temporally weighted aggregation of NO₂ exposure in weekend and weekdays,

$$E_{\text{NO}_2} = \frac{E_{\text{wd}} \cdot t_{\text{wd}} + E_{\text{we}} \cdot t_{\text{we}}}{t_{\text{wd}} + t_{\text{we}}} \quad (2)$$

where t_{wd} and t_{we} are the duration (h), calculated over one year, of weekdays and weekends, respectively. E_{wd} ($\mu\text{g}/\text{m}^3$) and E_{we} ($\mu\text{g}/\text{m}^3$) are the NO₂ exposures representative for weekdays and weekends, respectively.

For homemakers, E_{wd} is calculated as the C_{home} of weekdays. For commuters, E_{wd} is the median of NO₂ exposure calculated in each Monte Carlo run,

$$e_{\text{wd}} = \frac{C_{\text{home}} \cdot t_{\text{home}} + C_{\text{work}} \cdot t_{\text{work}} + \bar{C}_r \cdot t_{\text{road}}}{t_{\text{home}} + t_{\text{road}} + t_{\text{work}}} \quad (3)$$

where t_{home} , t_{work} and t_{road} indicate the duration when a commuter is at

home, work, and commuting, respectively.

The e_{wd} is calculated in 12 Monte Carlo realisations. The number of realisations is determined by randomly sampling 5000 residential locations from all the residential locations and running different numbers of Monte Carlo runs and comparing the distributions of the E_{wd} of sampled locations. In our study, 12 realisations are used as we found negligible differences in the E_{wd} distribution calculated using 12 to 30 runs.

For weekends, the outdoor activity window (section 2.2.2) forms the outdoor microenvironment when the person is outdoor. The E_{we} for both homemakers and commuters is calculated as:

$$E_{\text{we}} = \frac{C_{\text{outdoor}} \cdot t_{\text{outdoor}} + C_{\text{home}} \cdot t_{\text{home}}}{t_{\text{outdoor}} + t_{\text{home}}} \quad (4)$$

where the C_{outdoor} indicates the mean NO₂ concentration of the outdoor micro-environment averaging over the time t_{outdoor} when a person is doing outdoor activity.

2.2.5. Personal exposure distribution over the population

In pe, we calculated the E_{NO_2} of a person as a function of his or her home location. This is calculated for each home location and gives a map of personal exposure over the city, which enables studying the spatial pattern of personal exposures over the city. In addition to analysing this spatial pattern, it is relevant to study the distribution of personal exposures over the population. Retrieving actual distributions of homemakers or bike commuters would require knowing the group a resident belongs to. As this is currently not available at a sufficient level of detail, we perform the analysis assuming either all residents are homemakers or bike commuters. Under the assumption that there is no difference in the spatial pattern of homemakers and bike commuters, this still gives an indication of the difference in the distribution of the personal exposures between both groups.

To calculate personal exposures over the population, population data in a raster format at 100 m resolution is acquired from the national statistical office (CBS) for the year 2016. As our model assumes individuals living at each 20 m pixel location are exposed to the same air pollution, this 100 m population map is down-scaled to 20 m resolution. For each 100 m pixel, the population is distributed over the 20 m pixels containing residential buildings within the 100 m pixel, assuming that the same number of people lives in each 20 m pixel containing residential buildings within a 100 m pixel. Then, the E_{NO_2} of each grid is assigned to each person.

2.2.6. Implementation

The model is implemented in the data analysis environment and programming language Python (Team) with the PCRaster library (Karszenberg et al., 2010). The complete residential (home) and functional building (work) front door locations are from Dutch cadastral datasets Kadaster. The commuting routes (vectors) are retrieved using a routing engine (Contributors, b) on the OpenStreetMap bicycle profile (Contributors, c). The routes are converted to rasters and resampled to the 20 m grid also used for NO₂ concentration mapping. The OpenStreetMap is downloaded from (Contributors, a). The shortest route between home and work locations is used as commuting routes. All the route grids from home locations to work locations are stored in an HDF5-based database (de Bakker et al., 2017), to avoid re-calculating the routes during different commuting hours. The model output is a map that gives E_{NO_2} of a person living at a particular location on the map. The model is only evaluated for locations that contain residential buildings.

2.3. Sensitivity analysis

Sensitivity analysis is performed to evaluate the sensitivity of the calculated personal exposures to changes in S , T and R . To reduce run

times, the sensitivity analysis is applied on a spatially stratified random sample of 1000 people living at 1000 residential front door locations. The residential area is stratified into 50 equal-sized rectangular blocks; in each block, 20 random locations are drawn. The residential front door locations nearest to these locations form the sample dataset.

The annual average exposure E_{NO_2} of each of these 1000 people are assessed for different values of S , T , and R . For all the situations, one variable is controlled and the others are fixed, Table 2 shows the setting of the variables in each situation. Varying S and T may only negligible affect the E_{NO_2} as for a commuter only approximately 5–10% of the day is spent on commuting. To further understand the effects of T on E_{NO_2} , we additionally calculated NO_2 exposure for the period of time a person commutes.

2.4. Uneven probability of working locations

In practice, the probability that commuters work at certain locations is spatially uneven and this may affect the exposure that is assessed. To demonstrate that our model framework is flexible in modelling real-life scenarios and that it could be used to study different working groups, we implemented two additional scenarios. The first scenario defines the current bike commuter profile in more detail (only possible if data is available to do so) by incorporating additional information for more sophisticated representation of the work locations. The commuter profile is considered as consisting of two subgroups, university students and other commuters. In this scenario, we assume a probability of 0.15 that the commuter studies at university. The probability of 0.15 is estimated according to the number of students (Expatica) and the age structure (Utrecht) in Utrecht. The second scenario creates a smaller, more specific group, that is students only. In this scenario, we assumed all the commuters are students and commute to university. Both scenarios are run for a sample of 1000 persons that is used in the sensitivity analysis (sensit) and the default activity schedule is used.

2.5. Comparison with static exposure assessment

Epidemiological studies often assume human space-time activities are static and use long-term average air pollution concentration at front door home locations as air pollution exposure. As a comparison to exposures assessed using our agent-based modelling approach, we apply two different static exposure assessment techniques on two different static air pollution data sets. The first, most widely used, static exposure assessment technique assumes that exposure equals the long-term air pollution at the location of the front door. We calculate this exposure by reading the air pollution value at the location of the front door from the static air pollution map. As front-door exposure assessment neglects aggregation of ambient air pollution over a microenvironment visited by a person, we introduce a second static exposure assessment technique that calculates exposure as the average static air pollution within a $60\text{ m} \times 60\text{ m}$ window centred at the front door of a persons' home. The indoor exposure factor R is excluded from the static exposure

Table 2

Variable settings of cycling speed (S , km/h), time arriving at work (T , -), indoor ratio (R , -), to analyse the sensitivity of each variable. The standard run uses the centre value of each variable.

Testing variable	S	T	R
S	15	9 a.m.	0.7
	16	9 a.m.	0.7
	17	9 a.m.	0.7
T	16	8 a.m.	0.7
	16	9 a.m.	0.7
	16	10 a.m.	0.7
R	16	9 a.m.	0.6
	16	9 a.m.	0.7
	16	9 a.m.	0.8

assessments as this would require an estimate of the time spent inside, which is typically not done in static exposure assessment.

Both exposure assessment methods are applied on two different static 5 m resolution NO_2 concentration maps. One is derived from our dynamic NO_2 data set used in our agent-based modelling by averaging NO_2 over all hours of a year. In addition, we calculate exposures using an existing 5 m NO_2 concentration map Schmitz et al. (2019) derived from the ESCAPE LUR model (ESCAPE).

