Journal of Exposure Science & Environmental Epidemiology (2019) 29:258–266
https://doi.org/10.1038/s41370-018-0045-x

ARTICLE

Error in air pollution exposure model determinants and bias in health estimates

Jelle Vlaanderen[1] · Lützen Portengen[1] · Marc Chadeau-Hyam[2] · Adam Szpiro[3] · Ulrike Gehring[1] · Bert Brunekreef[1] · Gerard Hoek[1] · Roel Vermeulen[1]

Received: 31 January 2018 / Revised: 26 March 2018 / Accepted: 8 April 2018 / Published online: 8 June 2018

Abstract

Background Land use regression (LUR) models are commonly used in environmental epidemiology to assign spatially resolved estimates of air pollution to study participants. In this setting, estimated LUR model parameters are assumed to be transportable to a main study (the ''transportability assumption''). We provide an empirical illustration of how violation of this assumption can affect exposure predictions and bias health-effect estimates.

Methods We based our simulation on two existing LUR models, one for nitrogen dioxide, the other for particulate matter with aerodynamic diameter <2.5 μm. We assessed the impact of error in exposure determinants used in the LUR models on resultant air pollution predictions and on bias in an exposure-health-effect estimate assessed in a hypothetical cohort. We assigned error to predictors at monitoring sites (sites used to develop the LUR model) and at prediction sites (sites for which exposure predictions were needed), allowing for different error levels between site types.

Results Realistic error in the exposure determinants of the selected LUR models did not induce large additional error in exposure predictions and resulted in only minor (<1%) bias in health-effect estimates. Bias in the health-effect estimates strongly increased (up to 13.6%) when exposure determinant errors were different for monitoring sites than for prediction sites.

Conclusions These results suggest that only modest reductions in bias in estimated exposure health-effects are to be expected from reducing error in exposure determinants. It is important to avoid heterogeneous errors in exposure determinants between monitoring sites and prediction sites to satisfy the transportability assumption and avoid bias in estimated exposure health-effects.

Keywords Exposure modeling · Epidemiology · Empirical/statistical models

Introduction

Land use regression (LUR) models are commonly used in environmental epidemiology to assign spatially resolved estimates of air pollution markers to study participants [1]. For example, in the pan-European ESCAPE project of ±30 cohort studies including approximately 900,000 subjects, LUR-derived exposure estimates [2, 3] were shown to be associated with a variety of health outcomes such as birth weight, cardiovascular and respiratory disease, cancer incidence, and mortality [4–11].

Exposure prediction using LUR models proceeds in two steps. First, LUR models are estimated by regressing exposure measurement data collected at a limited number of sites (monitoring sites) on a set of geographical exposure determinants [2, 3]. Second, these LUR models are used to predict exposure levels for study subjects using the same set

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s41370-018-0045-x) contains supplementary material, which is available to authorized users.

✉ Jelle Vlaanderen
j.j.vlaanderen@uu.nl

[1] Division of Environmental Epidemiology & Veterinary Public Health, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

[2] MRC-PHE Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

[3] Department of Biostatistics, University of Washington, Seattle, WA, USA

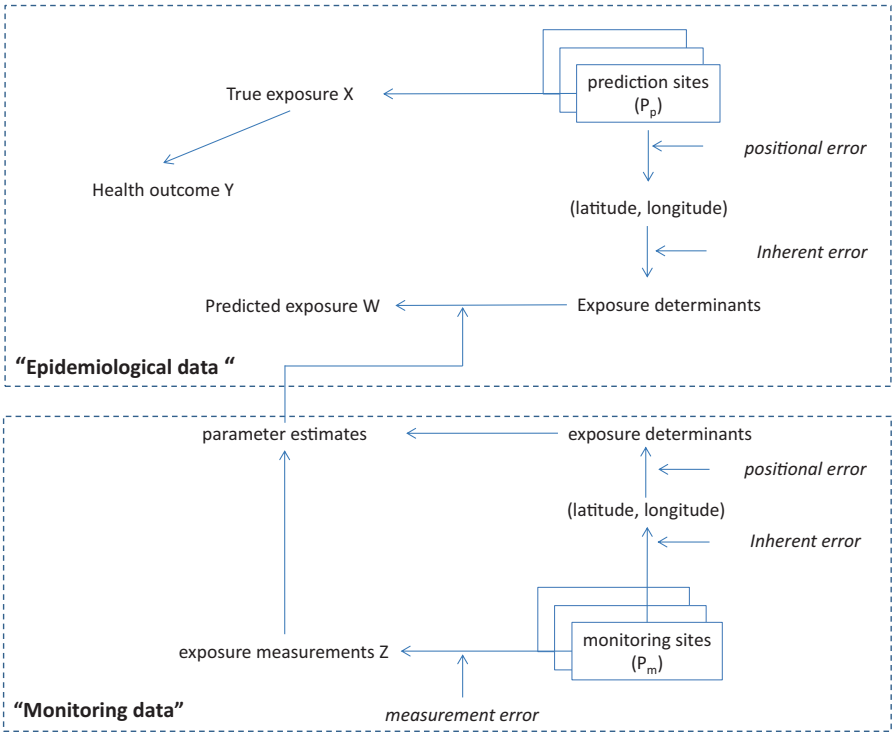

Fig. 1 Illustration of a two-stage analysis and aspects contributing to error in land use regression model predictions. Inherent error in exposure determinants occurs at both monitoring (Pm) and prediction (Pp) sites due to error in the data sources that were used to derive these exposure determinants. Positional error is the error resulting from error in the exact latitude and longitude of monitoring and prediction sites. Exposure measurements Z (affected by measurement error) are conducted at the monitoring sites and are used to derive parameter coefficients for the exposure determinants to allow prediction of exposure estimates W at the prediction sites. If true exposure X explains x percent of the variability in health outcome Y, the goal of the two stage analysis is for health effect β to approach x

of exposure determinants collected using the geographical location of their whereabouts (prediction sites, e.g., residence(s), work location). These predicted exposures are then used in an epidemiological analysis to assess the association between this exposure and a health outcome of interest (See Fig. 1 for a graphical representation).

