Modeling exposure to air pollution and cardiovascular mortality: the ESCAPE study

Meng Wang

Meng Wang, 2013
Modeling exposure to air pollution and cardiovascular mortality: the ESCAPE study
Thesis Utrecht University
ISBN: 978-90-393-6061-3

Cover design by Yiyun Zha, Faculty of Art and Design, University of Lapland
Layout by Meng Wang
Printed by Ridderprint Offsetdrukkerij BV, Ridderkerk, the Netherlands

Modeling exposure to air pollution and cardiovascular mortality: the ESCAPE study

Het modelleren van luchtverontreiniging en cardiovasculaire mortaliteit: de ESCAPE studie

(met een samenvatting in het Nederlands)

空气污染模型及心血管疾病死亡率的欧洲多城市队列研究(ESCAPE)

(附中文简介)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op donderdag 5 december 2013 des middags te 2.30 uur

door

Meng Wang

geboren op 31 juli 1983

te Kunming, Yunnan province, China

Promotor: Prof. dr. B. Brunekreef
Co-promotoren: Dr. G. Hoek
Dr. R.M.J. Beelen

The studies described in this thesis were funded by the European Community's Seventh Framework Program (FP7/2007-2011) projects ESCAPE (Grant agreement number: 211250) and TRANSPHORM (ENV.2009.1.2.2.1).

Contents

Chapter 1

General introduction

Air pollution is a complex mixture of particulate and gaseous pollutants originating from both natural and anthropogenic sources. To assess the risks of large population exposure to air pollution, exposure modeling is considered as a useful surrogate for exposure monitoring. Although the chronic effect of air pollution on cardiovascular mortality has been recognized in the past decades, little is known on which particulate components are most harmful. A well-established modeling approach and the combination of large cohort studies were used in this study to explore the long-term association between air pollution and cardiovascular mortality in Europe.

Exposure modeling

Exposure assessment is one of the key issues for health effect estimates in environmental epidemiology. Although many epidemiological studies collected detailed personal information on health outcomes and lifestyle factors, estimation of individual long-term exposure to air pollution remains a challenge. Two landmark cohort studies, the Harvard Six Cities[1] and the American Cancer Society (ACS) studies[2], compared mortality outcomes between cities using group level average exposures from central monitoring sites within each city. However, numerous studies have documented that air pollution concentrations of major air pollutants varied substantially on a small geographical scale[3-6]. For instance, between-site variation of nitrogen dioxide (NO_2) accounted for over 60% of total variation in four European cities[5]. Large spatial variability of fine particles has been characterized in Amsterdam, the Netherlands[7] and Beijing, China[8]. Traffic-related pollutants such as NO_2, fine particles and black carbon decreased substantially within 100m from a Los Angles highway and a highway in Somerville in the United Sates[6, 9]. One epidemiological study has suggested larger cardiovascular mortality risks of particulate matter variations within than between cities[10]. Because of the complexity of street configurations and pollution dispersion in a city, monitoring alone is not feasible to capture population exposures to outdoor air pollution at intra-urban scales. Moreover, routine monitoring stations often lack sufficient spatial density to allow intra-urban epidemiological studies.

Over the recent decade, exposure studies have therefore attempted to characterize spatial variation in urban air pollution. Various approaches have been developed to improve the quality of exposure estimation[11, 12]. Exposure modeling was considered as a useful surrogate for expensive and time-consuming exposure monitoring which is usually a major challenge in large epidemiological studies. Exposure assessment approaches used in previous studies have included simple indicator variables (e.g. traffic intensity in proximity to residence[13, 14] or distance to major road[15, 16] as exposure variables), interpolation techniques (e.g. kriging and inverse distance weighing)[17] and conventional dispersion models[18, 19] to assess human exposure to air pollution within a city. One of the limitations for the indicator variables is that the implicit assumption of air pollution dispersion pattern may be violated by wind patterns and topography conditions. Furthermore, no quantitative information on exposure is obtained. Interpolation techniques usually require spatially dense distribution of sampling sites and dispersion models are subject to

relatively low spatial resolution in meteorological data and absence of accurate emission data on a small scale[20].

With the improvement of the accuracy of geographical data, air pollution models incorporating geographical information system (GIS) are of increasing interest in exposure assessment[21, 22]. In recent epidemiological studies, land use regression (LUR) has been increasingly applied. Land use regression modeling is a GIS and statistics based method that exploits land use, geographic and traffic characteristics (e.g. traffic intensity, road length, population density) to explain spatial concentration contrasts at monitoring sites. The advantages of LUR modeling include its powerful computer mapping and spatial analyses technology which integrate a wide range of potential sources relating to air pollution. Unlike interpolation techniques solely relying on monitoring sites to smooth data, LUR modeling is capable to predict fine scale variations of air pollutants for individual level exposure by means of spatially refined explanatory predictor variables. Furthermore, LUR modeling requires considerably fewer resource inputs and less computing time than dispersion models[23]. Finally, LUR models are based on real measurements.

Land use regression models have been initially developed in the SAVIAH (Small Area Variation In Air quality and Health) study to investigate NO_2 spatial variation within a few cities in Europe[24] and subsequently applied in epidemiology in the past decade (Briggs et al. 2005). In the TRAPCA (Traffic-Related Air Pollution on Childhood Asthma) study, LUR models have been used to estimate $PM_{2.5}$ (particle dynamic diameter <2.5 µm) and absorbance (a marker for black carbon) at residences of three ongoing birth cohorts throughout the Netherlands, in Munich, Germany and Stockholm county, Sweden[25]. Nowadays, numerous epidemiological studies have widely expanded this approach throughout Europe, the United States and Canada for long-term exposure to air pollutions on diverse health outcomes[21]. Because of limitations in cost and logistics, most of the LUR studies focused on exposure to air pollutants that can be measured with passive sampler, particularly NO_2. There are also an increasing number of studies on particulate matter ($PM_{2.5}$)[21]. A few studies have investigated the development of LUR models for ultrafine particle[26, 27,28] or for species of volatile organic chemicals (VOCs) such as aromatic hydrocarbons (BTEX) in Spain[29-31] and total VOCs in Canada and the United States[32-34]. At the onset of this study, no particle composition LUR models (either metal or chemical compounds) had been developed, with the exception of elemental / black carbon[21]. LUR models can be applied to most pollutants if a suitably dense monitoring network and relevant source-related variable GIS datasets are available[25]. .

Even though the LUR modeling technique has been extensively studied in the past decade, performances of LUR models associated with prediction ability to outdoor exposure have not been well explored. In a review of LUR models, Hoek et al. (2008) raised several determinants such as site selection, precision of geographic data and modeling strategies which may directly influence model performances. The review noted that different methods to characterize prediction ability of models were used in LUR studies. All studies acknowledged that the model percentage explained variability (R^2) provides a too optimistic value.

Techniques that have been used include leave one out cross-validation (LOOCV) and hold out validation (HV). The LOOCV is used through successively excluding one data point and simultaneously estimating the model parameters on the remaining N-1 sites. The HV applies the models developed for the training sites to estimate concentrations at external test monitoring sites. There are no empirical LUR studies that have systematically compared results of LOOCV and HV evaluations using training and test sets of varying sample sizes. Recently, Johnson et al.(2011) estimated NOx, benzene and $PM_{2.5}$ concentrations in 318 census blocks in New Haven with a dispersion model and used the modeled concentrations for a systematic LUR model evaluation. Results suggested relatively poor external validity for models based on small numbers of training sites(20-40). However, since the air pollution concentrations at the sites in that study were not actually measured but modeled, there is a need to explore LUR model performance with varying sample sizes from real world measurements at sites specifically selected for LUR model building.

Finally, the efficiency of the application of LUR models would be increased significantly if models developed in one area could be used in another area. Hence, there is an increased interest to investigate the transferability of LUR models to other areas with similar predictor variables. As LUR models are empirical models, transferability is a more critical issue than it is in dispersion models based upon physical principles.

Air pollution and cardiovascular mortality

Associations between anthropogenic air pollution and human health effects have been acknowledged for a few decades (Brunekreef and Holgate, 2002). After several severe pollution events such as the London fog (1952)[35], mortality risks triggered by air pollution have been widely studied. Pollutants of interest included gaseous pollutants e.g. sulfur dioxide (SO_2), NO_2 and ozone and particulate matter including Black Smoke. In recent decades, a growing concern of potential harmful effects of ambient air pollutions on human health has been highlighted in terms of particulate matter (airborne dynamic diameters < 2.5μm, $PM_{2.5}$; <10μm, PM_{10}; and between 2.5μm and 10μm, coarse particles). Two large cohort studies, the Harvard Six Cities (Dockery et al. 1993) and the ACS study (Pope 3rd et al. 1995) in the United States found significant associations between long-term exposure to $PM_{2.5}$ and natural- cause and cardio-pulmonary mortality.

Airborne PM consists of a heterogeneous mixture of solid and liquid particles suspended in air, varying in size and chemical composition in space and time [36].Effects of air pollution on cardiovascular mortality often contribute a majority of the effects on all-cause mortality[37, 38]. Ambient PM pollution has been noticed as one of the major contributor to cardiovascular disease. Since PM exposure is ubiquitous worldwide, its accumulative burden on public health may be substantial[38, 39].

Effects of air pollution on cardiovascular mortality can be both acute and chronic. Therefore, relevant epidemiological studies have been designed in both short- and long-term scale. Research linking short-term exposure to PM to cardiovascular

mortality was more numerous and was often based on large populations in single cities or more recently multiple cities worldwide[37, 38]. One of the largest was the National Morbidity, Mortality and Air Pollution Study (NMMAPS) in the United States including 50 million people of the 20 largest cities[40]. In this study, the excess risk of cardiopulmonary mortality for each increase in the PM_{10} level of 10μg/m^3 was 0.68% (95% confidence interval (CI) 0.20-1.16%). Another large study which was conducted in Europe, the Air Pollution and Health: A European Approach (APHEA and extended APHEA-2) projects, investigated associations between fine particle and mortality outcomes within 43 million people in 29 European cities[41, 42]. The estimated increase in daily cardiovascular deaths was 0.69% (95% CI 0.31% to 1.08%) for each 10μg/m^3 increase in PM_{10}, which was consistent with the results of the NMMAPS study[40]. Recent studies have been expanded to Asian cities such as in China[43], Japan[44], South Korea[45] and Thailand[43], supporting the short-term effects of particles on excess cardiovascular mortality.