3. Results

3.1. Ambient NO_2 concentration mapping

The variable selection process selects four variables to predict NO_2 concentration: heavy traffic load within a 50 m buffer, total major road length within a 50 m buffer, and total road length within 1000 m and 5000 m buffers. The linear regression fitted for each hour resulted in 576 ($24 \times 2 \times 12$) different linear regression models, one for each hour of the day, for weekdays and weekends, and for each month of the year. The NO_2 maps are averaged to yearly average maps. The coefficients and the adjusted R^2 for each model, as well as the predictor maps, are provided at <https://github.com/pcraster/gghdc-spatio-temporal-lur-nl/blob/master/gghdc-dev-master.zip>.

3.2. NO_2 exposure assessment

An example of hourly NO_2 exposure from multiple realisations for a randomly selected commuter is shown in Fig. 1. The figure shows variation in exposure as a result of different working locations and routes. For almost all the realisations, the exposure is highest during commuting hours (around 8–9 a.m., and 5–6 p.m.).

Fig. 2 shows E_{NO_2} assessed for the commuters and homemakers. A clear spatial trend can be observed for homemakers. E_{NO_2} is high for people living along major roads and in the city centre and decreases with distance away from the city centre and main roads. The decrease in E_{NO_2} away from the city centre and major roads can be observed for commuters as well, but the magnitude of the variation is somewhat smaller compared to the homemakers. This is also shown by the spread of the E_{NO_2} distribution, the interquartile range, i.e., the difference between the value of the first and the third quartile, which is 1.9 for commuters and 2.9 for homemakers. The median E_{NO_2} for homemakers ($20.64\ \mu\text{g}/\text{m}^3$) is

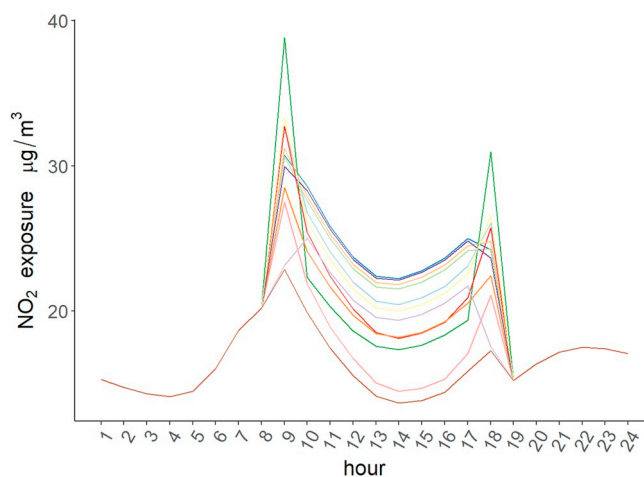


Fig. 1. An example time series of NO_2 exposure ($\mu\text{g}/\text{m}^3$) of a commuter in multiple realisations (i.e., Monte Carlo runs, in different colours) at a single residential front door location. The exposure is representative for the total average exposure over the 5 years. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

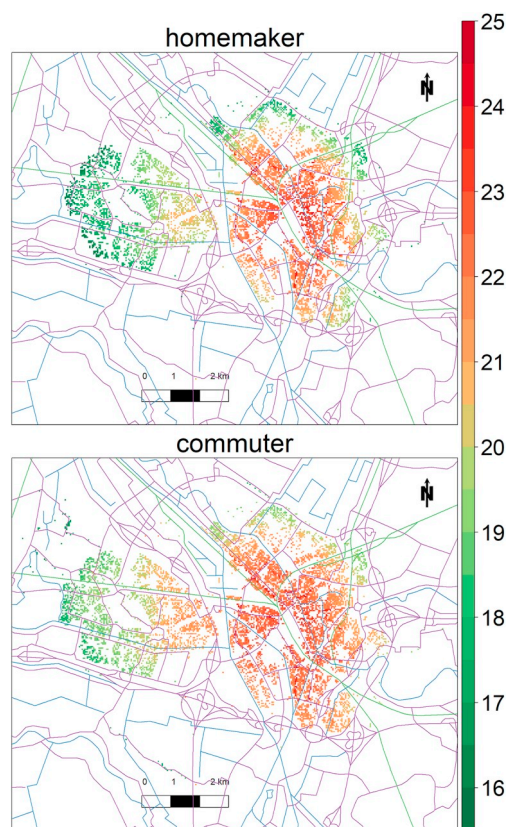


Fig. 2. NO_2 exposure ($\mu\text{g}/\text{m}^3$) of homemakers and commuters. The exposure is representative for the total average exposure over 5 years. The map shows for each location the exposure of a person living at the location. For a clearer visualisation, the higher and lower 1% quantiles are removed. The purple, green, and blue lines indicate major road, railway, and waterway, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

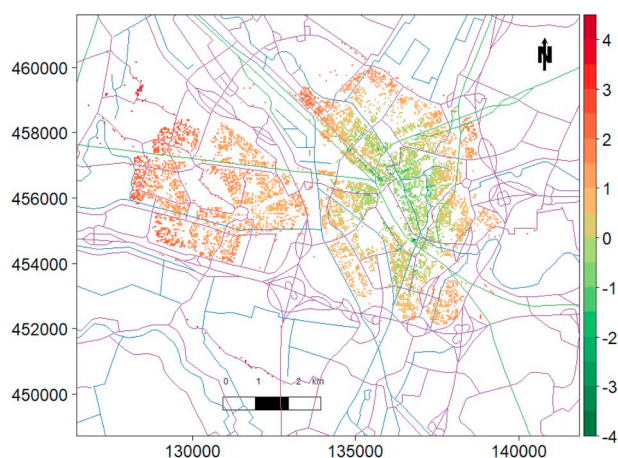


Fig. 3. The differences of NO_2 exposure (E_{NO_2} , $\mu\text{g}/\text{m}^3$) between commuters and homemakers (E_{NO_2} of commuters - E_{NO_2} of homemaker). The exposure is representative for the total average exposure over 5 years.

lower than for commuters ($21.35 \mu\text{g}/\text{m}^3$). The E_{NO_2} of commuters and homemakers is highly correlated (Pearson correlation coefficient: 0.98). Differences between the E_{NO_2} of homemakers and commuters (Fig. 3) are most distinctive in the suburban areas, where the commuters have a higher E_{NO_2} . The disparity increases at the city centre and close to roads.

Fig. 4 shows North-South and West-East transects of E_{NO_2} for

commuters and homemakers. For homemakers, the E_{NO_2} shows an increasing trend from rural areas in the west to the city centre, and a decreasing trend further away from the city centre. The trends on both transects are comparable, although E_{NO_2} along the North-South transect shows smaller variation. The spikes in E_{NO_2} along the North-South transect coincide with home locations that are very close to the major roads. Fewer fluctuations can be observed in E_{NO_2} for commuters compared to homemakers.

The distribution of the E_{NO_2} over the population of Utrecht shows a distinct difference between homemakers and bike commuters (Fig. 5). While on average homemakers and bike commuters are exposed to comparable values of NO_2 , the range in the personal exposures over the population of homemakers is larger, resulting in a larger proportion of the homemakers, compared to bike commuters, that is either exposed to relatively low or relatively high values of NO_2 . For instance, the proportion of homemakers with E_{NO_2} above $25 \mu\text{g}/\text{m}^3$ is higher than for bike commuters. More spikes can be observed in the left half of distribution for homemakers, indicating a large proportion of residences living in areas with relatively low NO_2 .

3.3. Sensitivity analysis

The sensitivity of NO_2 exposure calculated for different settings of S , T , and R shows the influences of changing these variables (Fig. 6). Varying the R leads to considerable changes in E_{NO_2} , the medians of E_{NO_2} are 15.4 , 18.2 , and $21.1 \mu\text{g}/\text{m}^3$ with R values of 0.6 , 0.7 and 0.8 , respectively. Changing S and T has a minor effect on E_{NO_2} . Setting the S to 15 , 16 , and $17 \text{ km}/\text{h}$ result in median values of E_{NO_2} of 18.25 , 18.20 , $18.15 \mu\text{g}/\text{m}^3$, respectively. Setting the T to 8 , 9 and 10 a.m. lead to the median values of E_{NO_2} of 18.21 , 18.20 , and $18.17 \mu\text{g}/\text{m}^3$, respectively. The NO_2 exposure during commuting times varies significantly with T (Fig. 7), which shows that assuming 8 h at work, a person exposes to the most NO_2 when arriving at work at 8 a.m. , and to the least NO_2 when arriving at work at 10 a.m.

