Atmospheric Environment 77 (2013) 598–606

Contents lists available at SciVerse ScienceDirect

Atmospheric Environment

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

Comparison of performance of land use regression models derived for Catalunya, Spain

Audrey de Nazelle [a,b,c,d,*], Inmaculada Aguilera [b,c,d], Mark Nieuwenhuijsen [b,c,d], Rob Beelen [e], Marta Cirach [b,c,d], Gerard Hoek [e], Kees de Hoogh [f], Jordi Sunyer [b,c,d,g], Jaume Targa [h], Bert Brunekreef [e,i], Nino Künzli [j,k], Xavier Basagaña [b,c,d]

[a] Centre for Environmental Policy, Imperial College London, London, United Kingdom
[b] Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
[c] IMIM (Hospital del Mar Research Institute), Barcelona, Spain
[d] CIBER Epidemiología y Salud Pública (CIBERESP), Spain
[e] Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands
[f] MRC-HPA Centre for Environment and Health, Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom
[g] Pompeu Fabra University, Barcelona, Spain
[h] 4sfera Innova, Girona, Spain
[i] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
[j] Swiss Tropical and Public Health Institute, Basel, Switzerland
[k] University of Basel, Basel, Switzerland

HIGHLIGHTS

- We validate three overlapping land use regression models with external datasets.
- Despite differences in sampling protocols validation R^2s are reasonably high.
- Predictions from different models at cohort addresses are relatively well correlated.

ARTICLE INFO

Article history:
Received 16 January 2013
Received in revised form
20 May 2013
Accepted 21 May 2013

Keywords:
Land use regression
NO_2
Performance
Residential exposure

ABSTRACT

Land use regression (LUR) models have become a popular tool to capture small-scale variations in air pollution exposures in epidemiological analyses, and have been developed with a variety of approaches with no clear indication of the most efficient and appropriate one. We evaluated the performance of the LUR model for NO_2 developed for the European Study of Cohorts for Air Pollution Effects (ESCAPE) for Catalunya, Spain, compared to two other LUR models derived locally and independently for two cohort studies (INMA-Sabadell and REGICOR-Girona) in different sub-areas of the ESCAPE domain. We made use of sampling campaigns from the three studies as independent sets of measurements by which to evaluate each model. We compared changes in R^2 and measures of bias when applying each model to its own dataset vs. the independent datasets from the other studies. The three studies differed principally in their scale of analysis (from urban area only for INMA-Sabadell to large province covering urban and rural areas for ESCAPE-Catalunya and REGICOR-Girona) and sampling protocol (e.g. site selection).

The LUR models performed similarly well in terms of their model adjusted R^2 and cross-validation R^2, ranging respectively from 0.62 and 0.63 to 0.75 and 0.73. The ESCAPE model performed well at the ESCAPE sites in Sabadell ($R^2 = 0.69$) and moderately well at the ESCAPE sites in Girona province ($R^2 = 0.53$). The ESCAPE model predicted the external sites less well: R^2 were 0.51 and 0.36 in Sabadell and Girona province. The INMA-Sabadell and REGICOR-Girona models showed a similar pattern: the R^2 for the INMA model dropped from 0.69 to 0.50 at INMA versus ESCAPE sites in Sabadell, while the R^2 for the REGICOR model dropped from 0.63 to 0.44 for REGICOR versus ESCAPE sites in Girona province. The drop in performance for external sites is likely a combination of overfitting and differences in the sampling campaigns (years, site selection). Agreement between models was 53%–74% for the

* Corresponding author. Centre for Environmental Policy, Imperial College London, London, United Kingdom.
E-mail addresses: anazelle@gmail.com, anazelle@imperial.ac.uk (A. de Nazelle).

1352-2310/$ – see front matter
http://dx.doi.org/10.1016/j.atmosenv.2013.05.054

classification of low, medium, and high levels of air pollution predicted at cohort addresses. Despite the drop in performance, the three models still explained a substantial fraction of the variation at independent sites, especially in Sabadell, supporting their use in epidemiological studies.