Usually, the locations of monitoring sites do not match the locations of the prediction sites, a situation known as spatial misalignment [12]. Spatial misalignment can induce error in the exposure estimates which in turn has an impact on the estimation of the exposure-health association.

How much error is introduced depends on the degree to which the exposure measurement error model is transportable, i.e., to which extent the measurement error model at the monitoring sites also holds true at the prediction sites [13].

Error in LUR model exposure determinants can be the result of inherent error in the source information that is used to generate a certain exposure determinant (e.g., in the database that is used to estimate the number of vehicles on the road), but can also be due to error in the measured geographical location of monitoring and prediction sites (See Fig. 1 for an illustration of where these errors occur).

Following standard "errors-in-variables" theory [14], classical error in the exposure determinants at monitoring sites will bias the LUR model parameter estimates. However, there is generally little intrinsic interest in the LUR model parameters, but only in the impact of these errors on the health-effect estimates that result from using a biased LUR model (i.e., a model with biased parameter estimates).

Here, using a Monte Carlo approach, we illustrate the impact of realistic error in the exposure determinants of the Netherlands/Belgium LUR models for nitrogen dioxide (NO_2) and for particulate matter with aerodynamic diameter <2.5 μm ($PM_{2.5}$) that were developed within the ESCAPE project. We illustrate the extent to which these errors are translated into error in predicted exposures at an external set of prediction sites and into subsequent bias of exposure-health-effect parameter estimates. We show the importance of balancing the error in exposure determinants at prediction sites with the error in exposure determinants at monitoring sites, the only situation that yields essentially unbiased health-effect estimates. We compare the impact of error in LUR model exposure determinants in our simulation to other sources that have been reported to contribute to measurement error in LUR model estimates.

Methods

We conducted a Monte Carlo simulation to illustrate how realistic error in exposure determinants of a LUR model could induce error in predicted exposures and bias exposure-health-effect estimates. We based our simulation on two LUR models, for NO_2 and $PM_{2.5}$, that were developed for the Netherlands and Belgium within the ESCAPE project

Table 1 Exposure determinants included in ESCAPE LUR models for the Netherlands and Belgium

	NO_2 model		$PM_{2.5}$ model	
Exposure determinant	β[a] (s.e)	ΔR^2[b]	β[a](s.e)	ΔR^2[b]
Intercept	−7.80 (4.18)		9.46 (2.65)	
Inverse distance to the nearest road (DTR)	$1.22*10^{1}$ (4.59)	1.3%		
Regional estimate (REG)[c]	1.18 ($2.02*10^{-1}$)	6.4%	$4.25*10^{-1}$ ($1.66*10^{-1}$)	6.1%
Population density[d] in a 5000 m buffer (PD)	$2.30*10^{-5}$ ($6.32*10^{-6}$)	2.5%		
Traffic load[e] in a 50 m buffer (TL50)	$2.47*10^{-6}$ ($6.65*10^{-7}$)	2.6%		
Road length of all roads in a 1000 m buffer (RL1000)	$1.06*10^{-4}$ ($3.30*10^{-5}$)	2.0%		
Heavy traffic load[e] in a 25 m buffer (HTL25)	$9.84*10^{-5}$ ($3.78*10^{-5}$)	1.3%		
Heavy traffic load[e] in a 25 m to 500 m buffer (HTL500)	$4.47*10^{-7}$ ($1.92*10^{-7}$)	1.0%		
Road length of all major roads in a 50 m buffer (MRL50)			$1.37*10^{-2}$ ($2.35*10^{-3}$)	3.1%
Traffic load[e] on major roads in a 1000 m buffer (TML1000)			$2.28*10^{-9}$ ($1.04*10^{-9}$)	4.5%
R^2	86.5%		66.6%	

[a]Coefficient (and standard error) of the exposure determinant included in the original model

[b]Change in R^2 due to removal of the exposure determinant from the regression model (does not add up to total R^2 due to correlation between the exposure determinants)

[c]Inverse distance weighted regional background concentration based on ten regional background sites

[d]Number of inhabitants in a buffer

[e]Sum of the length of a road segment$_*$the traffic intensity on that road segment for all road segments. in a buffer. Calculated for total traffic and heavy traffic, and for all roads and major

[2, 3]. The models were developed using measurement data from 80 (NO_2), and 40 ($PM_{2.5}$) monitoring sites classified as ''regional background'' (RB), ''urban background'' (UB), or ''street'' (S) [15, 16]. For each monitoring site three two-week-long measurements were available distributed over the seasons and corrected for temporal variation with the use of a reference site [15, 16]. A number of geographical exposure determinants were used to predict the spatial variation in NO_2 and $PM_{2.5}$ concentrations at the monitoring sites (Table 1). For each exposure determinant included in the LUR models we estimated the error in the exposure determinant that was introduced by error in the source data (inherent error). Furthermore, we estimated the error in each of the exposure determinants resulting from error in the geographical location of monitoring and prediction sites (positional error).