3.4. Uneven probability of working locations

The E_{NO_2} assessed varies with scenarios that assume different working location distribution probabilities. The scenario that assumes all the commuters are students (Fig. 8) shows the commuters generally are exposed to lower NO_2 compared to assuming 0.15 probability of commuters are university students and assuming an even probability of working locations. The difference between assuming 0.15 probability of commuters are university students and assuming an even probability of working locations is small and the E_{NO_2} assessed for the latter is slightly higher. Fig. 9 shows how frequent a road is passed by in the three scenarios. When the university buildings have higher probability, the roads to the university are more frequently taken.

3.5. Comparison with static exposure assessment

Exposures assessed through a static technique are always below those calculated by our activity based technique which is mainly due to the exclusion of the indoor proportion factor R in the calculation of the static exposures (Fig. 10). This is an important observation, but as variation in exposures between individuals is in many epidemiological studies at least as important as absolute values, we will focus our analysis on these variations, and how well static approaches are capable of reproducing the variation in exposure assessed by our activity based techniques.

Static exposure assessment techniques are capable of explaining 55% up to 99% of the variation in personal exposures assessed by our activity based exposure assessment technique. These percentages depend on the static exposure assessment technique that is used and in particular the NO_2 map from which exposures are calculated.

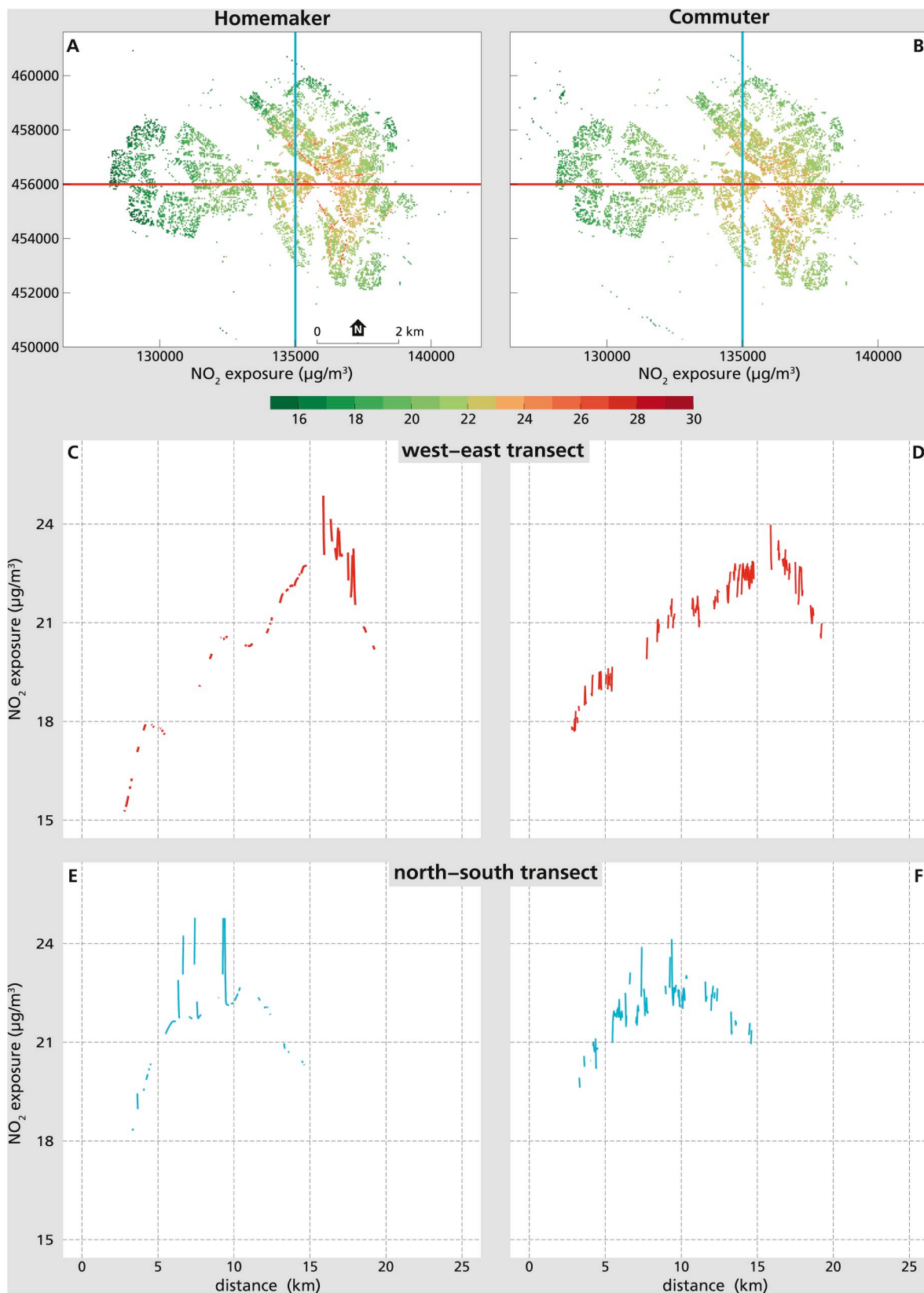


Fig. 4. NO₂ exposure (E_{NO_2} , $\mu\text{g}/\text{m}^3$) of homemakers and commuters assessed along east-west and north-south transects. The exposure is representative to the total average exposure of the 5 years. The red and blue lines in A and B indicate the location of the west-east and north-south transects. C, E: E_{NO_2} along the west-east transect, and the north-south transect, for homemakers. D, F the same for commuters. The E_{NO_2} is only plotted for pixels containing residential front door locations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

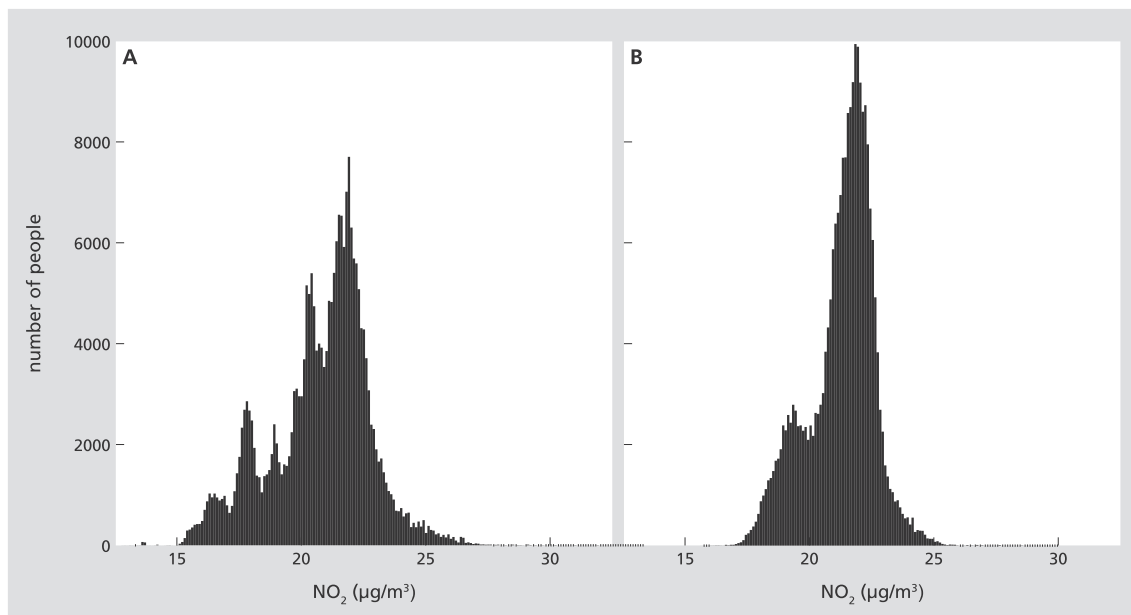


Fig. 5. Distribution of personal exposures (E_{NO_2} , $\mu\text{g}/\text{m}^3$) over the population, A: homemakers, B: commuters. The exposure is representative to the total average exposure of the 5 years.