Although short-term effects of air pollution on cardiovascular mortality have been extensively documented, additional effects of long-term exposure may occur[46]. Emphasis has therefore been on cohort studies, with efforts to examine the associations between long-term exposure to air pollution and chronic health effects. Important confounding variables of personal differences have been extensively controlled in several perspective cohort studies to minimize bias in air pollution effect estimates. Long-term exposure to air pollutant is crucial in assessing the true health-damaging effects of air pollution[47] such as the ascertainment of life expectancy. The first large prospective cohort study that demonstrated effects of long-term air pollution exposure on mortality was the Harvard Six Cities study in the United States[1]. This study suggested that fine particles played a vital role in excess cardiopulmonary mortality based on a cohort of 8111 adults with 14-16 years follow-up. The findings were supported by the much larger ACS study in which approximately 500,000 adults living in 50 states were followed up from 1982 to 1989[2]. Potential personal confounding variables such as tobacco smoking and other covariates have been well adjusted for in both studies. In the follow-up of the Harvard Six Cities and the ACS studies, cardiovascular mortality risk was separately analyzed from the cardiopulmonary combination, suggesting higher risks of $PM_{2.5}$ exposure for cardiovascular- than for pulmonary- and overall natural causes[48, 49]. In the extended analysis of the Harvard Six Cities study, Laden et al. (2006)[48] concluded that reduced $PM_{2.5}$ concentrations were associated with reduced cardiovascular mortality. The ACS extended study provided the first opportunity to examine a broad category of cardiovascular effects of which ischemic heart disease (IHD) showed the largest increase in risk (RR 1.18, 95% CI 1.14-1.23) per 10μg/m^3 and dominated the total death proportion[49].

Since the two landmark cohort studies, there is a growth of cohort studies attempting to study air pollution mortality associations in other American, Canadian and Asian cities. Meanwhile, studies on small-scale variations in air pollution concentrations became popular. Evidence of elevated risk of cardiovascular mortality associated with exposure to $PM_{2.5}$ was generally supported by the Women's Health Initiative (WHI)[10], Vancouver[50], Rome[51] and Canadian national cohort[52] studies but not by others (e.g. Netherlands national cohort (NLCS-AIR)[53],

Health Professionals[54], truckers[55] and California teacher studies[56] in the United States). A pooled meta-analysis by Hoek et al. (2013) calculated a significant relative risk of $PM_{2.5}$ for cardiovascular mortality of 10.6% (95% CI 5.4-16.0%) per 10µg/m^3, about twice as large as the risk for all-cause mortality. However, significant heterogeneity was observed across studies which could be partially driven by differences in particle composition.

Although mortality effects of PM mass concentrations have been found, little is known which constituents of particles are more toxic and are associated with higher risks. The impact of particle composition on the health effects of particles is currently one of the major research gaps. Several studies showed evidence of acute effects of PM components on cardiovascular mortality, but results differed between study areas[57-60]. Elevated risk effects of a few fuel combustion related elemental trace markers (e.g. vanadium and zinc) were observed in some large studies[61]. Recent reviews concluded that with the current studies it is not possible to identify specific components that are especially harmful[62]. Identification of health effects of specific components is complicated by often high correlations among components from similar sources and different degrees of measurement error[63]. A recent series of studies using a semi-experimental design identified different components for different endpoints including sulfate / nitrate, ultrafine particles, organic carbon and NO_2[64, 65].

There is hardly any study on long-term exposure to PM constituents and cardiovascular mortality except the California teacher studies which suggested that the constituents responsible for the IHD mortality risks derived from combustion of fossil fuel, biomass burning and crustal origin[66]. A major reason for the lack of studies is the lack of spatially resolved monitoring data of particle composition. LUR models for particle composition have not been published. Dispersion models are hampered by uncertainties in emission estimates for specific particle components.

The ESCAPE Project

The European Study of Cohorts for Air Pollution Effects (ESCAPE) is a multi-center European project, which aims at quantifying long-term health impacts of outdoor air pollution. It covers more than 30 ongoing European cohort studies including over 900,000 subjects across all ages. The aim of this study is to utilize harmonized protocols for exposure assessment including air pollution sampling and model development, and epidemiological data analyses with respect to statistical modeling and confounder specifications, followed by pooling of results of cohort-specific health analyses.

ESCAPE exposure assessment has been conducted since October 2008 in 36 European study areas. Air pollution measurements were simultaneously carried out in three two-week periods in each study area including NO_2, $PM_{2.5}$ and PM_{10} mass concentrations and their composition (absorbance and element content). LUR models were developed for each pollutant in individual study areas using a standardized approach and were subsequently applied to predict annual average concentrations of air pollutants at cohort residential addresses.

The ESCAPE epidemiological study included a broad representation of adverse

health effects in four categories: 1) pregnancy outcomes and children's respiratory and allergic outcomes; 2) respiratory morbidity; 3) cardiovascular morbidity; 4) mortality outcomes and cancer incidence. The mortality work package consists of more than 300,000 participants in 22 existing cohorts across 14 European countries. Natural and cardiovascular mortality and cancer incidence were estimated in association with ambient air pollutants.

Thesis aims and outline

This thesis was developed within the framework of the ESCAPE project and with the following specific aims:

1. To evaluate the performance of LUR models in terms of prediction ability
2. To develop LUR models for particle elemental composition
3. To estimate associations between long-term exposure to particle composition and cardiovascular mortality

In **chapter 2**, we compare different methods to express the predictive ability of LUR models, specifically model R^2, LOOCV R^2 and hold-out validation R^2. We examined the effects of the number of training sampling sites on LUR model performance in estimating NO_2 concentrations across the Netherlands. In **chapter 3**, we extend the findings of chapter 2 to evaluate the prediction ability of LUR models for NO_2 and two particle composition metrics ($PM_{2.5}$ absorbance and PM_{10} Cu) in 20 ESCAPE study areas, making use of independent NO_2 measurements. In **chapter 4**, we develop multi-city LUR models for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance using the ESCAPE database based on a much larger number of monitoring sites than the default study-area specific ESCAPE LUR models. We explore the performance in terms of predictive power and transferability. In **chapter 5**, we develop LUR models for PM elemental composition in each ESCAPE study area. I**n chapter 6**, we describe associations between elemental composition estimated by the LUR models and cardiovascular mortality in 19 ESCAPE cohorts. Finally, **chapter 7** discusses the main findings of this thesis and address additional issues and perspectives for future studies.

Reference

1. Dockery, D. W.; Pope, C. A., 3rd; Xu, X.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. G., Jr.; Speizer, F. E., An association between air pollution and mortality in six U.S. cities. *N Engl J Med* **1993,** *329*, (24), 1753-9.
2. Pope, C. A., 3rd; Thun, M. J.; Namboodiri, M. M.; Dockery, D. W.; Evans, J. S.; Speizer, F. E.; Heath, C. W., Jr., Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *Am J Respir Crit Care Med* **1995,** *151*, (3 Pt 1), 669-74.
3. Roorda-Knape, M. C.; Janssen, N. A.; de Hartog, J.; Van Vliet, P. H.; Harssema, H.; Brunekreef, B., Traffic related air pollution in city districts near motorways. *Sci Total Environ* **1999,** *235*, (1-3), 339-41.
4. Fischer, P. H.; Hoek, G.; van Reeuwijk, H.; Briggs, D. J.; Lebret, E.; van Wijnen, J. H.; Kingham, S.; Elliott, P. E., Traffic-related differences in outdoor and indoor concentrations of particles and volatile organic compounds in Amsterdam. *Atmospheric Environment* **2000,** *34*, (22), 3713-3722.
5. Lebret, E.; Briggs, D.; van Reeuwijk, H.; Fischer, P.; Smallbone, K.; Harssema, H.; Kriz, B.; Gorynski, P.; Elliott, P., Small area variations in ambient NO_2 concentrations in four European areas. *Atmospheric Environment* **2000,** *34*, (2), 177-185.
6. Zhu, Y. F.; Hinds, W. C.; Kim, S.; Shen, S.; Sioutas, C., Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmospheric Environment* **2002,** *36*, (27), 4323-4335.
7. Weijers, E. P.; Khlystov, A. Y.; Kos, G. P. A.; Erisman, J. W., Variability of particulate matter concentrations along roads and motorways determined by a moving measurement unit. *Atmospheric Environment* **2004,** *38*, (19), 2993-3002.
8. Wang, M.; Zhu, T.; Zheng, J.; Zhang, R. Y.; Zhang, S. Q.; Xie, X. X.; Han, Y. Q.; Li, Y., Use of a mobile laboratory to evaluate changes in on-road air pollutants during the Beijing 2008 Summer Olympics. *Atmospheric Chemistry and Physics* **2009,** *9*, (21), 8247-8263.
9. Durant, J. L.; Ash, C. A.; Wood, E. C.; Herndon, S. C.; Jayne, J. T.; Knighton, W. B.; Canagaratna, M. R.; Trull, J. B.; Brugge, D.; Zamore, W.; Kolb, C. E., Short-term variation in near-highway air pollutant gradients on a winter morning. *Atmospheric Chemistry and Physics* **2010,** *10*, (17), 8341-8352.
10. Miller, K. A.; Siscovick, D. S.; Sheppard, L.; Shepherd, K.; Sullivan, J. H.; Anderson, G. L.; Kaufman, J. D., Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med* **2007,** *356*, (5), 447-58.
11. Jerrett, M.; Burnett, R. T.; Ma, R.; Pope, C. A., 3rd; Krewski, D.; Newbold, K. B.; Thurston, G.; Shi, Y.; Finkelstein, N.; Calle, E. E.; Thun, M. J., Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* **2005,** *16*, (6), 727-36.
12. Briggs, D., The role of GIS: Coping with space (and time) in air pollution exposure assessment. *Journal of Toxicology and Environmental Health-Part a-Current Issues* **2005,** *68*, (13-14), 1243-1261.
13. Kramer, U.; Koch, T.; Ranft, U.; Ring, J.; Behrendt, H., Traffic-related air pollution is associated with atopy in children living in urban areas. *Epidemiology* **2000,** *11*, (1), 64-70.
14. Brunekreef, B.; Janssen, N. A.; de Hartog, J.; Harssema, H.; Knape, M.; van Vliet, P., Air pollution from truck traffic and lung function in children living near motorways. *Epidemiology* **1997,** *8*, (3), 298-303.
15. Nitta, H.; Sato, T.; Nakai, S.; Maeda, K.; Aoki, S.; Ono, M., Respiratory health associated with exposure to automobile exhaust. I. Results of cross-sectional studies in 1979, 1982, and 1983. *Arch Environ Health* **1993,** *48*, (1), 53-8.
16. Edwards, J.; Walters, S.; Griffiths, R. K., Hospital admissions for asthma in preschool children: relationship to major roads in Birmingham, United Kingdom. *Arch Environ Health* **1994,** *49*, (4), 223-7.
17. Beelen, R.; Hoek, G.; Pebesma, E.; Vienneau, D.; de Hoogh, K.; Briggs, D. J., Mapping of background air pollution at a fine spatial scale across the European Union. *Science of the Total Environment* **2009,** *407*, (6), 1852-1867.
18. Bellander, T.; Berglind, N.; Gustavsson, P.; Jonson, T.; Nyberg, F.; Pershagen, G.; Jarup, L., Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environmental Health Perspectives* **2001,** *109*, (6),