Inherent error in model exposure determinants

Full details of our approach to derive realistic error estimates for all exposure determinants are provided in the Supplemental Material. Briefly, error in the regional contribution of NO_2 or $PM_{2.5}$ at each site (REG) was estimated using the variation in measurement data at regional background sites across three sampling campaigns.. Error in the population density (PD) and distance to the nearest road (DTR) was estimated by calculating the standard deviation of the difference in estimates for each site from two independent data sources. Estimates for error in traffic load exposure determinants–traffic load in a 50 m buffer (TL50), heavy traffic load in a 25 m buffer (HTL25), heavy traffic load in a 25 m to 500 m buffer (HTL500), and traffic load on major roads in a 1000 m buffer (TML1000)–were taken from a study that reported relative standard deviations (RSDs) per road type for estimates of vehicle miles traveled (a quantity very similar to buffer variables used in the ESCAPE LUR models) from the U.S.A: 7.8% for street sites and urban background sites and 3.0% for regional background sites [17]. We assumed that very precise source data was available for exposure determinant road length (RL) and had no data available to reject this assumption. This exposure determinant was, therefore, not included in the Monte Carlo simulation.

Error in the geographical location of monitoring and prediction sites

Geographical location for prediction sites in the Netherlands and Belgium were geocoded In the ESCAPE study using individual building matching techniques based upon cadastral data. The accuracy of cadastral data is high. For example, in the Address Coordinates Netherlands database 93.5% of all coordinates are located at the centroid of the correct building, 6.0% located at the centroid of the correct parcel, and only 0.5% not located in the correct building or parcel [18]. For the simulations we assume that the average geographical location error in residential addresses in the Netherlands and Belgium was 4 m, based on the average

width and depth of a home in NL/BE [19]. We contrast this situation with that where the average error is 30 m, which has been reported for settings where sub-optimal spatial interpolation techniques have been used [20].

Monte Carlo simulation

The design of the simulation is described in detail in the Supplemental Material. Linking GIS data to the geographical coordinates is a time-consuming operation, and we, therefore, restricted the simulated study size to include 200 prediction sites (randomly sampled from residential locations of participants in the PIAMA birth cohort [21]), and 68 (NO_2 model), respectively, 34 ($PM_{2.5}$ model) monitoring sites. To avoid introducing unnecessary additional Monte Carlo error, we used a fully deterministic model for the outcome Y (i.e., the sampling variance σ_y^2 was set to 0).

For the true parameters of the exposure-health model we used an arbitrary intercept (β_0) of 1.8, and a slope (β_1) of −5. This slope resembles the percentage decrease in FEV1 in children with a roughly 20 μg/m^3 increase in NO_2 or a 5 μg/m^3 increase in $PM_{2.5}$ in the PIAMA study [22].

For each simulation, we sampled exposure determinant data contaminated with non-differential measurement error from a multivariate normal distribution based on mean zero (all variables were standardized) and the covariance matrix of the original exposure determinants at monitoring and prediction sites. This matrix was then used in all subsequent simulations to draw multivariate normal errors. We calculated the average β_1 across the simulations. To improve coverage we estimated 95% confidence intervals using a parametric bootstrap.

As a sensitivity analysis, we sampled exposure determinant data setting all off-diagonal values of the covariance matrix to zero (removing covariance between the exposure determinants). Similarly, we conducted analyses using the original covariance matrix of the error in the exposure determinants, and by setting all off-diagonal values to zero (removing covariance between the errors in the exposure determinants).

REG was included in the ESCAPE LUR models to capture regional variability in $PM_{2.5}$ and NO_2. REG is unlikely to change as a result of relatively small positional errors in the estimated geographical location of the monitoring and prediction sites, and we, therefore, only assessed the impact of error in the source information for this determinant.

Within the ESCAPE project all exposure determinants at the prediction sites were constrained to be within the range of values that were observed at the monitoring sites. We applied the same approach in our simulation, but assessed the sensitivity of our results to removing this constraint.

Code availability

Computing code required to replicate the results reported in this submission are available upon request from the corresponding author.

Results

Table 1 presents an overview of the exposure determinants that were included in the Netherlands/Belgium $PM_{2.5}$ and NO_2 LUR models developed within the ESCAPE project. We report the parameter estimates and the change in R^2 due to removal of the exposure determinant from the regression model (ΔR^2). The overall R^2 was 86.5% for the NO_2 model, and 66.6% for the $PM_{2.5}$ model, which is higher than the summed ΔR^2 for each individual exposure determinant because of correlations between exposure determinants. In Supplementary Tables S1 and S2 we report the error variance that we assigned to exposure determinants in the LUR models and in Supplementary Figs. S1–S8 we show the simulated error distribution of the exposure determinants at the monitoring sites.

Exposure simulation

Boxplots in Fig. 2 show the impact on exposure predictions of either inherent error in the source data of the geographical exposure determinants (either adding error to all model exposure determinants simultaneously or adding error to one of the exposure determinants while retaining the original values for the remaining exposure determinants) and of positional error introduced in all exposure determinants simultaneously as a result of error in the recorded geographical location of prediction sites (using either a 4 or 30 m error radius).