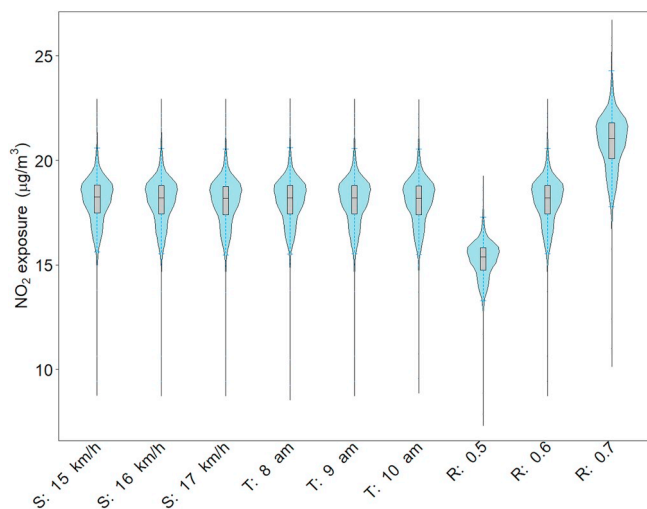


Fig. 6. Sensitivity analysis of the effects of cycling speed (S, km/h), time arriving at work (T, -), and indoor ratio (R, -) to NO_2 exposure ($\mu\text{g}/\text{m}^3$) from sampled residential front door locations. Each of the boxplots shows the distribution of the NO_2 exposure values calculated over sampled residential front door locations. The exposure is representative to the total average exposure of the 5 years. Line in the centre of the grey box: median; top and bottom boarders of the grey box: first and third quartile; blue violin shape: full distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Static exposures calculated from the same NO_2 data (but averaged over time) as used in our activity based approach are capable of explaining more than 93% of the variation in the exposures calculated by our agent-based model. Static exposures from our NO_2 data set almost completely correspond to our Homemaker exposures, which is as expected as the Homemaker profile does not include considerable mobility and thus is properly represented by NO_2 at the home location. Compared to the use of front door exposures, including a 60 m microenvironment in the static exposure assessment increases the similarity between the static exposures and activity based exposures.

The discrepancy between activity based exposure assessment and

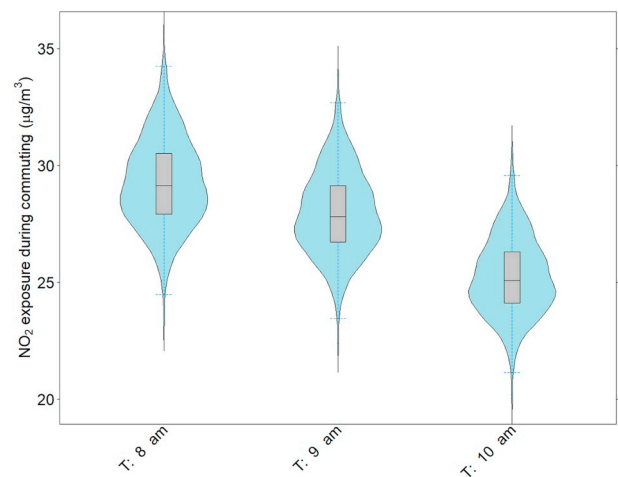


Fig. 7. Sensitivity analysis of the effects of time arriving at work (T, -) to average NO_2 exposure ($\mu\text{g}/\text{m}^3$) assessed *during commuting time*. The centre lines of the grey boxes indicate the median, the top and bottom boarders indicate the first and third quartile, and the violin shapes show the distribution of NO_2 exposure assessed for the 1000 sampled residential addresses. The exposure is representative to the total average exposure of the 5 years.

static exposure assessment considerably increases when using another, independent, air pollution data set for the calculation of the static exposures (Fig. 10). The static exposures calculated from ESCAPE explain 51%–53% of the variation (calculated as the square of the correlation) in the exposures calculated using the activity based model. Explained variation increases again, here with about 10%, when using the 60 m microenvironment instead of front door exposure. In general, the ESCAPE exposures are somewhat higher on average and show more variation than those calculated using our NO_2 data set. This is most likely due to the steeper increase in the NO_2 towards roads in ESCAPE compared to our data, which is due to the inclusion of a 25m buffer in the LUR of ESCAPE, which was not selected by our variable selection procedure used to create our NO_2 map.

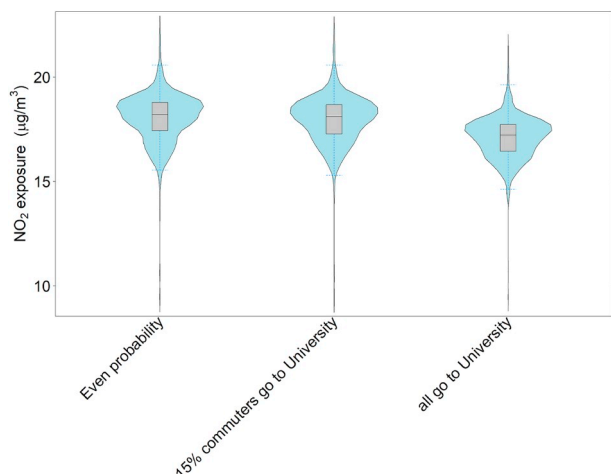


Fig. 8. NO₂ exposure ($\mu\text{g}/\text{m}^3$) assessed for the 1000 sampled residential addresses for the scenarios of commuters commuting to working locations with equal probability, 0.15 probability of commuters are university students, and all the commuters are university students, using the default activity schedule settings. The exposure is representative to the total average exposure of the 5 years.

4. Discussion

We proposed a process-based stochastic model to model human activity patterns and applied our model to hourly NO₂ concentration maps predicted using LUR to assess long-term NO₂ exposure of the personal exposure of commuters and homemakers. We implemented two scenarios, one brings the representation of general bike commuter closer to the reality and the other creates a more specific profile representing university students only. We found that the variation in personal exposures for homemakers was larger than for bike commuters. This is because exposure of a homemaker mainly depends on ambient air pollution concentration at the home location, which may vary

considerably between persons, while exposure of a bike commuter is an average of air pollution values at locations visited over the day and this average is more similar between bike commuters.