633-639.
19. Gram, F.; Nafstad, P.; Haheim, L. L., Estimating residential air pollution exposure among citizens in Oslo 1974-1998 using a geographical information system. *Journal of Environmental Monitoring* **2003,** *5*, (4), 541-546.
20. Liu, L. J. S.; Tsai, M. Y.; Keidel, D.; Gemperli, A.; Ineichen, A.; Hazenkamp-von Arx, M.; Bayer-Oglesby, L.; Rochat, T.; Kunzli, N.; Ackermann-Liebrich, U.; Straehl, P.; Schwartz, J.; Schindler, C., Long-term exposure models for traffic related NO_2 across geographically diverse areas over separate years. *Atmospheric Environment* **2012,** *46*, 460-471.
21. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; Briggs, D., A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* **2008,** *42*, (33), 7561-7578.
22. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; Morrison, J.; Giovis, C., A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology* **2005,** *15*, (2), 185-204.
23. Marshall, J. D.; Nethery, E.; Brauer, M., Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmospheric Environment* **2008,** *42*, (6), 1359-1369.
24. Briggs, D. J.; Collins, S.; Elliott, P.; Fischer, P.; Kingham, S.; Lebret, E.; Pryl, K.; VAnReeuwijk, H.; Smallbone, K.; VanderVeen, A., Mapping urban air pollution using GIS: a regression-based approach. *International Journal of Geographical Information Science* **1997,** *11*, (7), 699-718.
25. Brauer, M.; Hoek, G.; van Vliet, P.; Meliefste, K.; Fischer, P.; Gehring, U.; Heinrich, J.; Cyrys, J.; Bellander, T.; Lewne, M.; Brunekreef, B., Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* **2003,** *14*, (2), 228-39.
26. Abernethy, R. C.; Allen, R. W.; McKendry, I. G.; Brauer, M., A Land Use Regression Model for Ultrafine Particles in Vancouver, Canada. *Environmental Science & Technology* **2013,** *47*, (10), 5217-5225.
27. Rivera, M.; Basagana, X.; Aguilera, I.; Agis, D.; Bouso, L.; Foraster, M.; Medina-Ramon, M.; Pey, J.; Kunzli, N.; Hoek, G., Spatial distribution of ultrafine particles in urban settings: A land use regression model. *Atmospheric Environment* **2012,** *54*, 657-666.
28. Hoek, G.; Beelen, R.; Kos, G.; Dijkema, M.; van der Zee, S. C.; Fischer, P. H.; Brunekreef, B., Land Use Regression Model for Ultrafine Particles in Amsterdam. *Environmental Science & Technology* **2011,** *45*, (2), 622-628.
29. Aguilera, I.; Sunyer, J.; Fernandez-Patier, R.; Hoek, G.; -Alfaro, A. A.; Meliefste, K.; Bomboi-Mingarro, M. T.; Nieuwenhuijsen, M. J.; Herce-Garraleta, D.; Brunekreef, B., Estimation of outdoor NOx, NO_2, and BTEX exposure in a cohort of pregnant women using land use regression modeling. *Environmental Science & Technology* **2008,** *42*, (3), 815-821.
30. Kheirbek, I.; Johnson, S.; Ross, Z.; Pezeshki, G.; Ito, K.; Eisl, H.; Matte, T., Spatial variability in levels of benzene, formaldehyde, and total benzene, toluene, ethylbenzene and xylenes in New York City: a land-use regression study. *Environmental Health* **2012,** *11*.
31. Mukerjee, S.; Smith, L.; Neas, L.; Norris, G., Evaluation of Land Use Regression Models for Nitrogen Dioxide and Benzene in Four US Cities. *Scientific World Journal* **2012**.
32. Smith, L. A.; Mukerjee, S.; Chung, K. C.; Afghani, J., Spatial analysis and land use regression of VOCs and NO_2 in Dallas, Texas during two seasons. *Journal of Environmental Monitoring* **2011,** *13*, (4), 999-1007.
33. Atari, D. O.; Luginaah, I. N., Assessing the distribution of volatile organic compounds using land use regression in Sarnia, "Chemical Valley", Ontario, Canada. *Environ Health* **2009,** *8*, 16.
34. Su, J. G.; Jerrett, M.; Beckerman, B.; Verma, D.; Arain, M. A.; Kanaroglou, P.; Stieb, D.; Finkelstein, M.; Brook, J., A land use regression model for predicting ambient volatile organic compound concentrations in Toronto, Canada. *Atmospheric Environment* **2010,** *44*, (29), 3529-3537.
35. Bell, M. L.; Davis, D. L.; Fletcher, T., A retrospective assessment of mortality from the London smog episode of 1952: the role of influenza and pollution. *Environ Health Perspect* **2004,** *112*, (1), 6-8.
36. Seinfeld, J. H.; Pandis, S. N., Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. *New York, NY: John Wiley & Sons* **1998**, 1-1326.

37. Brook, R. D.; Franklin, B.; Cascio, W.; Hong, Y.; Howard, G.; Lipsett, M.; Luepker, R.; Mittleman, M.; Samet, J.; Smith, S. C., Jr.; Tager, I., Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association. *Circulation* **2004,** *109*, (21), 2655-71.
38. Brook, R. D.; Rajagopalan, S.; Pope, C. A., 3rd; Brook, J. R.; Bhatnagar, A.; Diez-Roux, A. V.; Holguin, F.; Hong, Y.; Luepker, R. V.; Mittleman, M. A.; Peters, A.; Siscovick, D.; Smith, S. C., Jr.; Whitsel, L.; Kaufman, J. D., Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **2010,** *121*, (21), 2331-78.
39. Lim, S. S.; Vos, T.; Flaxman, A. D.; Danaei, G.; Shibuya, K.; Adair-Rohani, H.; et al., A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **2012,** *380*, (9859), 2224-60.
40. Samet, J. M.; Dominici, F.; Curriero, F. C.; Coursac, I.; Zeger, S. L., Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. *N Engl J Med* **2000,** *343*, (24), 1742-9.
41. Analitis, A.; Katsouyanni, K.; Dimakopoulou, K.; Samoli, E.; Nikoloulopoulos, A. K.; Petasakis, Y.; Touloumi, G.; Schwartz, J.; Anderson, H. R.; Cambra, K.; Forastiere, F.; Zmirou, D.; Vonk, J. M.; Clancy, L.; Kriz, B.; Bobvos, J.; Pekkanen, J., Short-term effects of ambient particles on cardiovascular and respiratory mortality. *Epidemiology* **2006,** *17*, (2), 230-3.
42. Samoli, E.; Touloumi, G.; Zanobetti, A.; Le Tertre, A.; Schindler, C.; Atkinson, R.; Vonk, J.; Rossi, G.; Saez, M.; Rabczenko, D.; Schwartz, J.; Katsouyanni, K., Investigating the dose-response relation between air pollution and total mortality in the APHEA-2 multicity project. *Occup Environ Med* **2003,** *60*, (12), 977-82.
43. Wong, C. M.; Vichit-Vadakan, N.; Kan, H.; Qian, Z., Public Health and Air Pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. *Environ Health Perspect* **2008,** *116*, (9), 1195-202.
44. Omori, T.; Fujimoto, G.; Yoshimura, I.; Nitta, H.; Ono, M., Effects of particulate matter on daily mortality in 13 Japanese cities. *J Epidemiol* **2003,** *13*, (6), 314-22.
45. Son, J. Y.; Lee, J. T.; Kim, K. H.; Jung, K.; Bell, M. L., Characterization of fine particulate matter and associations between particulate chemical constituents and mortality in Seoul, Korea. *Environ Health Perspect* **2012,** *120*, (6), 872-8.
46. Kunzli, N.; Medina, S.; Kaiser, R.; Quenel, P.; Horak, F., Jr.; Studnicka, M., Assessment of deaths attributable to air pollution: should we use risk estimates based on time series or on cohort studies? *Am J Epidemiol* **2001,** *153*, (11), 1050-5.
47. WHO, Air quality guidelines, Global update 2005. *Copenhagen: WHO Regional Office for Europe* **2006**.
48. Laden, F.; Schwartz, J.; Speizer, F. E.; Dockery, D. W., Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *Am J Respir Crit Care Med* **2006,** *173*, (6), 667-72.
49. Pope, C. A., 3rd; Burnett, R. T.; Thurston, G. D.; Thun, M. J.; Calle, E. E.; Krewski, D.; Godleski, J. J., Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* **2004,** *109*, (1), 71-7.
50. Gan, W. Q.; Koehoorn, M.; Davies, H. W.; Demers, P. A.; Tamburic, L.; Brauer, M., Long-term exposure to traffic-related air pollution and the risk of coronary heart disease hospitalization and mortality. *Environ Health Perspect* **2011,** *119*, (4), 501-7.
51. Cesaroni, G.; Badaloni, C.; Gariazzo, C.; Stafoggia, M.; Sozzi, R.; Davoli, M.; Forastiere, F., Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ Health Perspect* **2013,** *121*, (3), 324-31.
52. Crouse, D. L.; Peters, P. A.; van Donkelaar, A.; Goldberg, M. S.; Villeneuve, P. J.; Brion, O.; Khan, S.; Atari, D. O.; Jerrett, M.; Pope, C. A.; Brauer, M.; Brook, J. R.; Martin, R. V.; Stieb, D.; Burnett, R. T., Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. *Environ Health Perspect* **2012,** *120*, (5), 708-14.
53. Beelen, R.; Hoek, G.; van den Brandt, P. A.; Goldbohm, R. A.; Fischer, P.; Schouten, L. J.; Jerrett, M.; Hughes, E.; Armstrong, B.; Brunekreef, B., Long-term effects of traffic-related air

pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environ Health Perspect* **2008,** *116*, (2), 196-202.
54. Puett, R. C.; Hart, J. E.; Suh, H.; Mittleman, M.; Laden, F., Particulate matter exposures, mortality, and cardiovascular disease in the health professionals follow-up study. *Environ Health Perspect* **2011,** *119*, (8), 1130-5.
55. Hart, J. E.; Garshick, E.; Dockery, D. W.; Smith, T. J.; Ryan, L.; Laden, F., Long-term ambient multipollutant exposures and mortality. *Am J Respir Crit Care Med* **2011,** *183*, (1), 73-8.
56. Lipsett, M. J.; Ostro, B. D.; Reynolds, P.; Goldberg, D.; Hertz, A.; Jerrett, M.; Smith, D. F.; Garcia, C.; Chang, E. T.; Bernstein, L., Long-Term Exposure to Air Pollution and Cardiorespiratory Disease in the California Teachers Study Cohort. *American Journal of Respiratory and Critical Care Medicine* **2011,** *184*, (7), 828-835.
57. Zhou, J.; Ito, K.; Lall, R.; Lippmann, M.; Thurston, G., Time-series analysis of mortality effects of fine particulate matter components in Detroit and Seattle. *Environ Health Perspect* **2011,** *119*, (4), 461-6.
58. Valdes, A.; Zanobetti, A.; Halonen, J. I.; Cifuentes, L.; Morata, D.; Schwartz, J., Elemental concentrations of ambient particles and cause specific mortality in Santiago, Chile: a time series study. *Environ Health* **2012,** *11*, 82.
59. Ito, K.; Mathes, R.; Ross, Z.; Nadas, A.; Thurston, G.; Matte, T., Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. *Environ Health Perspect* **2011,** *119*, (4), 467-73.
60. Cao, J.; Xu, H.; Xu, Q.; Chen, B.; Kan, H., Fine particulate matter constituents and cardiopulmonary mortality in a heavily polluted Chinese city. *Environ Health Perspect* **2012,** *120*, (3), 373-8.
61. Bell, M. L.; Ebisu, K.; Peng, R. D.; Samet, J. M.; Dominici, F., Hospital admissions and chemical composition of fine particle air pollution. *Am J Respir Crit Care Med* **2009,** *179*, (12), 1115-20.
62. Kelly, F. J.; Fussell, J. C., Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment* **2012,** *60*, 504-526.
63. Brunekreef, B.; Russell, A., Particulate matter and human health focus issue. *Environ Sci Technol* **2009,** *43*, (13), 4613-4.
64. Strak, M.; Janssen, N. A.; Godri, K. J.; Gosens, I.; Mudway, I. S.; Cassee, F. R.; Lebret, E.; Kelly, F. J.; Harrison, R. M.; Brunekreef, B.; Steenhof, M.; Hoek, G., Respiratory health effects of airborne particulate matter: the role of particle size, composition, and oxidative potential-the RAPTES project. *Environ Health Perspect* **2012,** *120*, (8), 1183-9.
65. Strak, M.; Hoek, G.; Godri, K. J.; Gosens, I.; Mudway, I. S.; van Oerle, R.; Spronk, H. M.; Cassee, F. R.; Lebret, E.; Kelly, F. J.; Harrison, R. M.; Brunekreef, B.; Steenhof, M.; Janssen, N. A., Composition of PM affects acute vascular inflammatory and coagulative markers - the RAPTES project. *Plos One* **2013,** *8*, (3), e58944.
66. Ostro, B.; Lipsett, M.; Reynolds, P.; Goldberg, D.; Hertz, A.; Garcia, C.; Henderson, K. D.; Bernstein, L., Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study. *Environ Health Perspect* **2010,** *118*, (3), 363-9.

Chapter 2

Systematic Evaluation of Land Use Regression Models for NO_2

Meng Wang
Rob Beelen
Marloes Eeftens
Kees Meliefste
Gerard Hoek
Bert Brunekreef

(Environmental Sciences & Technology 2012, 46, (8), 4481-4489)

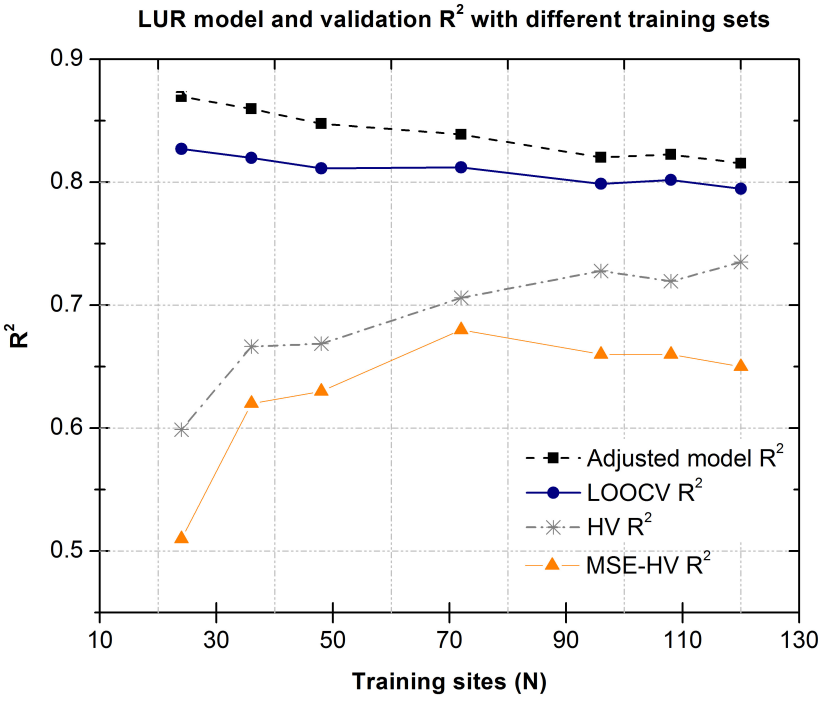
Abstract

Background: Land use regression (LUR) models have become popular to explain the spatial variation of air pollution concentrations. Independent evaluation is important.

Methods: We developed LUR models for nitrogen dioxide (NO_2) using measurements conducted at 144 sampling sites in the Netherlands. Sites were randomly divided into training datasets with a size of 24, 36, 48, 72, 96, 108 and 120 sites. LUR models were evaluated using (1) internal "leave-one-out-cross-validation (LOOCV)" within the training datasets and (2) external "hold-out" validation (HV) against independent test datasets. In addition, we calculated Mean Square Error based validation R^2s.

Results: The mean adjusted model and LOOCV R^2 slightly decreased from 0.87 to 0.82 and 0.83 to 0.79, respectively, with increasing number of training sites. In contrast, the mean HV R^2 was lowest (0.60) with the smallest training sets and increased to 0.74 with the largest training sets. Predicted concentrations were more accurate in sites with out of range values for prediction variables after changing these values to the minimum or maximum of the range observed in the corresponding training dataset.

Conclusion: LUR models for NO_2 perform less well, when evaluated against independent measurements, when they are based on relatively small training sets. In our specific application, models based on as few as 24 training sites, however, achieved acceptable hold out validation R^2s of, on average, 0.60.



(Table of content)

Introduction

Many epidemiological studies have shown that air pollution is associated with adverse health effects such as cardiovascular morbidity and mortality[1,2]. Concentrations of traffic related air pollutants have remarkable spatial variability at the urban scale[3-5]. For epidemiological studies on the effects of air pollution, routine monitoring networks are typically not sufficiently dense to represent small-scale spatial contrasts well. Therefore, interest has increased in exposure modeling incorporating Geographic Information System (GIS) data to capture small-scale spatial variability in air pollution concentrations [6,7].

Land Use Regression (LUR) modeling is a GIS based method that uses land use and traffic characteristics (e.g. traffic intensity, road length, population density) to explain the spatial variation of measured air pollution concentrations with the purpose to estimate long-term air pollution concentrations at unmeasured locations. Nitrogen dioxide (NO_2) is the pollutant for which most LUR models have been reported. The validity of the model predictions strongly relies on the quality of GIS predictor data, the selected monitoring sites, the complexity of the airshed, inherent variability in concentrations and the accuracy of measurements at given sites. The advantage of LUR modeling is the relatively simple input and low cost[7,8]. LUR models, especially when combined with geo-statistical methods such as Kriging may perform at least as well or better than dispersion models[9].

Evaluation is an essential part of model development. The model R^2 is too optimistic as linear regression optimizes the model for the data used whereas we are typically interested in how well the model predicts at unmeasured locations, such as addresses of participants of an epidemiological study. One evaluation approach is the leave-one-out-cross-validation (LOOCV), which is used through successively excluding one data point and simultaneously estimating the model parameters on the remaining N-1 sites. This is popular in many studies to test the internal validity. Another approach is the hold-out validation (HV) where a model is evaluated against measurements conducted at independent external sites[9]. However, few studies so far have reported hold-out validation results e.g. Beelen et al. [10] and Briggs et al. [11]. The required sample size for model-building has no strict minimum but depends on the local determinants of spatial variability[8]. To our knowledge there are no empirical LUR studies that have systematically compared results of LOOCV and HV evaluations using training and test sets of varying sample sizes. Recently, Johnson et al.[12] estimated NOx, benzene and $PM_{2.5}$ concentrations in 318 census blocks in New Haven with a dispersion model and used the modeled concentrations for a systematic LUR model evaluation. Results suggested relatively poor external validity for models based on small numbers of training sites. However, since the air pollution concentrations at the sites in that study were not actually measured but modeled, there is a need to explore LUR model performance with varying sample sizes from real world measurements at sites specifically selected for LUR model building.

For this purpose, we developed LUR models for datasets of varying sample sizes using data from a nation-wide monitoring campaign conducted in the Netherlands

in 2007. The goal was to compare the performance of LUR models assessed with LOOCV and HV, in dependence of the number of training as well as test sites.

Methods

Study area and air pollution measurements.

We used NO_2 concentrations from the 2007 TRACHEA campaign in the Netherlands for LUR model development. Details of the study design and measurement methods have been published before[13]. Briefly, 144 sites were measured spread over the Netherlands (Fig. S1 in the supplement). The sampling sites were divided over 26 regional background, 78 urban background and 40 traffic locations. Measurements were made at the façade of homes, as the goals were to develop a model for residential exposures. Measurements were conducted in four one-week sampling periods in January, April, June and September 2007, covering the four seasons of the year. Sampling took place in the same week for all 144 sites. The mean concentration for each site was calculated and used for model development.

NO_2 has been measured using Ogawa passive samplers and following the Ogawa analysis protocol (Ogawa&Co V 3.98, USA, Inc.). For each batch of 40 filters, the average value of four laboratory blank filters was subtracted.

Predictor variables.

Geographic variables were generated and stored in a Geographic Information System (GIS) using ArcGIS 9.3(ESRI, Redlands, California). In total, 76 potential predictor variables were selected for model development (Table S1 in supplement). For some of the geographical variables, we calculated a number of different buffers around each of the sampling sites.

Land use variables were derived from the CORINE 2000 (COordination and INformation on the Environment programme, European Comission) database with a 100 meter grid cell unit. Specifically, 3 land use variables were considered for model development: low-density residential, industry, and urban green combined with forests and agriculture. Buffer sizes of 100, 300, 500, 1000, 5000 m were evaluated. We did not consider port and airport variables because the number of sites with non-zero values was very small, especially for small training datasets with 24 or 36 measurement sites. For the same reason, we also excluded the industry variable in buffer sizes of 100 and 300 m and urban green combined with forests and agriculture variable within 100 m.

Household and population density were available from a national database with 100m grid cell unit obtained from the National Institute of Public Health and the Environment. We calculated buffer sizes of 100, 300, 500, 1000 and 5000 m for household and population density.

The local road network database (National Road Database (the National Wegen Bestand, NWB)) was used to calculate distance to the nearest road, distance to the nearest major road (≥10.000 mvh/24h), total traffic intensity for the same type of roads, from which intensity of heavy traffic was also evaluated separately. NWB is the most complete and geographically precise road network in the Netherlands. For

all roads, traffic intensity was assigned. Circular buffers were calculated for a radius of 25, 50, 100, 300, 500 and 1000 m. Traffic loads variables were calculated by multiplying the sum of traffic intensity with each road segment within the buffer, in a unit of veh/(day*meter).