When applied simultaneously, realistic error in the exposure determinants of the NO_2 model resulted in median RSDs of the predicted NO_2 concentrations of 6.6%, 4.1%, and 3.2% at regional background (RB), urban background (UB), and street (S) locations, respectively. Error in NO_2 predictions due to inherent error in exposure determinant sources was primarily due to error in DTR, PD, REG, while error in TL50, HTL25, and HTL500 had only minor impact. An error radius of 4 m in the geographical location of the monitoring sites resulted in median RSDs of 2.0%, 1.8%, and 1.8%, at RB, UB, and S sites, respectively, while a 30 m error radius resulted in much larger mean RSDs of 9.5, 5.8, and 10.2%.

Combined error in exposure determinant sources resulted in median RSDs of predicted $PM_{2.5}$ concentrations of 1.5%, 1.4%, and 1.2% at RB, UB, and S locations, respectively. The error in these predictions was primarily the result of

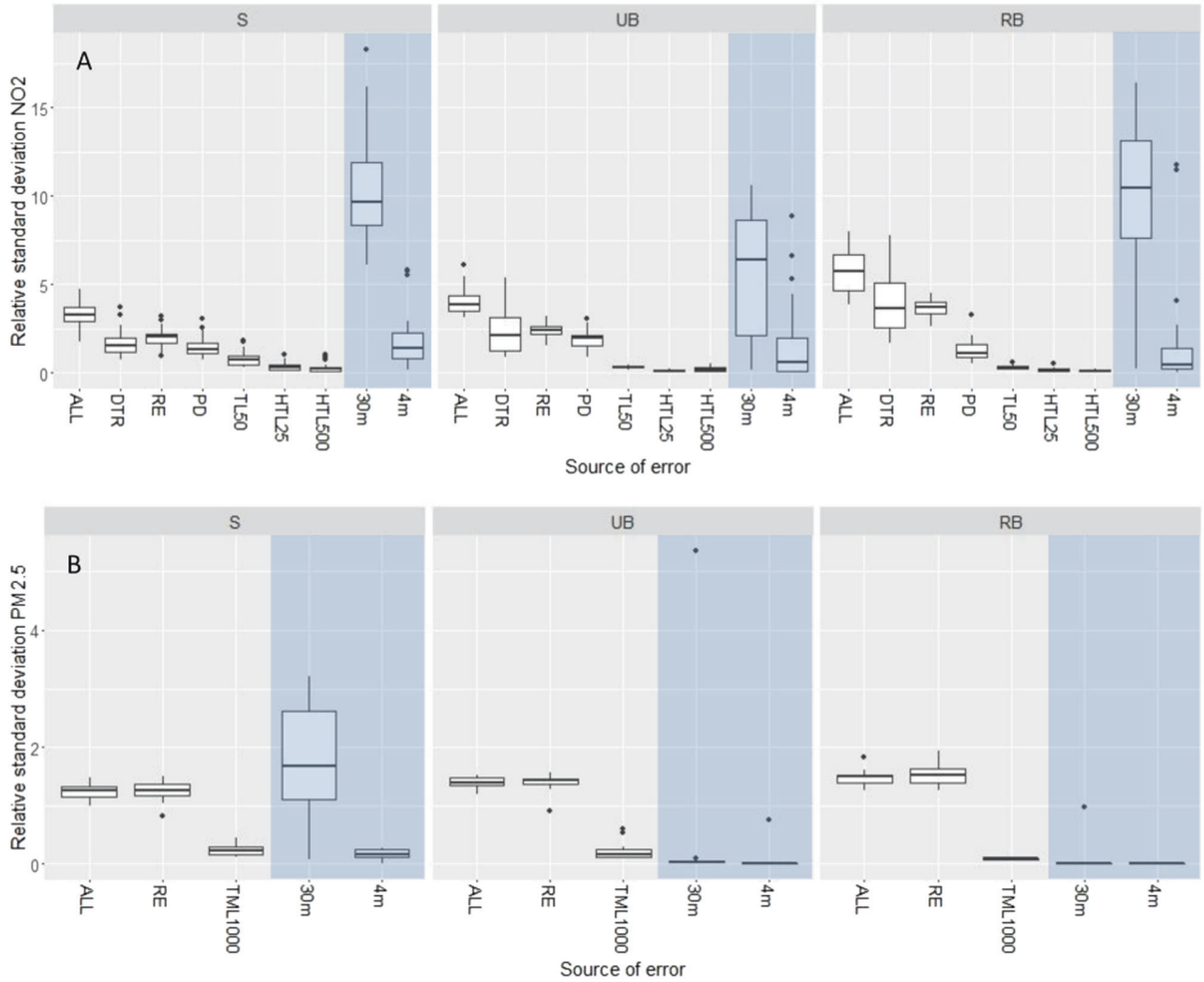

Fig. 2 Impact of realistic error in exposure determinants on NO_2 (plot A) and $PM_{2.5}$ (plot B) exposure levels predicted by the ESCAPE NL/BE LUR model. Relative standard deviation (expressed in percentage and obtained by multiplying the standard deviation by 100 and dividing this product by the average) of the exposure predictions is calculated across 200 simulations for 68 sites for NO_2 and 34 sites for $PM_{2.5}$. Boxplots are ordered by site type (street site (S), urban background site (UB), and regional background site (RB)) and source of error. The blue shaded area contains results for the impact of positional error (30 and 4 meters). The boxplot for 'ALL' represents the error as the result of the combined error in the all exposure determinants (not incorporating positional error) of the LUR model that were included in the simulation. DTR (inverse distance to the nearest road), REG (regional estimate), PD (population density in a 5000 m buffer), TL50 (traffic load in a 50 m buffer), RL1000 (road length of all roads in a 1000 m buffer), HTL25 (heavy traffic load in a 25 m buffer), HTL500 (heavy traffic load in a 25 m to 500 m buffer), MRL50 (road length of all major roads in a 50 m buffer), TML1000 (traffic load on major roads in a 1000 m buffer)