Including space-time activity patterns in air pollution exposure assessment enables the integration of spatiotemporal varying air pollution along tracks visited by individuals. This has mainly been done in studies that have access to measured tracks, for instance from persons equipped with GPS receivers, or in studies that have access to rich information on activity patterns from which tracks can be derived, for instance using travel diaries. In this study, we have shown that the inclusion of space-time activity patterns is also feasible in data poor situations where, for instance, and as in our study, only home addresses and all potential working locations of the population are known. The exposure assessment, then, has to rely on assumed travel behaviour of specific socio-economic groups, combined with probabilistic modelling of the unknown factors. The latter has been applied to our study for the working locations of the bike commuters. Although we did compare our personal exposures against those calculated at front door locations, we had no access to observed personal exposures in our study population and validation of our results against empirical data is not possible. An essential step for future research, thus, is to compare modelled personal exposures against observed exposures, throughout a wide range of socio-economic groups. At the same time, there is room for considerable improvement of and extensions upon the approach presented here. One is to run simulations for more specific socio-economic groups potentially based on different occupations and ages (Lee et al., 2013) which allows to better constrain model inputs. Following our approach to simulate students commuting to the university campus, the set of potential working locations could be made specific for other, smaller, socio-economic groups, where blue-collar workers, for instance, would commute to areas containing offices. This is however only possible if data are available on the distribution of various work locations over the city. In a similar fashion, the daily activity calendar used as input to the simulation could be made more specific for particular socio-economical groups. Other activities, for example shopping, could be added to the activity calendar. The activity schedule regarding grocery shopping could differ between different socio-economical groups and depend on

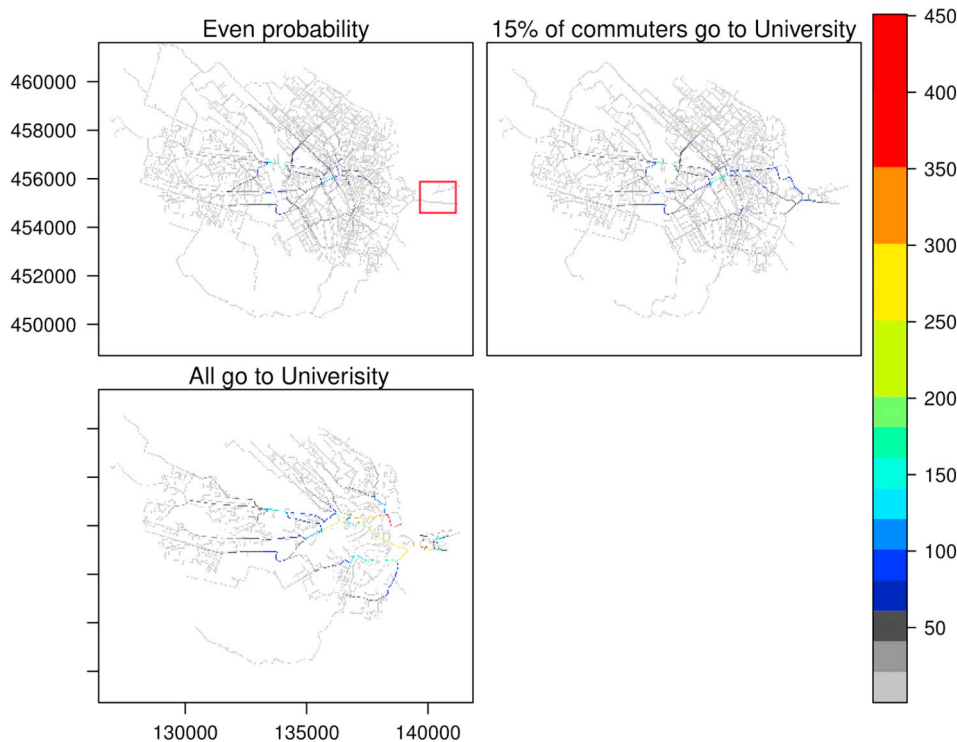


Fig. 9. The number of times that roads are visited during trips between home and working locations. The values represent the number of individuals (out of the 1000 randomly sampled individuals for which the simulation has been run) that visit a road segment. For instance, a value of 500 indicates that half of the persons cross that road segment. A: Commuters commuting to working locations with equal probability, B: 0.15 probability that a commuter is a university student, C: all the commuters are university students. The red box in A indicates the university campus. The exposure is representative for the total average exposure over 5 years. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

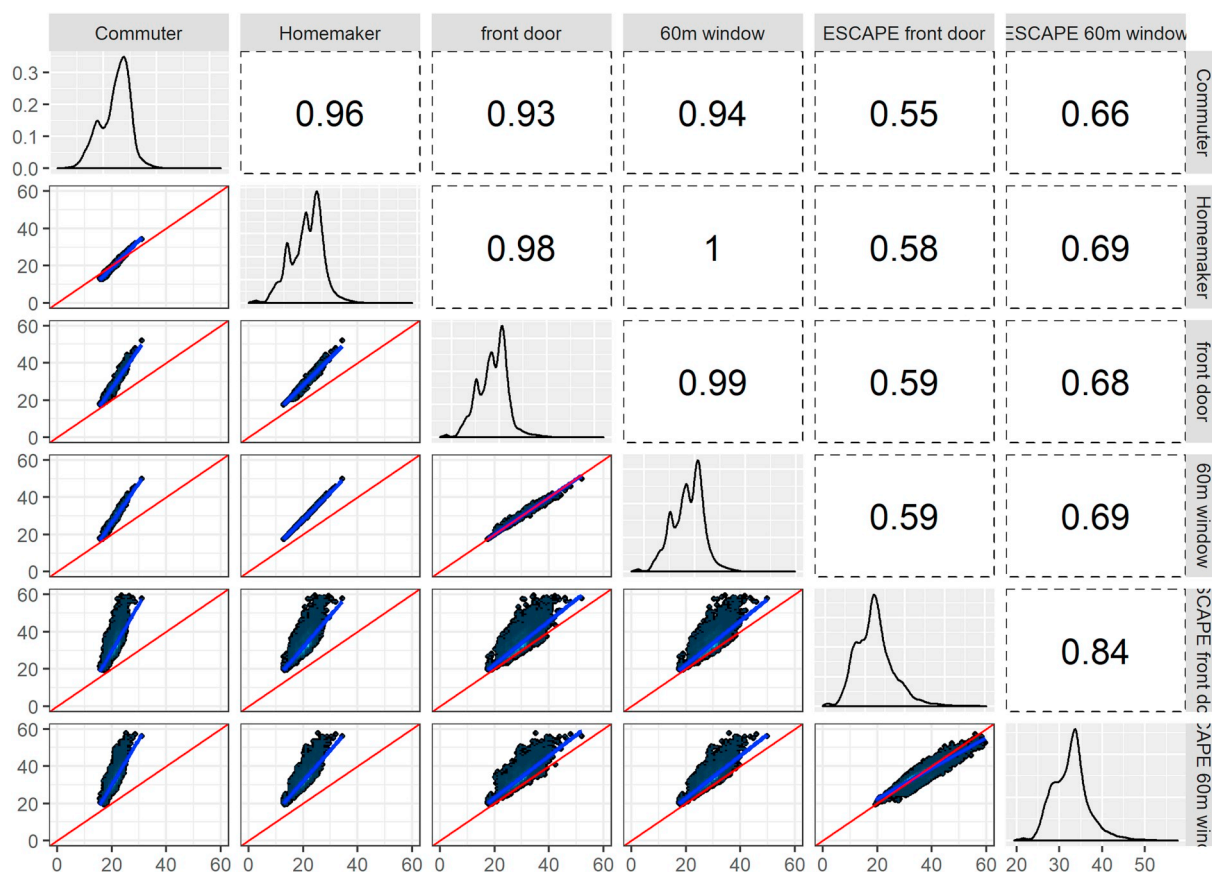


Fig. 10. Comparison between the NO₂ exposure calculated using our method for bike commuters (commuter), homemakers (homemakers), conventional static method (front door) using time-averaged exposure at front door locations, our 5 m NO₂ prediction averaged in 60 m windows (60m window), the original 5 m resolution NO₂ map predicted using the LUR model from the ESCAPE project (ESCAPE front door), the ESCAPE 5 m NO₂ prediction averaged in 60 m windows (ESCAPE 60m window). The exposure is representative to the total average exposure of the 5 years. The diagonal plots show the distribution of NO₂ exposures using different methods. The values on the right panels are the R² calculated between the paired predictions. The red line is the 1 to 1 line and the blue line is the regression line. Light blue indicates a high density of points. Distributions are shown over all the grid cells consisting of home addresses, altogether 36741 values. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the food environment. Another example would be to separately model part-time workers and full-time workers as these will have different activity patterns. Also, we currently neglect effects of seasonality and weather on the space-time activities of individuals. In warmer seasons, people are more likely to stay outdoor, especially in the weekends. Our model could be used to study how the air pollution exposure is affected by the seasonality of space-time behaviour.

