Environment International 62 (2014) 64–71

Contents lists available at ScienceDirect

Environment International

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

Implementation and validation of a modeling framework to assess personal exposure to black carbon

Evi Dons [a,b,*], Martine Van Poppel [a], Bruno Kochan [b], Geert Wets [b], Luc Int Panis [a,b]

[a] VITO (Flemish Institute for Technological Research), Boeretang 200, 2400 Mol, Belgium
[b] IMOB (Transportation Research Institute), Hasselt University, Wetenschapspark 5 bus 6, 3590 Diepenbeek, Belgium

ARTICLE INFO

Article history:
Received 14 June 2013
Accepted 5 October 2013
Available online 25 October 2013

Keywords:
Personal exposure
Activity-based model
Land use regression
Black carbon
Microenvironments
Traffic

ABSTRACT

Because people tend to move from one place to another during the day, their exposure to air pollution will be determined by the concentration at each location combined with the exposure encountered in transport. In order to estimate the exposure of individuals in a population more accurately, the activity-based modeling framework for Black Carbon exposure assessment, AB^{2}C, was developed. An activity-based traffic model was applied to model the whereabouts of individual agents. Exposure to black carbon (BC) in different microenvironments is assessed with a land use regression model, combined with a fixed indoor/outdoor factor for exposure in indoor environments. To estimate exposure in transport, a separate model was used taking into account transport mode, timing of the trip and degree of urbanization.

The modeling framework is validated using weeklong time–activity diaries and BC exposure as revealed from a personal monitoring campaign with 62 participants. For each participant in the monitoring campaign, a synthetic population of 100 model-agents per day was made up with all agents meeting similar preconditions as each real-life agent. When these model-agents pass through every stage of the modeling framework, it results in a distribution of potential exposures for each individual. The AB^{2}C model estimates average personal exposure slightly more accurately compared to ambient concentrations as predicted for the home subzone; however the added value of a dynamic model lies in the potential for detecting short term peak exposures rather than modeling average exposures. The latter may bring new opportunities to epidemiologists: studying the effect of frequently repeated but short exposure peaks on long term exposure and health.

1. Introduction

High-resolution personal measurements are the best way of getting information on the exposure of individuals; which combined with geo-information enables exposure to be assessed in space and time. Unfortunately personal measurements are expensive and a burden to individuals; therefore most current studies are limited to short periods of personal measurements in a small sample. These snapshots of exposure are not sufficient to provide epidemiologists with data on long term exposure of large cohorts. Building exposure models is an obvious alternative, but this often results in simple models not capable of estimating exposure with enough reliability.

Ideally personal exposure to air pollution should be modeled as a combination of two interacting geographies: a moving population and a continuously changing air quality (Briggs, 2005). Many studies only take into account residential location and ignore time–activity patterns while air pollution is typically modeled using annual average concentrations without any daily variation. Several more complex exposure models have been developed: e.g. SHEDS (Stochastic Human Exposure and Dose Simulation Model, (Burke et al., 2001)), pCNEM (Probabilistic Version of the NAAQS Exposure Model, (Zidek et al., 2005)), STEMS (Space–Time Exposure Modeling System, (Gulliver and Briggs, 2005)), EMI (Exposure Model for Individuals, (Breen et al., 2010)), MEEM (MicroEnvironmental Exposure Model, (Mölter et al., 2012)). The first two models estimate population exposure; the other models assess exposure of individuals. Models estimating exposure of individuals are up till now not designed to model exposure of a full population as specific information on individuals, not available on an aggregated level, is required. Population exposure models generate diaries from activity databases, often without a geographical dimension, and often deterministic in nature. Beckx et al. (2009) and Hatzopoulou et al. (2011) were the first to use an activity-based transportation model to predict stochastic time–location–activity diaries and use them in an air pollution exposure model. Lefebvre et al. (2013) and Dhondt et al. (2012) used activity-based models to evaluate the effect of air quality measures on population exposure and health. An important issue is that the final outcome of the aforementioned population exposure models has never been fully validated; the validation was limited to the validation of submodels. Personal monitoring of a random sample of n individuals from a target

* Corresponding author at: VITO (Flemish Institute for Technological Research), Boeretang 200, 2400 Mol, Belgium. Tel.: +32 14 33 51 90; fax: +32 14 58 26 57.
E-mail address: evi.dons@vito.be (E. Dons).

0160-4120/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envint.2013.10.003

population is a valuable tool to examine the validity of human exposure models (Duan, 1991; Gerharz et al., 2013; Mölter et al., 2012). Models for individual exposure assessment expect diaries or GPS tracks to be known: validating these kinds of models in fact means validating only the air quality models, and not the diaries themselves. In the case of population exposure models, the space–time predictions need to be validated as well, together with the air quality models.