error in REG. Error in traffic on a major road in a 1000 m buffer (TML1000) contributed to some (<0.5%) error at S and UB sites, but not at RB sites. Positional error within a 4 m radius in the geographical location had very little impact on predicted $PM_{2.5}$ concentrations (RSD < 0.2%), regardless of site type. Impact of positional error in a 30 m radius was restricted almost completely to street sites (median RSD 1.7%). In Supplementary Figs. S9 and S10 we show the impact of positional error within a 4 m and 30 m radius on the variance of the standardized exposure determinants at both monitoring and exposure determinant sites.

Health simulation

For each simulation we calculated as a reference the results from a two-stage analysis where the true (i.e., error-free) exposure determinants were available at both monitoring and prediction sites (first column with estimates in Tables 2 and 3). The results indicate that the slope estimates (β_1) from this analysis are slightly biased downwards (ranging from −4.85 to −4.95), probably due to error in the LUR model parameter estimates. Average β_1 for the $PM_{2.5}$ -health model were generally slightly lower than the average β_1 for

Table 2 Impact on β_1 and coverage of its bootstrapped confidence intervals of inherent error in exposure determinants at monitoring (P_m) and prediction (P_p) sites of two land use regression models (NO_2 and $PM_{2.5}$). True beta is −5.00

Model[a,b]	Exposure determinants constrained[c]	No error on P_m, no error on P_p[a,b,d]	Error on P_m, error on P_p[a,b,d]	Error on P_p, no error on P_m[a,b,d]	Error on P_m, no error on P_p[a,b,d]
NO_2	Yes	−4.94 (−5.33, −4.61) *92.5%*	−4.90 (−5.39, −4.45) *92.5%*	−4.42 (−4.84, −3.94) *44.5%*	−5.03 (−5.50, −4.59) *94.5%*
NO_2	No	−4.94 (−5.34, −4.56) *99.5%*	−4.93 (−5.37, −4.47) *93.0%*	−4.27 (−4.79, −3.75) *35.5%*	−5.06 (−5.57, −4.61) *97.0%*
$PM_{2.5}$	Yes	−4.94 (−6.13, −4.03) *91.5%*	−4.89 (−6.06, −4.05) *93.0%*	−4.58 (−5.83, −3.71) *80.0%*	−5.11 (−6.28, −4.19) *94.5%*
$PM_{2.5}$	No	−4.92 (−6.07, −4.10) *89.5%*	−4.93 (−6.07, −4.09) *92.5%*	−4.56 (−5.54, −3.75) *79.5%*	−5.14 (−6.42, −4.17) *95.5%*

[a]Mean, 5th and 95th percentile of β_1 estimated in 200 simulations. True beta is −5.00. Parameter reflects a realistic estimate in percentage decrease in FEV1 in children for a roughly 5 µg/m^3 increase in $PM_{2.5}$and a 20 µg/m^3 increase in NO_2 reported in the PIAMA cohort

[b]Italic values: percentage of 95% confidence intervals of β_1, calculated during each simulation, that contains the true beta

[c]''Yes'': exposure determinants of the land use regression (LUR) model were constrained within the minimum and maximum values that were observed during the monitoring campaign, ''no'': exposure determinants were not constrained

[d]P_m monitoring sites, P_p prediction sites

Table 3 Impact on β_1 and coverage of its bootstrapped confidence intervals of positional error on monitor (P_m) and prediction (P_p) sites of two land use regression models (NO_2 and $PM_{2.5}$)

Model[a,b]	Positional error[c]	Exposure determinants constrained[d]	No error on P_m, no error on P_p[a,b,e]	Error on P_m, error on P_p[a,b,e]	Error on P_p, no error on P_m[a,b,e]	Error on P_m, no error on P_p[a,b,e]
NO_2	4 m	Yes	−4.92 (−5.31, −4.54) *91.0%*	−4.92 (−5.30, −4.54) *92.5%*	−4.88 (−5.28, −4.50) *91.5%*	−4.94 (−5.32, −4.57) *93.0%*
NO_2	4 m	No	−4.95 (−5.36, −4.60) *95.5%*	−4.94 (−5.36, −4.55) *92.0%*	−0.95 (−3.12, −0.21) *8.0%*	−4.98 (−5.41, −4.57) *95.5%*
NO_2	30 m	Yes	−4.92 (−5.31, −4.54) *91.0%*	−4.91 (−5.38, −4.45) *92.5%*	−4.46 (−4.91, −4.00) *55.0%*	−5.02 (−5.49, −4.54) *93.5%*
NO_2	30 m	No	−4.95 (−5.36, −4.60) *95.5%*	−4.92 (−5.36, −4.54) *93.5%*	−0.20 (−0.46, 0.01) *1.5%*	−5.02 (−5.46, −4.61) *96.5%*
$PM_{2.5}$	4 m	Yes	−4.90 (−6.42, −3.77) *91.0%*	−4.90 (−6.38, −3.78) *90.5%*	−4.90 (−6.41, −3.76) *91.5%*	−4.90 (−6.39, −3.79) *91.5%*
$PM_{2.5}$	4 m	No	−4.85 (−6.38, −3.80) *91.5%*	−4.85 (−6.41, −3.83) *90.5%*	−4.84 (−6.37, −3.80) *91.5%*	−4.85 (−6.42, −3.83) *92.0%*
$PM_{2.5}$	30 m	Yes	−4.85 (−6.42, −3.70) *86.0%*	−4.83 (−6.56, −3.62) *85.5%*	−4.57 (−5.96, −3.43) *82.0%*	−5.10 (−7.06, −3.84) *93.5%*
$PM_{2.5}$	30 m	No	−4.85 (−6.38, −3.80) *91.5%*	−4.83 (−6.57, −3.66) *91.5%*	−4.56 (−6.15, −3.58) *80.5%*	−5.10 (−6.87, −3.85) *95.5%*