LUR Model development and evaluation

We divided the 144 sites into two groups: training datasets and test datasets. The training datasets were used for model development and were divided in fractions of 1/6, 1/4, 1/3, 1/2, 2/3, 3/4, 5/6 of the total number of 144 sites, corresponding to 24, 36, 48, 72, 96, 108 and 120 sites. The complementary sites were used as test datasets for external model evaluation following the approach used by Johnson et al.[12]. The test dataset was not fixed because we did not want the results to depend on one specific dataset. For instance (for details, see Fig.S2 in supplement), the total of 144 sites was equally divided into four subsets of 36 sites. Each subset was alternately picked up as training dataset for modeling and the remaining 108 sites were treated for external prediction. So in this step, four models were developed. Site selection was repeated ten times to be able to study the sampling variability of our procedures which produced 40 different training datasets of 36 sites each. Subsequently, we exchanged the training dataset (n=108) and the test dataset (n=36) for the same model constructions and evaluations. Similar procedures were conducted to obtain training sets of 72, 48 and 24 sites respectively. Hence, a total of 280 training datasets and corresponding test datasets were allocated for model development, as follows: 24 training sites (60 models), 36 sites (40 models), 48 sites (30 models), 72 sites (20 models), 96 sites (30 models), 108 sites (40 models), 120 sites (60 models). In addition, we conducted two sensitivity analyses to compare results obtained using different test set sizes for a fixed training set size, and vice versa. We fixed test sets to 72 to evaluate models based on 24, 36, 48 and 72 training sites. The test and training sets were randomly selected thirty times and therefore 120 models were built. We also built models based on randomly selected fixed training sets (N=24, 36, 48, 72). Each model was evaluated against four test sets (N=24, 36, 48, 72) which were randomly selected from the remaining sites. This work has been repeated thirty times as well, generating 120 models (30 for each training set group)

We performed two selections: stratified and random. The stratified selection took into account that the original 144 sites were deliberately chosen to reflect region of the country, regional background, urban background and street locations. In the random selection we ignored the design of the sampling campaign completely to investigate whether model performance was affected by this difference in selection of training sites.

Procedures of model development were similar to recent studies in the Netherlands[13, 14]. A supervised stepwise regression was used to develop the LUR model. Firstly, all the potential variables were entered separately and the variable that explained the largest percentage of variability in measured concentrations was included. Secondly, the remaining variables were added individually and we assessed whether the adjusted R^2 was increased by at least 1%. This procedure was repeated until no more variables met this criterion. New variables were only

entered in the model if the sign of the slope had the a priori specified direction (Table S1), for example, positive for traffic intensity and negative for urban green. Finally, geographic coordinate variables (X coordinate, Y coordinate, X+Y, X-Y) were included using the same entry criterion. Variables were deleted in the final stage when they were not statistically significant at the 0.10 level.

Several regression diagnostics were performed. We investigated the Cook's distance to examine potential influential observations, and excluded variables from model development which were affected most by problematic sites that resulted in high Cook's distance value (>1). Variables with variance inflation factor (VIF) larger than 3 were also excluded to avoid multicollinearity, starting with excluding the variable with highest VIF.

We used two evaluation approaches: 1) Leave-one-out-cross-validation (LOOCV), which successively leaves out one site from the training dataset and estimates models based on the remaining N-1 sites. In this procedure, the variables in the model were the same as identified using the full training data set, only the coefficients of the model changed. This is a common procedure in LUR model evaluation[15]. 2) Hold-out-validation, which applies the models developed for the training sites to estimate concentrations at the external test dataset monitoring sites. Two prediction errors were estimated: a) regression based R^2 (HV R^2) which was derived from correlations between predicted and observed values. b) mean square error based R^2 (MSE-HV R^2), taking into account absolute values in terms of mean squared prediction error rather than merely correlation as stated in a). The formula was defined as:

$$MSE - HV \;\; R^2 = 1 - \frac{MSE}{\left(\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \overline{y_t} \right)^2 \right)} \qquad (1)$$

where $\overline{y_t}$ is the average prediction in the training sample[16, 17]. MSE-HV R^2 can yield negative values when, in the evaluation (test) set, the average of the observed values performs better, in terms of mean squared error, than the predictions of the model. Both evaluation approaches were performed once for each model and therefore, were iteratively carried out for 280 times each. Prediction errors for both evaluation methods were estimated by root mean squared error (RMSE).

In the paper, we report particularly the percentage explained variability (R^2) calculated for the model and the two evaluation methods (LOOCV and HV).

In addition to the national analysis, we also performed an analysis of the 79 sites in the west and middle region because this study area corresponds more to the typical metropolitan area scale used in many LUR studies.

All analyses were conducted with SAS version 9.2.

Results

Descriptive statistics of the measured NO_2 concentrations are shown in Table 1. One street site in the city center of Utrecht was excluded from further analysis because it was a narrow street canyon-type site which had the highest

concentration of all (63.7 μg/m^3), and serves as a two-way bus lane, carrying diesel-powered buses which made it unlike any other street site in the data base. It did not have particularly high traffic volumes otherwise. When in test sets, we found that many models strongly under-predicted concentrations at this site. As we did not want our exercise results to be dominated by an atypical site, we decided to remove it from further analysis.

The LUR models typically included three to eight independent variables. The mean number of predictors ranged from 4 for the n=24 training sets to 6 for the n=120 datasets (Table S2). As shown in Figure S3, 69 variables appeared at least once in models based on 24 training sites while only 24 variables appeared at least once in models based on 120 training sites. In the latter models, the most frequently included variables were the X+Y coordinate (100%), Population density in a 5000 meter buffer (95%), Industry in a 500 meter buffer (66%) and Inverse distance to major road (50%).

Fig.1 and Table 2 show the variation of adjusted model R^2, LOOCV R^2, HV R^2 and MSE-HV R^2 in relation to the number of sites in the stratified training datasets for both the national and the regional scale. The mean adjusted model R^2 was weakly inversely associated with the number of training sites. At the national scale, the mean adjusted model R^2 decreased from 0.87 (n=24) to 0.82 (n=120). The mean LOOCV R^2 was slightly lower, ranging from 0.83 (n=24) to 0.79 (n=120). In contrast, the mean HV R^2 was positively associated with the number of sites in the training dataset. The models based on 24 training sites explained on average 60% of the variation in the independent test datasets. This increased to 74% when using 120 training sites. The MSE-HV R^2 was 3%-9% lower than HV R^2. Only 2 out of 280 models had negative values of the MSE-HV R^2 (both for models based on 24 training sites), indicating that almost always, models predicted the measured concentrations in the test sets better than the sample mean. The differences between the mean LOOCV and HV R^2, MSE HV R^2 were relatively large for models using 24 training sites with a difference of 0.23 and 0.32, but when using the 120 training sites this difference was reduced to 0.05 and 0.14 respectively. The variability of R^2 across models for the same N was higher for hold-out validation

Table 1 Distribution of measured NO_2 concentrations (μg/m^3) from the TRACHEA study stratified by site types (regional background, urban background and street)

Region	Site[c]	Number	Min	P25	P50	P75	Max
NL[a]	rb	26	10.3	14.5	17.6	20.9	30.1
	ub	78	12.1	19.4	24.6	27.7	40.4
	s	39	18.5	28.4	35.7	42.3	62.7
	Total	143	10.3	19.2	25.7	30.8	62.7
W+M[b]	rb	12	14.5	18.3	20.7	24.1	30.1
	ub	45	17.7	24.9	27.0	29.1	40.4
	s	21	24.5	31.2	40.0	42.8	53.1
	Total	78	14.5	24.5	27.8	32.6	53.1

NL[a]: the whole dataset of the Netherlands;
W+M[b]: west and middle region;
Site[c]: rb-regional background; ub-urban background; s-street.

than for cross-validation and model development. Variability decreased with increasing sample size of the training dataset for model and cross-validation R^2. For hold-out validation, variability was largest for the small training datasets and the largest training datasets. The latter is likely due to the smaller test datasets. The mean RMSE varied in predictable fashion with training set size and evaluation method (Table 3), never exceeding 5.37 which is relatively small compared to the measured range of about 10 to more than 60 $\mu g/m^3$. In the analyses restricted to the regions Middle and West, model performance was similar to the national one. Similar variations of adjusted model, LOOCV, HV and MSE-HV R^2 were observed in models using random site selections for the training datasets, compared to models based on a stratified selection of training sites. For 24 training sites, adjusted model, LOOCV, HV and MSE-HV R^2 were 0.87, 0.82, 0.60 and 0.50 respectively. For 120 training sites, adjusted model, LOOCV, HV R^2 and MSE-HV R^2 were 0.82, 0.80, 0.73 and 0.62 respectively.

In our sensitivity analyses using fixed training set sizes or fixed test set sizes, we obtained results which were very comparable to those found in our main analyses (Figure S5 and Table S5).

For some models, especially the models based on small number of training sites, the ranges of values for some independent variables in the model were smaller than the ranges in values of these variables in the test datasets. As a result, predictions for such sites in the test data set may become unrealistic, especially when predictors such as inverse distance are non-linear. Two possible solutions to this are to remove the site entirely, or to recode the value of the predictor variable to the upper or lower limit of the range observed in the training sites. Figure 2 and Table 4 compare the R^2 and the regression slopes of hold-out validations before and after this range restriction. This procedure improved the HV and MSE HV R^2 by 3% and 9% respectively in small sets but the improvement decreased or vanished in large sets. The regression coefficients of observed vs. predicted concentrations, however, became much closer to unity showing that the models not only became

Table 2 Comparison of model adjusted model R^2, leave one out cross validation (LOOCV) R^2 and holdout (HV) R^2 in relation to size of training dataset in the whole Netherlands and Middle and West region (with stratified selection by site type)

Region	Training Sites (n)	Mean (SD)			
		Model	LOOCV	HV	MSE-HV
NL[a]	24	0.87 (0.05)	0.83 (0.06)	0.60 (0.08)	0.51(0.20)
	36	0.86 (0.04)	0.82 (0.05)	0.67 (0.07)	0.62(0.11)
	48	0.85 (0.03)	0.81 (0.04)	0.67 (0.07)	0.63(0.12)
	72	0.84 (0.03)	0.81 (0.04)	0.71 (0.05)	0.68(0.07)
	96	0.82 (0.02)	0.80 (0.03)	0.73 (0.06)	0.66(0.15)
	108	0.82 (0.02)	0.80 (0.02)	0.72 (0.08)	0.66(0.14)
	120	0.82 (0.02)	0.79 (0.02)	0.74 (0.08)	0.65(0.18)
W+M[b]	20	0.89 (0.05)	0.85 (0.08)	0.60 (0.10)	0.51(0.22)
	40	0.86 (0.03)	0.83 (0.04)	0.71 (0.06)	0.64(0.10)

NL[a]: the whole dataset of the Netherlands;

W+M[b]: West and Middle region.