Another important extension, required to make our exposure assessment technique applicable to a wide range of studies, is the extension of the number of socio-economic groups such that the exposure can be assessed for the complete population. This would require including other means of commuting (train, car) (Beckx et al., 2009).

Our results show that NO₂ exposure of a person is dependent on the socio-economic group someone belongs to as well as the residential location. The residential location is most important, at least when considering only bike commuters and homemakers. Long-term personal exposures vary between 16 and 25 $\mu\text{g}/\text{m}^3$ depending on the location of living. The socio-economic group is relevant as well and homemakers may have an exposure up to 4 $\mu\text{g}/\text{m}^3$ higher (in the city centre) or lower (in suburbs) compared to bike commuters living at the same location. The 4 $\mu\text{g}/\text{m}^3$ has a limited health impact for an individual but may need attention on the public health level. The combined effect of spatial variation and differences in activity schedules, however, result in small differences between homemakers and bike commuters when considering the distribution of expected NO₂ over the city population. The results of our study show that the differences between homemakers and

bike commuters are mainly in a lower E_{NO_2} range for commuters, while the difference in the median between homemakers and bike commuters is small. Other studies find results in terms of the changes in NO₂ magnitude and variation when mobility is considered similar or dissimilar to ours, and it seems this depends on the exposure assessment technique used. Techniques that aggregate exposures at different locations as a result of assumed human behaviours have shown consistent findings with ours, examples include Park and Kwan (2017), whose NO₂ exposure model is based on pre-assumed activity schedules and Tang et al. (2018), who assess NO₂ exposure using a time-weighted aggregation to represent human mobility. Techniques that use activity simulation models based on surveys and activity diaries have shown inconsistent results when the mobility of persons are included. Considerably lower variation and magnitude in NO₂ when mobility is modelled is found by Smith et al. (2016) and higher variation and lower magnitude when mobility is modelled is found by Mølter et al. (2012), who used a micro-environment exposure assessment model to assess NO₂ exposure of children. The different conclusions reached between using different techniques could be caused by how the activities are modelled and the surveys that are included in the activity simulation models. In our scenario of bike commuters, every commuter is assumed to have the same activity schedule, which differs from the study of Smith et al. (2016) that assigned different activity schedules to different persons. This may increase the variation in the NO₂ exposure between commuters. Our model is developed based on the assumption that there is insufficient survey or activity diary data available (i.e., in the data pool

situation) while the activity simulation models are based on survey and activity diary data as inputs and these field observations will affect the model results. Furthermore, [Smith et al. \(2016\)](#) uses a spatially varying indoor-outdoor ratio while our study and [Park and Kwan \(2017\)](#) considered a fixed indoor-outdoor ratio. Moreover, [Smith et al. \(2016\)](#) considers different transportation modes and in our study we only implemented the transportation scenario of cyclists. Also, the residential locations and the air pollution map may also contribute to the differences. In our study and [Smith et al. \(2016\)](#), despite different NO₂ prediction models are used to predict NO₂ (i.e., LUR and mechanic model, respectively), the distributions of the number of people over predicted NO₂ concentration are both close to Gaussian, indicating the effects from residential locations and air pollution maps to the differences may be small. In future studies, sensitivity analysis should be conducted to systematically compare these methods.

Although validation of our modelled exposures is required against long term measurements of exposure, one can argue that almost by definition the personal exposures calculated by our activity based model are closer to each scenario compared to those calculated using static approaches, as our activity based method incorporates mobility as well as temporal variation in air pollution. If we follow this rationale, the question then is how well static approaches, which are commonly easier to implement, are capable of reproducing the exposures modelled using our activity based model. Our evaluation of the static approaches shows that, when using the same source data for the air pollution mapping, static approaches give results that are only somewhat different from our activity-based approach, and the differences may be acceptable when using exposures in epidemiological studies. However, this depends on the spatial pattern in air pollution - when spatial variation in air pollution is higher, static approaches may perform worse compared to our simulations. This partly explains the much lower correlations between static exposure calculated from the independent ESCAPE model and activity-based exposure calculated from our NO₂ data set. As the ESCAPE NO₂ map shows stronger increasing trends in NO₂ towards roads compared to our data set, static exposure assessment with ESCAPE results for residential locations close to roads in considerably higher values for the exposure compared to the activity based approach which averages out NO₂ over larger areas, representing activity spaces of persons. The relevance of the spatial aggregation is confirmed by comparing static exposures calculated at the location of the front door with those averaged over a 60 m × 60 m window centred at the front door. For ESCAPE, the R² value between these two exposures is only 0.84, which shows that the spatial aggregation has a considerable effect. To conclude, our comparison with static approaches show that much care should be taken with the use of front door exposures, as these do not in any way take into account the fact that humans are mobile. However, a static approach that uses a small window (e.g. 60 m) as used here may give acceptable results for epidemiological studies but further research is needed here on the effect of the quality of the air pollution mapping v. s. the sophistication in the representation of space-time activities in exposure assessment.

Successful implementation of activity-based exposure assessment is contingent on the availability of hourly air pollution concentrations maps representing long term average temporal trends and presumed daily routines of the socio-economic groups considered. In our study, we used hourly NO₂ exposure predicted from LUR models. The result of our model will vary with different input air pollution concentration maps from different air pollution predicting techniques, and this evaluation is our next step. The model can be used to quantify joined effects on personal exposures of human mobility patterns and spatiotemporal variation of ambient air pollution concentration. [Yoo et al. \(2015\)](#) studied these joined effects using simulated data; an interesting comparison could be made to test these effects with spatiotemporal trajectories modelled with agent-based models and air pollution concentration maps predicted with various statistical and dispersion models.

A number of improvements can be made to our approach. As the

indoor factor affects significantly the accuracy of air pollution assessed and varies between seasons of the year, the model could be extended to consider different indoor factors during different seasons of the year. In addition, different indoor factors are to be used for different types of building (e.g., residential buildings vs. office buildings) as the penetration rate of ambient air pollution may be different between buildings. Also, the representation of the commute trip could be improved by including multiple possible routes instead of the shortest route as well as a more precise representation of the time of the commute trip, as during rush hours, temporal variation in air pollution is high ([Zhang et al., 2011](#)). Based on the socio-economical group we defined and studied, the contribution of commute trips to the long-term average personal exposure is very small to almost negligible, as shown by our sensitivity analysis. This result can be explained by the short duration of the commute trips compared to other activities. When peak exposures need to be considered in the health assessment, however, commute trips do become very relevant in the analysis as a result of the high air pollution values along the roads. Our model is capable of assessing personal exposure of other air pollutants. The estimation accuracy of air pollution exposure may vary between different pollutants e.g., traffic related pollutants sourced from motor vehicles emissions and have a short degradation period (e.g. NO₂) may have high spatiotemporal variation ([Johansson et al., 2015](#); [Elliot et al., 2000](#)). In addition, different temporal aggregations associate differently to health effects ([Darrow et al., 2011](#)), e.g. the peak concentration of a day, and requires knowledge of human space-time mobility.

In our model framework, we used a simulation time step of 1 h, and thus air pollution as well as personal activity could vary at this level of temporal detail. This time step is chosen because this level of aggregation enables the representation of most of the spatial and temporal variation both in the air pollution and in the activity of each person. At higher temporal resolutions, the simulations would become intractable, in particular because of long run-times, given the fact that a very large number of routes and exposures need to be calculated. In our study, if the commuting time is shorter than 1 h, the rest of the time the commuter is assumed to stay at the original place. If the commuting time is longer than 1 h, we assumed the commuter left the original place earlier.

Computationally, a challenge for applying our method to a larger scale (e.g. country scale) is to save the routes of all people to avoid re-acquisition of the routes for dynamically calculating the NO₂ over time. The file to store routes can be huge for large-scale studies and more Monte Carlo simulations. In our study case, we stored each route as a dense raster and the routes file is 230 GB for all of the 36741 residential locations for one realisation. The storage will be greatly reduced if these routes can be stored as sparse arrays or vectors.