[a]Mean, 5th and 95th percentile of β_1 estimated in 200 simulations. True beta is −5.00. Parameter reflects a realistic estimate in percentage decrease in FEV1 in children for a roughly 5 µg/m^3 increase in $PM_{2.5}$ and a 20 µg/m^3 increase in NO_2 reported in the PIAMA cohort

[b]Italic values: percentage of 95% confidence intervals of β_1, calculated during each simulation, that contains the true beta

[c]Error (in meters) in the geographical location of monitoring sites (P_m) and prediction sites (P_p)

[d]''Yes'': exposure determinants of the land use regression (LUR) model were constrained within the minimum and maximum values that were observed during the monitoring campaign, ''No'': exposure determinants were not constrained

[e]P_m monitoring sites, P_p prediction sites

the NO_2-health model, reflecting the lower coefficient of determination (R^2 66.6%) of the $PM_{2.5}$ exposure model. Coverage of the parametric bootstrap derived confidence intervals varied around 95% (deviations due to the limited number of simulations).

Compared to the reference scenario, we observed only limited additional bias in slopes for the scenario where we added the same degree of error to exposure determinants at both monitoring and prediction sites. For inherent error (Table 2) mean additional bias was <0.8% for the NO_2 model and <1.0% for the $PM_{2.5}$ model. For positional error (Table 3) the mean additional bias was between −0.2% and −0.6% for the NO_2 model and between 0.2% and −1.0% for the $PM_{2.5}$ model. For both pollutants constraining of the exposure determinants had a small but unpredictable (both up- and downwards) impact on the mean additional bias. Mean additional bias increased with increasing error in estimated geographical location of the prediction sites (4 m vs. 30 m). Coverage of 95% confidence intervals was generally unaffected for all settings.

Considerable additional negative bias (compared to the reference scenario) was observed in the scenario where we added error to the exposure determinants at prediction sites, but not at monitoring sites. For inherent error (Table 2) the additional biases were −10.5 and −13.6% for the NO_2 model and −7.3% (twice) for the $PM_{2.5}$ model. For positional error (Table 3) large additional mean bias was observed for the NO_2 model when exposure determinants were not constrained (−80.8% for 4 m error and −96.0% for 30 m). Average bias was restricted when exposure determinants were constrained (−0.8% (4 m) and −9.3% (30 m)). For the $PM_{2.5}$ model, mean bias ranged between 0.0% and −6.0%, increased with increasing positional error of the prediction sites (4 m vs. 30 m) and was restricted when exposure determinants were constrained. Coverage of the 95% confidence intervals changed in parallel to the bias induced in β_1.

Considerable additional positive bias (compared to the reference scenario) was observed in the scenario where we added error to the exposure determinants on the monitoring

sites, but not on the exposure determinant sites (mean additional bias ranging from +0.4% to +2.6% for the NO_2 model and from +0% to +5.6% for the $PM_{2.5}$ model). Mean bias increased with increasing positional error of the prediction sites, while constraining of the exposure determinants had minimal impact. Coverage of the 95% confidence intervals changed in parallel to the bias induced in β_1.

Results in Supplementary Tables S3–S6 indicate that removing the covariance between exposure determinants resulted in stronger bias in all scenarios were exposure determinants were affected by error. Removing the covariance between the errors of the exposure determinants had only negligible impact under all scenarios.

Discussion

Impact of error in determinants on exposure predictions

Our analysis demonstrates that, perhaps with the exception of the regional contribution of NO_2 or $PM_{2.5}$ at each site (REG), added error in the exposure determinants that were used in our LUR models did not induce large error in exposure predictions. Realistic error in exposure determinant REG had the largest influence on error in predictions of both NO_2 and the $PM_{2.5}$. Based on its large contribution to the coefficient of determination of the LUR models used in this exercise this was expected, and it shows that accurate measurement of this exposure determinant is important. Within the ESCAPE study, this was done by conducting three 2-week sampling campaigns at 20 sites (10 for $PM_{2.5}$) to estimate the long-term average regional contribution to ambient air pollution levels [2, 3]. An approach to reduce the influence of the error in the REG would be to further increase the number of measurements per measurement site to reduce the error in the average. Alternatively, the REG could be better characterized by utilizing additional data to represent regional variation in ambient air pollution such as satellite data or modeled regional background levels from chemical transport models.