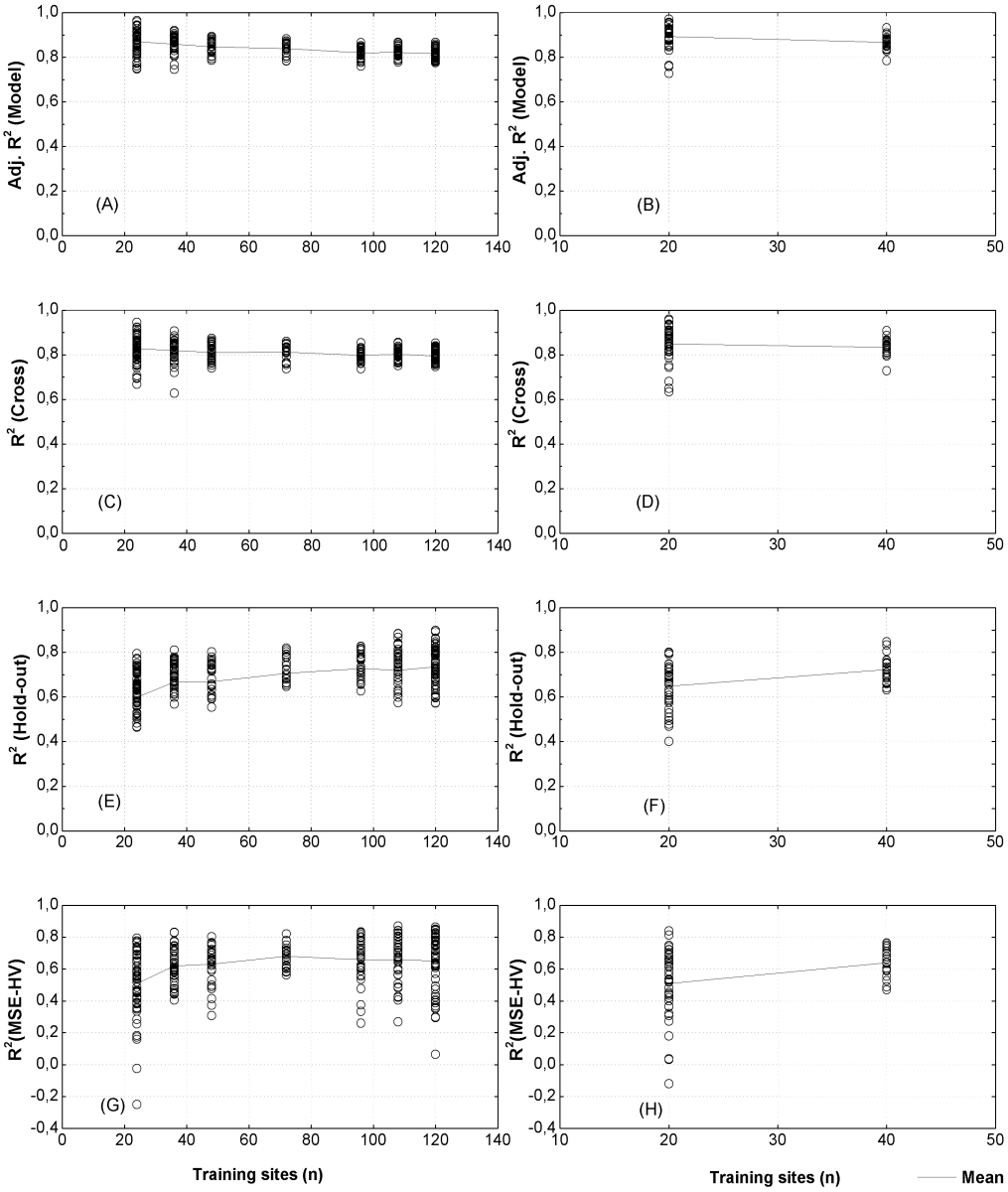


Figure 1 Model adjusted, LOOCV, HV and MSE-HV R^2 with increasing stratified selected training sites in the whole Netherlands(A, C, E, G) as well as in the small region (middle+west area)(B, D, F, H). The black lines show the trends of the plots which are connected by mean values in training groups.

more precise but also more accurate. Figure 3 shows residual plots with and without range restriction and it illustrates that range restriction removed the large residuals, particularly for models based on small numbers of training sites. Similar results were found in the analysis of sites in the west and middle of the country.

Table 3 Mean and standard deviation of RMSE in LOOCV and HV with increasing stratified selected training sites in the whole Netherlandsas well as in the smaller region (Middle+West area)

Region	Training Sites (n)	MEAN (SD)	
		LOOCV	HV
NL[a]	24	3.70(0.63)	5.37(0.62)
	36	3.78(0.43)	4.87(0.42)
	48	3.89(0.34)	4.86(0.49)
	72	3.86(0.30)	4.63(0.43)
	96	4.01(0.25)	4.51(0.47)
	108	3.97(0.19)	4.55(0.66)
	120	4.04(0.21)	4.53(0.68)
W+M[b]	20	3.13(0.85)	4.81(0.71)
	40	3.34(0.49)	4.29(0.45)

NL[a]: the whole dataset of the Netherlands;
W+M[b]: West and Middle region.

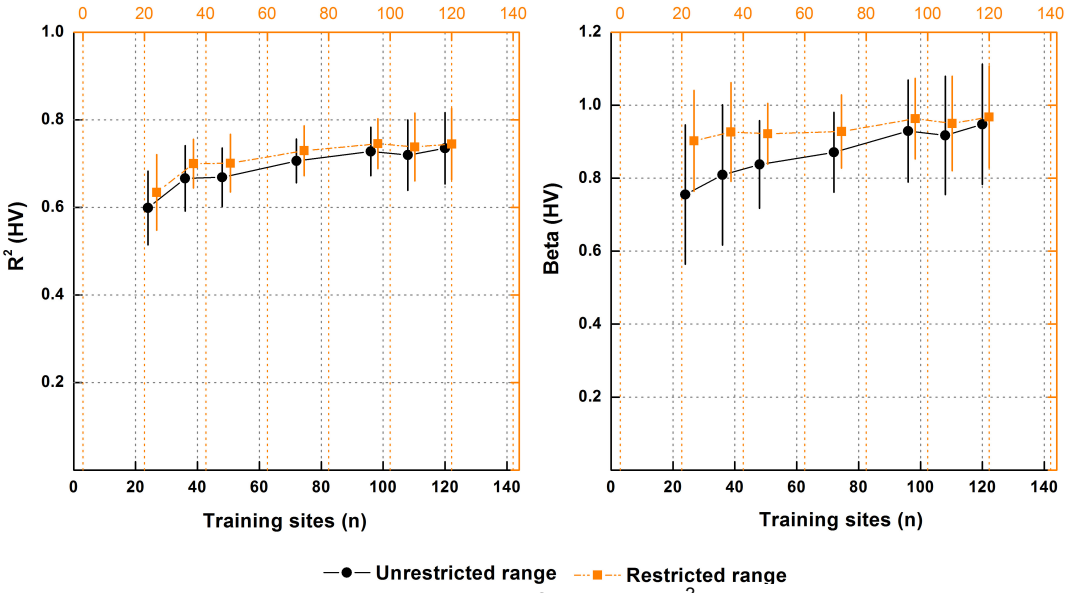


Figure 2 Mean and standard deviations of the HV R^2 and corresponding regression coefficients of association between measured and modeled concentration with increasing number of training sites using stratified site selections, and unrestricted and restricted ranges for predictor values at test sites.

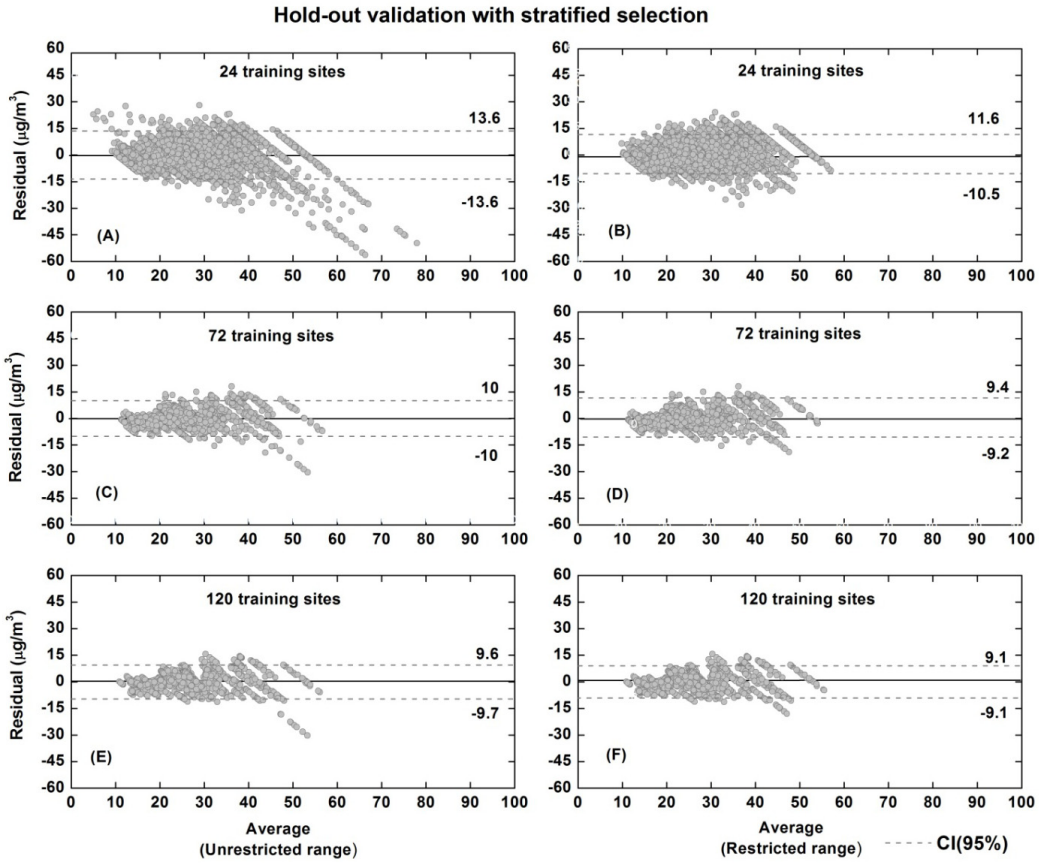


Figure 3 Variations of model residuals (predicted – observed NO_2 concentrations) as a function of averaged NO_2 concentrations ((predicted + observed NO_2 concentration) /2) before (A, C, E) and after (B, D, F) range restrictions in hold-out validation with stratified site selection. The dashed lines show the 95% Confidence Intervals of the residuals.