1. Introduction

Near-road traffic related air pollution has been associated in many epidemiologic studies with a multitude of health effects ranging from reproductive outcomes to all-cause mortality (HEI, 2010). While these studies indicate that living close to busy roads is associated with adverse health effects, findings are not easily translated into policies and generalization from one city to another is limited as "proximity" does not necessarily reflect the same types and levels of pollution. Thus, alternative objective measures of local traffic-related pollution are warranted. Land Use Regression (LUR) modelling is often the chosen exposure assessment methodology to capture small-scale differences in air pollution concentrations, particularly from traffic sources, with medium implementation costs (Jerrett et al., 2005). Numerous epidemiological studies use LUR models to investigate the health effects of air pollution. Nitrogen dioxide (NO_2) has most frequently been used as marker of near-road traffic related pollutants. A landmark activity of the ESCAPE project (European Study of Cohorts for Air Pollution Effects) is the adoption of uniform air pollution exposure assessment methodologies to assess the spatial variability of traffic-related air pollutants (Beelen et al., 2013; Cyrys et al., 2012). The same methods of measurements and modelling were applied to 36 study areas across Europe in sites where pre-existing cohort studies were available to provide health outcome data. NO_2 was one of the pollutants chosen as an indicator of traffic-related exposures to develop LUR estimates to then assign exposures to the participants of local studies.

A major challenge in the ESCAPE LUR development was the selection of the feasible number of measurement sites for a large number of European study areas within a limited budget. As previously shown, the number of measurement sites used to develop LUR models also impacts the performance of those models (Basagana et al., 2012; Wang et al., 2013). These studies have shown that for models developed on a small number of sites, the model and cross-validation R^2 overestimate predictive ability at independent test sites. This highlights the importance of validating models with independent datasets. A second – and partly related – challenge was the ambitious spatial coverage demanded all across Europe for regions of heterogeneous sizes, namely varying from intra-city to large regional domains. One example of a single large but geographically diverse region is Catalunya, Spain, where ESCAPE provided a LUR model for several cohort studies from different towns and cities. For two cohorts – namely REGICOR (Girona Heart Registry) in the region of Girona (Rivera et al., 2012), and INMA (Environment and Childhood) in the city of Sabadell (Aguilera et al., 2008) – independent exposure assessment studies were undertaken to build NO_2 LUR models for both study areas prior to the start of ESCAPE. The existence of parallel measurement campaigns and LUR models offers an interesting opportunity to compare the performance of the regional ESCAPE LUR model for NO_2 with LUR models derived locally for each of these cohorts based on a higher density of measurement sites.

Comparisons between LUR models developed for the same area are scant. Dijkema et al. (2011) compared two LUR models developed at different scales (large area and city-specific, encompassing the same core-area of Amsterdam) and using different monitoring campaigns. They found in both cases a drop in model performance in terms of adjusted R^2 when applying the model to the other model's monitoring sites, and highlighted in their conclusions the importance of a sampling location strategy purposefully designed to reflect locations where models are to be applied.

The first objective of our study was to study the performance of the LUR model by comparing predictions with observed values at locations that were not used for model development.

Because the ultimate goal of LUR modelling in epidemiology is to assign estimates of air pollution exposure to participants of health studies, our second objective was to compare the different model predictions at the residential addresses of the two local cohort studies.

2. Methods

Three distinct LUR models were developed within the region of Catalunya for three different epidemiological studies: ESCAPE (Cyrys et al., 2012), INMA-Sabadell (Aguilera et al., 2008), and REGICOR-Girona (Rivera et al., 2012). The model domain and the methodology and concepts of model development varied for each of the study models. As shown in Fig. 1, the ESCAPE domain encompasses the domain of the other two studies, as it was developed to assess exposures for participants of three European cohorts, including INMA-Sabadell and REGICOR-Girona. Basic model development steps were common in all three cases, i.e. air pollution was monitored at a variety of sites across the model domain, and resulting concentrations were regressed against land use, traffic, and other geographic variables characterizing the measurement sites. The three studies differed however in the monitoring protocols and modelling procedures, as detailed below.

2.1. Study area

The Catalan ESCAPE LUR model domain (Fig. 1) was 513 km^2, including the INMA study area of Sabadell (38 km^2), the REGICOR study area of Girona province (375 km^2, including Girona city and 11 towns), and the city of Barcelona (100 km^2). A wide variety of topographical and meteorological conditions are present, especially in the Girona province with altitude ranging from 1 to 500 m, with towns on the coast and inland, and land uses ranging from urban to rural, while Sabadell is mostly urban with altitude ranging from 100 to 350 m.

2.2. Monitoring campaigns

Table 1 provides a summary of the three NO_2 monitoring campaigns for each study. The major differences lay in the site selection procedures, number of monitoring sites and monitoring equipment. All three studies used passive samplers, but each used a different specific instrument (Ogawa, Radiello, Palmes). Both ESCAPE and INMA-Sabadell campaigns followed a site selection protocol which aimed at representing the gradient of various land use and traffic characteristics, with additional requirements on proximity to traffic intersections and other emission sources for ESCAPE. The goal was to assess residential exposures in all three studies. In ESCAPE, monitors were placed at the façade of homes on