In our simulation the impact on exposure predictions of realistic error in model exposure determinants on monitoring and exposure determinant sites was higher for the NO_2 model than for the $PM_{2.5}$ model. As the LUR models primarily captured small scale spatial variability in ambient air pollution, these observations correspond to the notion that spatial variability in NO_2 concentrations, for which local traffic is a primary source, occurs on a smaller scale than spatial variability in $PM_{2.5}$ concentrations, to which many sources contribute in addition to local traffic [23]. This is also reflected in the higher number of exposure determinants of small scale variation that is included in the NO_2 model compared to the $PM_{2.5}$ model.

Impact of error in exposure determinants on bias in health-effect estimates

Using a simulation framework we showed that realistic (non-differential) error in the exposure determinants of the ESCAPE LUR models introduced bias in health-effect estimates when there was an imbalance in the degree of error between exposure determinants at the monitoring sites and exposure determinants at the prediction sites.

We explain this observation as follows: error in the exposure determinants of the exposure model results in a wider distribution of the measured exposure determinant values compared to the true exposure determinant values at the monitoring sites. Consequently, parameter estimates for the measured exposure determinants derived from a linear regression model will be biased. When these biased parameter estimates are used on a new set of exposure determinants at the exposure determinant sites that are equally affected by error (have a similar error structure) the resulting predicted exposure distribution will approach the true exposure distribution and health-effect estimates will be largely unbiased. However, when exposure determinants at the monitoring sites are not affected by error and exposure determinants at the prediction sites are, parameter estimates for the measured exposure determinants will not be biased, but the exposure distribution predicted for the prediction sites will be wider than the true exposure distribution and health-effect estimates will be biased downwards. Furthermore, in the (admittedly unrealistic) situation of exposure determinants at the monitoring sites are affected by error, but exposure determinants at the prediction sites being unaffected, the exposure distribution predicted for the prediction sites will be narrower than the true exposure distribution and health-effect estimates will be biased upwards.

Our observation contributes to the realization that one sided investments in improving the quality of exposure determinants at monitoring sites while similar improvements cannot be made at the prediction sites will introduce bias in health-effect estimates. An example of such a situation would be the location of the monitoring sites being determined using higher quality information (e.g., using Global Positioning System equipment) and, therefore, measured with less error than the locations of the prediction sites (e.g., only determined with ‘‘building matching techniques’’). In this example bias in health-effect estimates will be reduced when similar techniques are used on both monitoring and prediction sites.

Results from our simulation are also consistent with patterns that have been observed in studies that assessed the impact of the quality of geocoding techniques on the

magnitude of the association between exposure to air pollution and (simulated) health outcomes. In these studies the strength of the relationship between air pollution and disease decreased with decreasing geocoding accuracy [20, 24].

Supplementary Figs. S9 and S10 illustrate that in this simulation a similar degree of error in the location of a site had a differential effect on the error in the model exposure determinants between monitoring and prediction sites. This phenomenon occurred because there was a difference in the distribution of site characteristics between monitoring and prediction (e.g., on average more monitoring than prediction sites in an urban area or vice versa) and highlights the importance of the exposure determinants at monitoring sites having the same distribution as the exposure determinants at prediction sites (no mis-specification) [25], a situation that is not necessarily reached when monitoring sites are selected to achieve highest contrast in exposure determinants (as was done in ESCAPE and other studies) [1]. A weighting approach (using weights representing exposure determinant distributions from the study population in the LUR development phase) might be a viable approach to correct for this source of bias.

A number of other sources of measurement error in modeled exposure estimates have been described in air pollution epidemiology studies using LUR models [13, 26–29]. For example, Basagaña et al. [30] showed that the number of monitoring sites, the number of available exposure determinants for model selection, and the amount of explainable variability in the true exposure had a large impact on the estimated LUR model parameters and assigned air pollution exposure, and consequently substantially attenuated health-effect estimates (up to 78% in the worst case scenario). Alexeeff et al. [31] demonstrated the impact of model mis-specification, showing that a LUR models with low *R*-squared sometimes yielded biases ranging from 60% upward bias to 70% downward in the health-effect estimates. The degree of bias we report here is much smaller than what has been reported in these two studies, thereby suggesting that error in model exposure determinants is not a major contributor to bias in air pollution epidemiology studies based on LUR models.

Several methods have been developed to correct health-effect estimates for the bias and error introduced by the use of a model that does not account for all sources of the ''true'' exposure (model mis-specification) or from error in estimating exposure model parameters (see refs [25, 28, 29] for some examples). However, the bias due to imbalance in error in the model exposure determinants described here cannot be corrected for by these methods.

One limitation of our simulation was a necessarily limited number of iterations ($n = 200$) that was possible in each scenario. Using a *t*-test we compared the distributions of the generated health-effect estimates in each scenario (Supplementary Tables S3 and S4). We observed that, with the exception of the scenario in which 4 m positional error was introduced, sufficient statistical power was obtained in all other scenarios.

In conclusion, our results suggest that modest reductions in bias in estimated effects of LUR modeled air pollutants on health outcomes are to be expected from reducing error in model exposure determinants. To minimize bias in health-effect estimates it is important that (errors in) exposure determinants at monitoring sites and at prediction sites have a similar distribution thereby adhering to the assumption of transportability.

Data availability

Data and computing code required to replicate the results reported in this submission are available upon request from the corresponding author.