Table 4 HV R^2 and corresponding model regression coefficients in the Netherlands and West and Middle region using stratified site selections after range restriction

Region	Training sites (n)	Beta[c]	Beta[d]	Mean (SD) HV $R^{2\,d}$	MSE-HV $R^{2\,d}$
NL[a]	24	0.75 (0.19)	0.90 (0.14)	0.63 (0.09)	0.60 (0.21)
	36	0.80 (0.19)	0.93 (0.14)	0.70 (0.06)	0.67 (0.12)
	48	0.84 (0.12)	0.92 (0.08)	0.70 (0.07)	0.64 (0.12)
	72	0.87 (0.11)	0.93 (0.10)	0.73 (0.06)	0.68 (0.06)
	96	0.92 (0.14)	0.96 (0.11)	0.75 (0.06)	0.66 (0.15)
	108	0.92 (0.16)	0.95 (0.13)	0.74 (0.08)	0.66 (0.14)
	120	0.95 (0.17)	0.96 (0.13)	0.74 (0.08)	0.65 (0.18)
W+M[b]	20	0.74 (0.17)	0.92 (0.14)	0.65 (0.10)	0.63(0.31)
	40	0.84 (0.13)	0.96 (0.10)	0.72 (0.06)	0.67(0.11)

NL[a]: the whole dataset of the Netherlands; W+M[b]: West and Middle region;
Beta[c]: Regression coefficient of the measured and predicted data before range restriction;
Beta[d]: Regression coefficient of the measured and predicted data after range restriction;
HV $R^{2\,d}$: HV R^2 after the range restriction.
MSE HV $R^{2\,d}$: MSE HV R^2 after the range restriction

Discussion

In this study, we evaluated the performance of land-use regression models using internal cross-validation and independent hold-out validation and investigated the impact of increasing numbers of training sites. Hold-out validation R^2 increased as expected with the number of training sites and was lower on average than the corresponding cross-validation R^2, especially for models based on the smallest number of observations. Evaluation against fully independent datasets showed better results with increasing number of training sites, but the improvements were small beyond models based on 36 or 48 training sites. Constraining ranges of independent variables in the test datasets to the ranges observed for these variables in the training datasets improved the precision and especially the accuracy of the models.

The model predictive power and LOOCV R^2 were highest with the smallest training sets, decreasing only slightly by 4-5% towards the largest training sets. Other studies found similar, relatively small differences between model R^2 and LOOCV R^2 [13, 18, 19]. However, based on our results, the use of LOOCV is overly optimistic to estimate model predictive power especially when the number of training sites is relatively small. No studies made a systematic comparison of model, LOOCV and HV R^2 in relation to the size of the training sets for model development based on empirical data. Our work was inspired by a similar analysis by Johnson et al.[12]. Their work found larger differences in adjusted model R^2 going from small (n=25; model adjusted R^2 = 0.79) to large training sets (n=285; model adjusted R^2=0.63). The variation of hold-out validation R^2 was even more dramatic, changing from 0.28 with 25 training sites to 0.63 with 285 training sites. However, input concentrations were produced by a dispersion model producing average concentrations by census block, and not from measurements at carefully selected sites such as in our study. In a smaller empirical study, Madsen et al. [9] used training sets of 20 and 40 sites for model-building, and independent validation sets of 29-60 sites and found adjusted model R^2 of 0.77 and HV R^2 of 0.71-0.72 for measurements conducted in 2005, n=80; and adjusted model R^2 of 0.74, HV R^2 of 0.64-0.67 for measurements conducted in 2008, n=69.

The larger model adjusted R^2 combined with lower HV R^2 in the smaller datasets is likely explained by the phenomenon of over fitting[20, 21]. With many candidate predictors, and relatively small numbers of training sites, there is a recognized risk of over fitting models with artificially high model R^2 and relatively poor predictive power for independent validation data sets[20, 21]. The problem is aggravated with automatic model selection methods as the number of degrees of freedom is not properly characterized by the final model[21]. Our modeling approach consisted of offering 76 predictor variables using a supervised modeling approach, which may explain the modest change of R^2 with increasing sampling size. We restricted selection of predictors to a few with plausible coefficient signs and significant model improvement, based upon adjusted R^2 and not R^2. We fixed the shape of the relationship to linear (most variables) or another a priori shape (e.g. inverse distance). Our results suggest that model and evaluation R^2 did not depend on the number of predictors in the models (table S2 in supplement). Our results do

suggest, however, that the models based on the smallest size training sets were less successful in predicting concentrations at independent test sites than models based on larger training sets. Several investigators such as Bayak[21] suggest as a rule of thumb to have at least 10 observations per predictor in the model. Particularly models based on small training sets did not meet this criterion. However, in a sensitivity analysis restricting the number of predictor variables to 2 or 3 for models based on 24 training sites and 3 or 4 for models based on 36 training sites (Figure S6 and Table S6), we found that as the number of variables increased, not only the model adjusted R^2 and LOOCV R^2 increased (as expected), but the HV and MSE HV R^2 increased as well. This suggest that in the models developed in our main analyses, there was no over fitting compared to more parsimonious models. This is likely the result of paying careful attention to evaluating variables only when based on prior knowledge, a contribution is expected, and to include it only when the contribution is in the expected direction. Also, in choosing our monitoring sites, we paid careful attention to covering as much as possible the expected ranges of potential predictor variables in the study populations for which the models were being developed.

Especially when using small training sets, many of the evaluated predictor variables were selected at least once into a model. This is probably due to the fact that predictor variables are often correlated (e.g., traffic densities in buffers of varying sizes). Evaluation of regression coefficients is not very useful in multiple regression with correlated predictors. In the end we are interested in the capacity of the models to precisely and accurately predict concentrations measured at independent test sites.

Two approaches have been used for hold-out validation in previous studies. Some studies have set aside a random or stratified selection of the study sites measured and selected with the same procedure for model testing[9, 10, 11], whereas some studies have used other datasets using different monitoring methods and site selections[22, 23]. The latter studies can be useful if training set sizes were small. In principal, the selected test sites should be equally representative for the purpose of application of the LUR model. If the goal is assessment of air pollution exposures at the residential address, sites near the façade of homes are more useful than hotspots located at kerbsides of the busiest roads. We are aware of one example where a model developed from residential address-type sites did less well in predicting concentrations measured at kerbsides for regulatory purposes [22].

The observation that variability in holdout R^2 across models increased for the large training sets and thus small test sets, points to the importance of a sufficiently large test set. This may explain the finding in the original SAVIAH study of higher hold-out validation R^2 in test sets including 8-10 sites from regulatory monitoring networks than in the model development training sets[11]. Our analysis suggests that one should interpret these evaluation results with caution.

The purpose of any decent site selection procedure for monitoring to support LUR model development will always be to represent the population to which the model is going to be applied. With the usually limited number of monitoring sites that can realistically be included, it is inevitable that in the usually much larger study population, the home addresses of some subjects will have out-of-range predictor

variable values, resulting in unrealistic predictions. The observation that restricting the model predictions to the range of predictor variables observed in the training dataset improved hold-out validation results both in terms of precision and accuracy has implications for the application of LUR models. In principle, there are three ways in which this information can be used when estimating concentrations at home addresses of study populations: (1) by ignoring it completely – this will produce unrealistically high or low predicted values for at least some of the addresses, which in turn may influence subsequent epidemiological analyses; (2) by recoding the out-of-range values for these predictor values to the highest or lowest observed value in the model training set, as in our exercise – this will produce more realistic, albeit somewhat biased predictions for these addresses which are unlikely to affect epidemiological analyses much; (3) by removing all addresses with such out-of-range predictor values from further analyses to avoid any bias. A reasonable suggestion is to do both (2) and (3) and compare results. In most applications so far (including several of our own), we suspect that (1) has been applied, and we are not aware of published systematic comparisons of (2) and (3).

In our study, we did not observe much difference in the results obtained by stratified and random site selection. We anticipated that with random site selection, evaluation results would be poorer as some training / test datasets could have smaller contrast in (traffic) predictor variables, resulting in less stable models. It should, however, be noted that the full dataset was not a random sample, but a sample in which traffic locations were over represented. Especially in the larger training sets, one would expect random selection to more or less represent the predefined site categorization which is maintained in the stratified selection.

We found little difference in model explanatory power when comparing the regional model to the country wide model. Also in this smaller area, model and cross-validation R^2 were smaller for the n=40 sites compared to the n=20 sites whereas HV R^2 was higher for the n=40 sites. In both the national and the regional analysis, the largest gain in hold out validation occurred between 24 (20) and 36 (40) observations in the training dataset. The R^2 and LOOCV R^2 of especially models based upon fewer than 40 sites must be interpreted carefully. This does not imply that these models are unreliable. Even for a model based upon 24 sites across the Netherlands, a hold-out validation R^2 of 0.60 was obtained. The quantitative results of this study may not apply to other study areas, depending e.g on the complexity of the area with respect to sources and geography.

In summary, LUR model performance for NO_2 varies with the number of training sites. Hold-out validation R^2 was lower than the corresponding cross-validation R^2, especially for the smallest training sets. Truly independent evaluation data are especially useful when LUR models are developed from small training sets where we have shown the adjusted model and LOOCV R^2s to deviate most from the hold out validation R^2s. In our specific application, models based on as few as 24 training sites, however, achieved acceptable hold out validation R^2s of, on average, 0.60.

References

1. Brunekreef, B.; Holgate, S. T., Air pollution and health. Lancet 2002, 360 (9341), 1233-1242.
2. Pope, C. A.; Dockery, D. W., Health effects of fine particulate air pollution: Lines that connect. J. Air Waste Manage. Assoc. 2006, 56 (6), 709-742.
3. Bukowiecki, N.; Dommen, J.; Prevot, A. S. H.; Richter, R.; Weingartner, E.; Baltensperger, U., A mobile pollutant measurement laboratory-measuring gas phase and aerosol ambient concentrations with high spatial and temporal resolution. Atmos. Environ. 2002, 36 (36-37), 5569-5579.
4. Weijers, E. P.; Khlystov, A. Y.; Kos, G. P. A.; Erisman, J. W., Variability of particulate matter concentrations along roads and motorways determined by a moving measurement unit. Atmos. Environ. 2004, 38 (19), 2993-3002.
5. Westerdahl, D.; Fruin, S.; Sax, T.; Fine, P. M.; Sioutas, C., Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles. Atmos. Environ. 2005, 39 (20), 3597-3610.
6. Briggs, D., The role of GIS: coping with space (and time) in air pollution exposure assessment. J Toxic. Environ. Health A 2005, 68 (13-14), 1243-61.
7. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; Morrison, J.; et al., A review and evaluation of intraurban air pollution exposure models. J. Expo. Anal. Environ. Epidemiol. 2005, 15 (2), 185-204.
8. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; et al., A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 2008, 42 (33), 7561-7578.
9. Madsen, C.; Gehring, U.; Haberg, S. E.; Nafstad, P.; Meliefste, K.; Nystad, W.; et al., Comparison of land-use regression models for predicting spatial NOx contrasts over a three year period in Oslo, Norway. Atmos. Environ. 2011, 45 (21), 3576-3583.
10. Beelen, R.; Hoek, G.; Pebesma, E.; Vienneau, D.; de Hoogh, K.; Briggs, D. J., Mapping of background air pollution at a fine spatial scale across the European Union. Sci. Total Environ. 2009, 407 (6), 1852-1867.
11. Briggs, D. J.; Collins, S.; Elliott, P.; Fischer, P.; Kingham, S.; Lebret, E.; et al., Mapping urban air pollution using GIS: a regression-based approach. Int. J. Geogr. Inf. Sci. 1997, 11 (7), 699-718.
12. Johnson, M.; Isakov, V.; Touma, J. S.; Mukerjee, S.; Ozkaynak, H., Evaluation of land-use regression models used to predict air quality concentrations in an urban area. Atmos. Environ. 2010, 44 (30), 3660-3668.
13. Eeftens, M.; Beelen, R.; Fischer, P.; Brunekreef, B.; Meliefste, K.; Hoek, G., Stability of measured and modelled spatial contrasts in NO_2 over time. Occup. Environ. Med. 2011, 68 (10), 765-770.
14. Beelen, R.; Hoek, G.; Fischer, P.; van den Brandt, P. A.; Brunekreef, B., Estimated long-term outdoor air pollution concentrations in a cohort study. Atmos. Environ. 2007, 41 (7), 1343-1358.
15. Hastie, T., Tibshirani, R. and Friedman, J.H. The elements of statistical learning : data mining, inference, and prediction. Springer series in statistics. Springer, New York, 2001, pp 214.
16. Harrell, F.E.,. Regression modeling strategies : with applications to linear models, logistic regression, and survival analysis. Springer series in statistics. Springer, New York, 2001, xxii, 568 p. pp.
17. Szpiro, A.A., Paciorek, C.J. and Sheppard, L.,. Does more accurate exposure prediction necessarily improve health effect estimates? Epidemiology, 2011, 22 (5), 680-685.
18. Vienneau, D.; de Hoogh, K.; Beelen, R.; Fischer, P.; Hoek, G.; Briggs, D., Comparison of land-use regression models between Great Britain and the Netherlands. Atmos. Environ. 2010, 44 (5), 688-696.
19. Novotny, E. V.; Bechle, M.; Millet, D. B.; Marshall, J. D., National Satellite-based Land Use Regression: NO_2 in the United States. Environ. Sci. Technol. 2011, 45 (10), 4407-4414.
20. Craig, M. H.; Sharp, B. L.; Mabaso, M. L. H.; Kleinschmidt, I., Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure. Int.l J. Health Geogr. 2007, 6.