5. Conclusion

We proposed an activity-based stochastic model to quantify long-term all-individual air pollution exposure. The uncertainty of unknown personal working locations is addressed using a Monte Carlo simulation. Our model continuously assesses air pollution exposure over time, separating between weekdays and weekends and can be extended to inter-annual activities such as holidays and vacancies. Compared to homemaker, the bike commuters are in general exposed to slightly higher NO₂ due to their exposure during cycling, and there is less variability among residences. The value of the indoor factor affects significantly the exposure assessed. Bike commute trips have a relatively minor contribution to the long-term average exposure as they are short compared to other activities. The model assessed exposure varies depending on the space-time activities, particularly the amount of time a person spends outdoor and indoor. Our model is the first step towards modelling long-term personal air pollution exposure. The model can be extended to model air pollution exposure of specific space-time mobility patterns of different social economic groups and the modelled exposure

could be applied to cohorts studies to study the relationship between air pollution exposure and health under different space-time activity scenarios.

Acknowledgement

This research is funded by the Global Geo Health Data Center (Utrecht University) and the Startimpulsprogramma Meten en Detecteren van Gezond Gedrag (Dutch Science Foundation). We are grateful to the contributions from the reviewers. We thank Ton Marcus for his advises and contributions on improving the figures.

References

- Alessandretti, L., Sapiezynski, P., Lehmann, S., Baronchelli, A., 2017. Multi-scale spatio-temporal analysis of human mobility. *PLoS One* 12 (2), e0171686.
- Batty, M., DeSyllas, J., Duxbury, E., 2003. The discrete dynamics of small-scale spatial events: agent-based models of mobility in carnivals and street parades. *Int. J. Geogr. Inf. Sci.* 17 (7), 673–697.
- Baxter, L.K., Dionisio, K.L., Burke, J., Sarnat, S.E., Sarnat, J.A., Hodas, N., Rich, D.Q., Turpin, B.J., Jones, R.R., Mannshardt, E., 2013. Exposure prediction approaches used in air pollution epidemiology studies: key findings and future recommendations. *J. Expo. Sci. Environ. Epidemiol.* 23 (6), 654.
- Beckx, C., Panis, L.L., Arentze, T., Janssens, D., Torfs, R., Broekx, S., Wets, G., 2009. A dynamic activity-based population modelling approach to evaluate exposure to air pollution: methods and application to a Dutch urban area. *Environ. Impact Assess. Rev.* 29 (3), 179–185.
- CBS. Kaart van 100 meter bij 100 meter met statistieken. <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografisch/%20data/kaart-van-100-meter-bij-100-meter-met-statistieken>. (Accessed 7 May 2018).
- Chen, H., Goldberg, M.S., 2009. The effects of outdoor air pollution on chronic illnesses. *McGill J. Med.: MJM* 12 (1), 58.
- Chen, R., Chu, C., Tan, J., Cao, J., Song, W., Xu, X., Jiang, C., Ma, W., Yang, C., Chen, B., 2010. Ambient air pollution and hospital admission in shanghai, China. *J. Hazard Mater.* 181 (1–3), 234–240.
- Contributors, O. Geofabrik. <http://geofabrik.de>. (Accessed 9 October 2018).
- Contributors, O. ModernC++ routing engine for shortest paths in road networks. <https://github.com/Project-OSRM>. last. (Accessed 9 October 2018).
- Contributors, O. Open street route machine (osrm) profiles. <https://github.com/Project-OSRM/osrm-backend/tree/master/profiles>. (Accessed 9 October 2018).
- Darrow, L.A., Klein, M., Sarnat, J.A., Mulholland, J.A., Strickland, M.J., Sarnat, S.E., Russell, A.G., Tolbert, P.E., 2011. The use of alternative pollutant metrics in time-series studies of ambient air pollution and respiratory emergency department visits. *J. Expo. Sci. Environ. Epidemiol.* 21 (1), 10.
- de Bakker, M.P., de Jong, K., Schmitz, O., Karssenber, D., 2017. Design and demonstration of a data model to integrate agent-based and field-based modelling. *Environ. Model. Softw* 89 (Suppl. C), 172–189.
- Deffner, V., Küchenhoff, H., Maier, V., Pitz, M., Cyrys, J., Breitner, S., Schneider, A., Gu, J., Gerschkat, U., Peters, A., 2016. Personal exposure to ultrafine particles: two-level statistical modeling of background exposure and time-activity patterns during three seasons. *J. Expo. Sci. Environ. Epidemiol.* 26 (1), 17.
- Dias, D., Tchepel, O., 2018. Spatial and temporal dynamics in air pollution exposure assessment. *Int. J. Environ. Res. Public Health* 15 (3), 558.
- Dons, E., Panis, L.L., Van Poppel, M., Theunis, J., Willems, H., Torfs, R., Wets, G., 2011. Impact of time-activity patterns on personal exposure to black carbon. *Atmos. Environ.* 45 (21), 3594–3602.
- Elliott, P., Wakefield, J.C., Best, N.G., Briggs, D.J., 2000. *Spatial Epidemiology: Methods and Applications*. Oxford University Press.
- ESCAPE. Escape - european study of cohorts for air pollution effects. <http://www.escapeproject.eu/manuals/>. last. (Accessed 16 May 2019).
- Expatica. Utrecht university: facts & figures. <https://www.expatica.com/nl/education/higher-education/utrecht-university-facts-figures-107021/>. (Accessed 22 August 2019).
- for Public Health, N.I., the Environment. Air quality monitoring network. <https://www.luchtmeetnet.nl/>. (Accessed 9 October 2018).
- Gehring, U., Gruzjeva, O., Agius, R.M., Beelen, R., Custovic, A., Cyrys, J., Eeftens, M., Flexeder, C., Fuertes, E., Heinrich, J., 2013. Air pollution exposure and lung function in children: the ESCAPE project. *Environ. Health Perspect.* 121 (11–12), 1357.
- Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.-L., 2008. Understanding individual human mobility patterns. *Nature* 453 (7196), 779.
- Gulliver, J., Briggs, D.J., 2005. Time-space modeling of journey-time exposure to traffic-related air pollution using GIS. *Environ. Res.* 97.
- Gurram, S., Stuart, A.L., Pinjari, A.R., 2015. Impacts of travel activity and urbanicity on exposures to ambient oxides of nitrogen and on exposure disparities. *Air Qual. Atmos. Health* (1), 97–114.
- Hankey, S., Marshall, J.D., 2015. On-bicycle exposure to particulate air pollution: particle number, black carbon, PM_{2.5}, and particle size. *Atmos. Environ.* 122, 65–73.
- Hertel, O., De Leeuw, F.A., Jensen, S.S., Gee, D., Herbarth, O., Pryor, S., Palmgren, F., Olsen, E., 2001. Human exposure to outdoor air pollution (iupac technical report). *Pure Appl. Chem.* 73 (6), 933–958.
- Hoek, G., Krishnan, R.M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., Kaufman, J.D., 2013. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ. Health* 12 (1), 43.
- Johansson, K.A., Balmes, J.R., Collard, H.R., 2015. Air pollution exposure: a novel environmental risk factor for interstitial lung disease? *Chest* 147 (4), 1161–1167.
- Kadaster. Dataset: basisregistratie adressen en gebouwen (BAG). <https://www.pdok.nl/introductie/-/article/basisregistratie-adres-en-gebouwen-ba-1>. (Accessed 17 September 2019).
- Kan, H., Chen, R., Tong, S., 2012. Ambient air pollution, climate change, and population health in China. *Environ. Int.* 42, 10–19.
- Karssenber, D., Schmitz, O., Salamon, P., de Jong, K., Bierkens, M.F., 2010. A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environ. Model. Softw* 25 (4), 489–502.
- Law, S., Sakr, F.L., Martinez, M., 2014. Measuring the changes in aggregate cycling patterns between 2003 and 2012 from a space syntax perspective. *Behav. Sci.* 4 (3), 278–300.
- Lee, S., Shin, J., Lee, G., Moon, I.-C., 2013. Impact of population relocation to city commerce: micro-level estimation with agent-based model. In: *Proceedings of the Agent-Directed Simulation Symposium*. Society for Computer Simulation International, p. 11.
- Li, H.Z., Gu, P., Ye, Q., Zimmerman, N., Robinson, E.S., Subramanian, R., Apte, J.S., Robinson, A.L., Presto, A.A., 2019. Spatially dense air pollutant sampling: implications of spatial variability on the representativeness of stationary air pollutant monitors. *Atmos. Environ.* X 2, 100012.
- Luengo-Oroz, J., Reis, S., 2019. Assessment of cyclists' exposure to ultrafine particles along alternative commuting routes in edinburgh. *Atmos. Pollut. Res.*
- Meier, R., Eeftens, M., Phuleria, H.C., Ineichen, A., Corradi, E., Davey, M., Fierz, M., Ducret-Stich, R.E., Aguilera, I., Schindler, C., 2015. Differences in indoor versus outdoor concentrations of ultrafine particles, PM_{2.5}, PM_{absorbance} and NO₂ in Swiss homes. *J. Expo. Sci. Environ. Epidemiol.* 25 (5), 499.
- Miller, E., Roorda, M., 2003. Prototype model of household activity-travel scheduling. *Transp. Res. Rec.: J. Transp. Res. Board* (1831), 114–121.
- Miller, H.J., 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *Int. J. Geogr. Inf. Syst.* 5 (3), 287–301.
- Minet, L., Liu, R., Valois, M.-F., Xu, J., Weichenenthal, S., Hatzopoulou, M., 2018. A Comparison of Air Pollution Exposure Surfaces Developed via Mobile and Fixed-Site Monitoring Protocols against Personal Exposures. Technical Report.
- Möller, A., Lindley, S., de Vocht, F., Agius, R., Kerry, G., Johnson, K., Ashmore, M., Terry, A., Dimitroulopoulou, S., Simpson, A., 2012. Performance of a microenvironmental model for estimating personal no₂ exposure in children. *Atmos. Environ.* 51, 225–233.
- Nguyen, A.D., Sénac, P., Ramiro, V., Diaz, M., 2011. Steps - an approach for human mobility modeling. In: *International Conference on Research in Networking*. Springer, pp. 254–265.
- Park, Y.M., Kwan, M.-P., 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health Place* 43, 85–94.
- Rivas, I., Viana, M., Moreno, T., Bouso, L., Pandolfi, M., Alvarez-Pedrerol, M., Forns, J., Alastuey, A., Sunyer, J., Querol, X., 2015. Outdoor infiltration and indoor contribution of UFP and BC, OC, secondary inorganic ions and metals in PM_{2.5} in schools. *Atmos. Environ.* 106, 129–138.
- Schmitz, O., Beelen, R., Strak, M., Hoek, G., Soenario, I., Brunekreef, B., Vaartjes, I., Dijkstra, M.J., Grobbee, D.E., Karssenber, D., 2019. High resolution annual average air pollution concentration maps for The Netherlands. *Sci. Data* 6, 190035.
- Shekarrizfard, M., Faghieh-Imani, A., Tetreault, L.-F., Yasmin, S., Reynaud, F., Morency, P., Plante, C., Drouin, L., Smargiassi, A., Eluru, N., 2017. Regional assessment of exposure to traffic-related air pollution: impacts of individual mobility and transit investment scenarios. *Sustain. Cities Soc.* 29, 68–76.
- Smith, J.D., Mitsakou, C., Kitwiroon, N., Barratt, B.M., Walton, H.A., Taylor, J.G., Anderson, H.R., Kelly, F.J., Bevers, S.D., 2016. London hybrid exposure model: improving human exposure estimates to NO₂ and PM_{2.5} in an urban setting. *Environ. Sci. Technol.* 50 (21), 11760–11768.
- Soenario, I., Helbich, M., Schmitz, O., Hoek, G., van der Molen, M., Lu, M., Karssenber, D., 2019. Land Use Regression Models Revealing Spatio-Temporal Covariation in Combustion Related Air Pollutants in the Netherlands. submitted to *Atmospheric Environment*.
- Steinle, S., Reis, S., Sabel, C.E., 2013. Quantifying human exposure to air pollution—moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Sci. Total Environ.* 443, 184–193.
- Strak, M., Janssen, N., Beelen, R., Schmitz, O., Karssenber, D., Houthuijs, D., van den Brink, C., Dijkstra, M., Brunekreef, B., Hoek, G., 2017. Associations between lifestyle and air pollution exposure: potential for confounding in large administrative data cohorts. *Environ. Res.* 156, 364–373.
- Tang, R., Tian, L., Thach, T.-Q., Tsui, T.H., Brauer, M., Lee, M., Allen, R., Yuchi, W., Lai, P.-C., Wong, P., 2018. Integrating travel behavior with land use regression to estimate dynamic air pollution exposure in Hong Kong. *Environ. Int.* 113, 100–108.
- Team, P., Python software foundation. *python language reference, version 2.7*. <http://www.python.org>. (Accessed 7 May 2018).
- Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* 267–288.
- travel forecasting resource. Albatross. <http://tfresource.org/ALBATROSS>. last. (Accessed 1 June 2018).
- Utrecht, V. Age structure. <https://www.volksgezondheidsmonitor.nl/en/age-structure-utrecht/page6.html>. last. (Accessed 22 August 2019).
- Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., De Baets, B., 2015. Mobile monitoring for mapping spatial variation in urban air quality:

- development and validation of a methodology based on an extensive dataset. *Atmos. Environ.* 105, 148–161.
- WHO. WHO guidelines for indoor air quality: selected pollutants. http://www.euro.who.int/_data/assets/pdf_file/0009/128169/e94535.pdf. (Accessed 8 May 2019).
- Woodcock, J., Edwards, P., Tonne, C., Armstrong, B.G., Ashiru, O., Banister, D., Beevers, S., Chalabi, Z., Chowdhury, Z., Cohen, A., 2009. Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport. *The Lancet* 374 (9705), 1930–1943.
- Yang, L.E., Hoffmann, P., Scheffran, J., Rühle, S., Fischereit, J., Gasser, I., 2018. An agent-based modeling framework for simulating human exposure to environmental stresses in urban areas. *Urban Sci.* 2 (2), 36.
- Yang, S., Yang, X., Zhang, C., Spyrou, E., 2010. Using social network theory for modeling human mobility. *IEEE Netw.* 24 (5).
- Yang, W., Lee, K., Chung, M., 2004. Characterization of indoor air quality using multiple measurements of nitrogen dioxide. *Indoor Air* 14 (2), 105–111.
- Yoo, E., Rudra, C., Glasgow, M., Mu, L., 2015. Geospatial estimation of individual exposure to air pollutants: moving from static monitoring to activity-based dynamic exposure assessment. *Ann. Assoc. Am. Geogr.* 105 (5), 915–926.
- Yu, H., 2006. Spatio-temporal gis design for exploring interactions of human activities. *Cartogr. Geogr. Inf. Sci.* 33 (1), 3–19.
- Zhang, K., Batterman, S., Dion, F., 2011. Vehicle emissions in congestion: comparison of work zone, rush hour and free-flow conditions. *Atmos. Environ.* 45 (11), 1929–1939.
- Zhou, W., Yuan, D., Ye, S., Qi, P., Fu, C., Christiani, D.C., 2001. Health effects of occupational exposures to vehicle emissions in Shanghai. *Int. J. Occup. Environ. Health* 7 (1), 23–30.
- Zou, B., Wilson, J.G., Zhan, F.B., Zeng, Y., 2009. Air pollution exposure assessment methods utilized in epidemiological studies. *J. Environ. Monit.* 11 (3), 475–490.