Acknowledgements This work was supported by grant 211250 (ESCAPE) and grant 308610 (EXPOSOMICS) from the European Community's Seventh Framework Program (FP7/2007–2011).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Hoek G, et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos Environ. 2008;42:7561–78.
2. Eeftens M, et al. Development of land use regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. Environ Sci Technol. 2012;46:11195–205.
3. Beelen R, et al. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe—The ESCAPE project. Atmos Environ. 2013;72:10–23.
4. Beelen R, et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet. 2014;383:785–95.
5. Fuertes E, et al. Associations between particulate matter elements and early-life pneumonia in seven birth cohorts: Results from the ESCAPE and TRANSPHORM projects. Int J Hyg Environ Health. 2014;217:819–29.
6. Stafoggia M, et al. Long-term exposure to ambient air pollution and incidence of cerebrovascular events: results from 11 European cohorts within the ESCAPE project. Environ Health Perspect. 2014;122:919–25.
7. Fuks KB, et al. Arterial blood pressure and long-term exposure to traffic-related air pollution: an analysis in the European Study of Cohorts for Air Pollution Effects (ESCAPE). Environ Health Perspect. 2014;122:896–905.

8. Cai Y, et al. Cross-sectional associations between air pollution and chronic bronchitis: an ESCAPE meta-analysis across five cohorts. Thorax. 2014;69:1005–14.
9. Adam M, et al. Adult lung function and long-term air pollution exposure. ESCAPE: a multicentre cohort study and meta-analysis. Eur Respir J. 2014;45:38–50.
10. Cesaroni G, et al. Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project. BMJ. 2014;348:f7412.
11. Pedersen M, et al. Ambient air pollution and low birthweight: A European cohort study (ESCAPE). Lancet Respir Med. 2013;1: 695–704.
12. Gryparis A, Paciorek CJ, Zeka A, Schwartz J, Coull BA. Measurement error caused by spatial misalignment in environmental epidemiology. Biostatistics. 2009;10:258–74.
13. Spiegelman D. Approaches to uncertainty in exposure assessment in environmental epidemiology. Annu Rev Public Health. 2010; 31:149–63.
14. Carroll, RJ, Ruppert, D, Stefanski, LA, Crainiceanu, CM. Measurement error in nonlinear models: A modern perspective, 2nd ed. 2006. Boca Raton, FL: Chapman & Hall/CRC Press.
15. Cyrys J, et al. Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. Atmos Environ. 2012;62:374–90.
16. Eeftens M, et al. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2—Results of the ESCAPE project. Atmos Environ. 2012;62:303–17.
17. Mendoza D, et al. Implications of uncertainty on regional CO2 mitigation policies for the U.S. onroad sector based on a high-resolution emissions estimate. Energy Policy. 2013;55:386–95.
18. Kadaster. Address Coordinates Netherlands (ACN)—Quality survey 2000 [Adres Coordinaten Nederland (ACN)—Kwaliteitsonderzoek 2000]. 2001. Apeldoorn, The Netherlands: Kadaster.
19. Beekhuizen J, et al. Impact of input data uncertainty on environmental exposure assessment models: A case study for electromagnetic field modelling from mobile phone base stations. Environ Res. 2014;135C:148–55.
20. Jacquemin B, et al. Impact of geocoding methods on associations between long-term exposure to urban air pollution and lung function. Environ Health Perspect. 2013;121:1054–60.
21. Brunekreef B, et al. The prevention and incidence of asthma and mite allergy (PIAMA) birth cohort study: Design and first results. Pediatr Allergy Immunol. 2002;13(Suppl 1):55–60.
22. Gehring U, et al. Air pollution exposure and lung function in children: The ESCAPE project. Environ Health Perspect. 2013; 121:1357–64.
23. Hagemann R, et al. Spatial variability of particle number concentrations and NOx in the Karlsruhe (Germany) area obtained with the mobile laboratory 'AERO-TRAM'. Atmos Environ. 2014;94:341–52.
24. Mazumdar S, Rushton G, Smith BJ, Zimmerman DL, Donham KJ. Geocoding accuracy and the recovery of relationships between environmental exposures and health. Int J Health Geogr. 2008;7:13.
25. Szpiro AA, Sheppard L, Lumley T. Efficient measurement error correction with spatially misaligned data. Biostatistics. 2011;12: 610–23.
26. Szpiro AA, Paciorek CJ. Measurement error in two-stage analyses, with application to air pollution epidemiology. Environmetrics. 2013;24:501–17.
27. Lopiano KK, Young LJ, Gotway CA. A comparison of errors in variables methods for use in regression models with spatially misaligned data. Stat Methods Med Res. 2011;20:29–47.
28. Lopiano KK, Young LJ, Gotway CA. Estimated generalized least squares in spatially misaligned regression models with Berkson error. Biostatistics. 2013;14:737–51.
29. Chang HH, Peng RD, Dominici F. Estimating the acute health effects of coarse particulate matter accounting for exposure measurement error. Biostatistics. 2011;12:637–52.
30. Basagaña X, et al. Measurement error in epidemiologic studies of air pollution based on land-use regression models. Am J Epidemiol. 2013;178:1342–6.
31. Alexeeff SE, et al. Consequences of kriging and land use regression for PM2.5 predictions in epidemiologic analyses: Insights into spatial variability using high-resolution satellite data HHS Public Access. J Expo Sci Env Epidemiol. 2015;2540: 138–44.