In this study, a personal exposure modeling framework using an activity-based model will be implemented and its performance will be compared to real-life personal exposures. Most submodels have already been presented elsewhere (Bellemans et al., 2010; Dons et al., 2013a, 2013b), but in this study all the models are integrated in one framework for the first time: the AB^{2}C model (An Activity-Based modeling framework for Black Carbon exposure assessment). Personal measurements are compared against a distribution of personal exposure estimates and serve as a validation for the complete framework. The AB^{2}C model is developed to estimate population-wide exposure to black carbon (BC), a pollutant suspected of having health effects (WHO, 2012; Zanobetti and Schwartz, 2006). Introducing individual mobility in population exposure models for BC is highly relevant because moving from one place to another can significantly alter exposure through the steep concentration gradients observed near BC sources. Individual exposure assessment on a population level, as aimed for by the AB^{2}C model, can contribute to health policy: highly exposed individuals or groups of individuals can be identified; the impact of age, gender or socio-economic class on exposure can be verified or the number of individuals exposed to levels above a certain threshold can be determined. Moreover policy scenarios can be calculated and the effect on personal and population exposure to BC can be assessed, and ultimately extended with appropriate exposure–response functions to perform health impact assessments.

2. Materials and methods

The modeling framework AB^{2}C consists of an activity-based model, hourly land use regression models, an in-traffic personal exposure model, and an indoor air model. Separate models are described first; how models fit in a consistent framework is discussed afterwards. Real-life measurements serve as a validation of the entire model chain.

2.1. Activity-based model

Activity-based models simulate activities and trips for a synthetic population based on thousands of revealed diaries (Arentze and Timmermans, 2004; Davidson et al., 2007). Trips are derived from the activities performed and their physical location: if subsequent activities are in a different geographical zone, a trip between the origin and the destination is required. In Flanders, FEATHERS (Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS) was developed as a simulation platform to implement activity-based models (Bellemans et al., 2010; Kochan et al., 2013). For every adult in the population, approximately 6 million people, a diary is built using a series of decision trees and constraints. It results in a model predicting activities (home-based activities, work/education, business, bring/get, shopping, services, social, leisure, touring, other), their timing, duration and location for 24 h for every agent in a synthetic population. The study area is divided into 2386 subzones, also called Traffic Analysis Zones (TAZ; terminology used in transportation planning models), with an average size of 5.7 km^2; the size of a subzone is related to the number of inhabitants. The subzone where an activity will be performed is determined in FEATHERS by taking into account land use data such as the number of work places and the number of inhabitants in each TAZ to estimate the level of production and attraction. Agents have a preference of performing an activity closer to home for shopping trips, but the distance can be larger for work locations, etc. For all 168 h of an average week, a dynamic population density is calculated, not based on static addresses, but based on the locations that agents actually visit during each hour. For trips, the transport mode (car driver, car passenger, bike/walk, public transport), the duration of the trip, and subzone of origin and destination will be predicted by FEATHERS. Those trips are then assigned to a road network using an equilibrium assignment, thus taking into account congestion effects, in TransCAD software (Caliper, 2013). Outputs from the FEATHERS model used in subsequent steps are: hourly traffic flows for motorized trips, hourly dynamic population densities taking into account population mobility, and 24 h-diaries for a subset of agents with specific characteristics (see Supplemental Table SI1 for an example output of a modeled diary).

2.2. Hourly LUR models

In a land use regression (LUR) model statistical associations are developed between potential predictor variables and measured pollutant concentrations as a basis for predicting concentrations at unsampled sites (Hoek et al., 2008; Ryan and LeMasters, 2007). In 2010 and 2011, BC was measured on 63 locations in Flanders with a high temporal resolution (5-min) using micro-aethalometers (AethLabs, 2012; Dons et al., 2013b). Measurement sites were carefully selected to have enough variation in expected concentrations (13 street sites, 25 urban traffic sites, 11 urban background sites, 14 rural sites) and to geographically cover the whole study area. On most locations measurements were repeated in a contrasting season, and raw measurements were adjusted (using one continuous monitor) to be representative of annual average concentrations. Mean annual concentrations ranged from 846 ng/m^3 to 4184 ng/m^3, with hourly concentration peaks of nearly 10,000 ng/m^3. Hourly LUR models were derived for weekdays (Mon–Fri, 24 models) and for the weekend (Sat–Sun, 24 models) to capture intraday variation in the spatial concentration pattern of the study area (Dons et al., 2013b). Seasonal variability is not accounted for as this variability is generally driven by meteorological conditions and meteorological variables are not collected, whereas diurnal variability is principally derived from human activities. The hourly models were estimated independent of each other according to a fixed model development algorithm (Eeftens et al., 2012; Henderson et al., 2007). Weekday hourly models performed well during the day and on traffic peak hours, explaining 60 to 80% of variability using mainly traffic variables (Dons et al., 2013b). At night and in the weekend the models were less predictive using only 1 or 2 predictors (mainly land use variables in larger buffers), but RMSE values were low indicative of homogeneous BC concentrations (Dons et al., 2013b). Traffic and population variables from the activity-based model were only sporadically included, e.g. traffic intensity on the nearest road was significant only on traffic peak hours. Although hourly models were developed independently of adjacent hours, similar variables do often return in consecutive models demonstrating the robustness of the models.

In this study, concentrations were calculated for 10 randomly chosen buildings in each subzone using the hourly LUR models; the median concentration is the exposure assigned to people present in that subzone during that hour (10 buildings proved to be sufficient to reliably estimate median concentrations, see supplemental material for details on the methodology). Where estimates exceeded the maximum measured concentrations by more than 20%, concentration estimates were capped (Henderson et al., 2007).

2.3. In-traffic personal exposure model

While exposure in geographical places can be modeled using hourly LUR models, simple LUR-like regression models were also developed to