21. Babyka, M. A., What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. Psychos. Med. 2004, 66, 411 - 421.
22. Dijkema, M. B.; Gehring, U.; van Strien, R. T.; van der Zee, S. C.; Fischer, P.; Hoek, G.; et al., A comparison of different approaches to estimate small-scale spatial variation in outdoor NO concentrations. Environ. Health Perspect. 2011, 119 (5), 670-5.
23. Stedman, J. R.; Vincent, K. J.; Campbell, G. W.; Goodwin, J. W. L.; Downing, C. E. H., New high resolution maps of estimated background ambient NOx and NO_2 concentrations in the UK. Atmos. Environ. 1997, 31 (21), 3591-3602.

Chapter 2

Appendix

(Online: http://pubs.acs.org/doi/suppl/10.1021/es204183v)

Description of the stratification of the sites in training/test sites selections

The monitoring sites have been selected in three types of locations, regional background, urban background, and traffic, expecting relatively low values in regional background sites and high values in traffic sites. In addition, sites were distributed throughout the Netherlands. After the monitoring, sites were categorized into four regions, North, Middle, West, and South areas according to the distributions of cities and spatial distributions of concentrations, for instance, high concentrations were in the Middle and West regions and lower concentrations in the North area. Therefore, for this evaluation, training sites were selected based on two strata: region and site type. First, the number of training sites in each region was chosen to be in proportion to the overall distribution of sites across regions. Second, the number of training sites in each site type in each region was chosen to be in proportion to overall distribution of sites across types within each region. Therefore, the required number of sites in each type in each region was calculated.

As shown in Figure S2a, a total of 144 sites were equally divided into four fractions with each subset of 36 sites. Each subset was alternately picked up as training dataset for modeling and the remaining 108 sites were treated for external prediction. Therefore, four models were developed and performed for evaluations. Site selections were performed ten times to be able to study the sampling variability of our procedures which produced 40 training datasets. Subsequently, we exchanged the training dataset (n=108) and the test dataset (n=36) for the same model constructions and validations (Figure S2b). Similar produces were conducted for the two, three and six divisions with each training sites of 72, 48 and 24 sites. In total, 60 models were produced by 24 training sites, following 40 models by 36 training sites, 30 models by 48 training sites, 20 models by 72 training sites, 30 models by 96 training sites, 40 models by 108 training sites and 60 models by 120 training sites. Therefore, a total of 280 models were generated in the final stage for the whole of the Netherlands. Note that after removal of one site, 143 sites remained. As a result, in each individual model evaluation the training set OR the test set contained one less site, depending on whether the excluded site was allocated to the training set or the test set. .

Figure S3 and Table S3 together show the frequency distributions of variables that appear in the models in different training sets. More variables were selected into models based on a small number of training sites than into models based on larger numbers of training sites. For instance, 69 variables appear at least once in models based on 24 training sites while 24 variables appear at least once in models based on 120 training sites. The most frequent variables in these models were X+Y coordinate (100%), population density in 5000 meters buffer (95%), Industry in 500 (66%) meters buffer, Inverse distance to major road (50%).

As a sensitivity analysis, test sets were fixed to 72 sites, and the remaining sites were used for model building using training sets of 24, 36, 48 and 72 sites. The test sets and training sets were randomly selected thirty times and therefore 120 models were built. As shown in Figure S4 and Table S4, we found that for same test sets, model adjusted R^2, LOOCV R^2 decreased as the number of training sites increased. The median HV and MSE-HV R^2 increased from 0.65 to 0.74 and 0.53 to 0.67 respectively with increasing numbers of training sites. These results were completely in line with the results of our main analyses which used varying rather than fixed sizes for training and test sets respectively.

As a further sensitivity analysis, models were built based on randomly selected fixed training sets (N=24,36,48,72). For each model, four test sets (N=24, 36, 48, 72) were evaluated which were randomly selected from the remaining sites of the 144 sites. This work has been repeated for thirty times and therefore generated 120 models (30 of each

training set groups) . As shown in Figure S5 and Table S5, for a fixed size of the training sets, the median HV and MSE-HV R^2 did not vary much with the number of test sites. For instance, the median HV and MSE-HV R^2 were 0.60, 0.58, 0.61, 0.64; and 0.57, 0.51, 0.55, 0.55 for 24, 36, 48 and 72 test sites used for evaluation of models based on 24 training sites.

As another sensitivity analysis, we built models based on 24 and 36 training sites. We restricted the maximum number of variables to 2 and 3 for 24 training sites, and 3 and 4 for 36 training sites respectively and compared results with those of our main analyses in which we did not restrict the number of variables allowed to be selected into the models. Figure S6 and Table S6 show the distributions of model adjusted R^2, the LOOCV, HV and MSE-HV R^2 based on fixed number of variables. In total, 100 data sets (60 for N=24, 40 for N=36) were used for modeling. The MSE-based HV R^2 was calculated according to formula (1). When restricting the models based on training sets of 24 sites to 2 predictor variables, model, LOOCV, HV as well as MSE-HV R^2 were clearly lower than when using 3 or more variables. The difference between models using 3, and models using 3 or more variables were much less. There was no indication that HV and MSE_HV R^2 were improved when restricting the number of prediction variables, suggesting that models based on larger numbers of prediction variables did not do worse, in terms of precision and accuracy of predictions, than more parsimonious models. When using 36 training sites, differences between models based on 3, 4 or more predictor variables were fairly small.

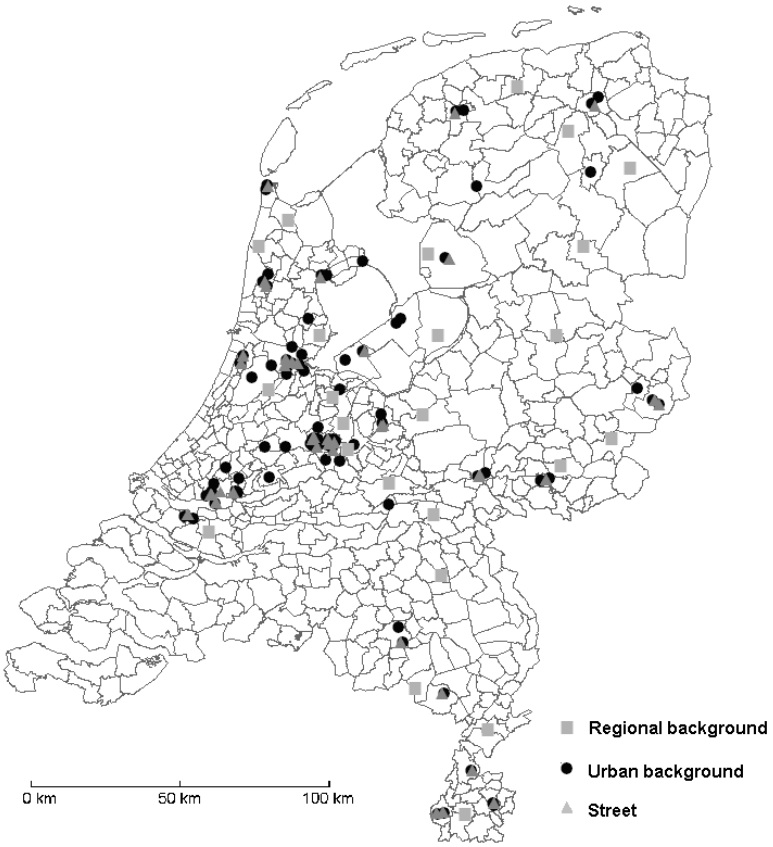


Figure S1 Map of TRACHEA monitoring sites in the Netherlands

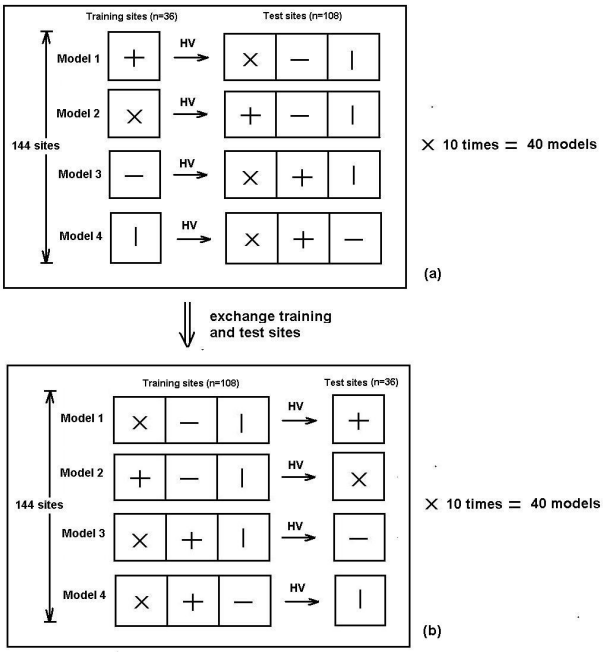


Figure S2 Allocation scheme of monitoring sites for training and test datasets. The boxes with the same mark in different models show the same data sets. The example is for models based on 36 training sites and 108 test sites, and for 108 training sites and 36 test sites respectively.

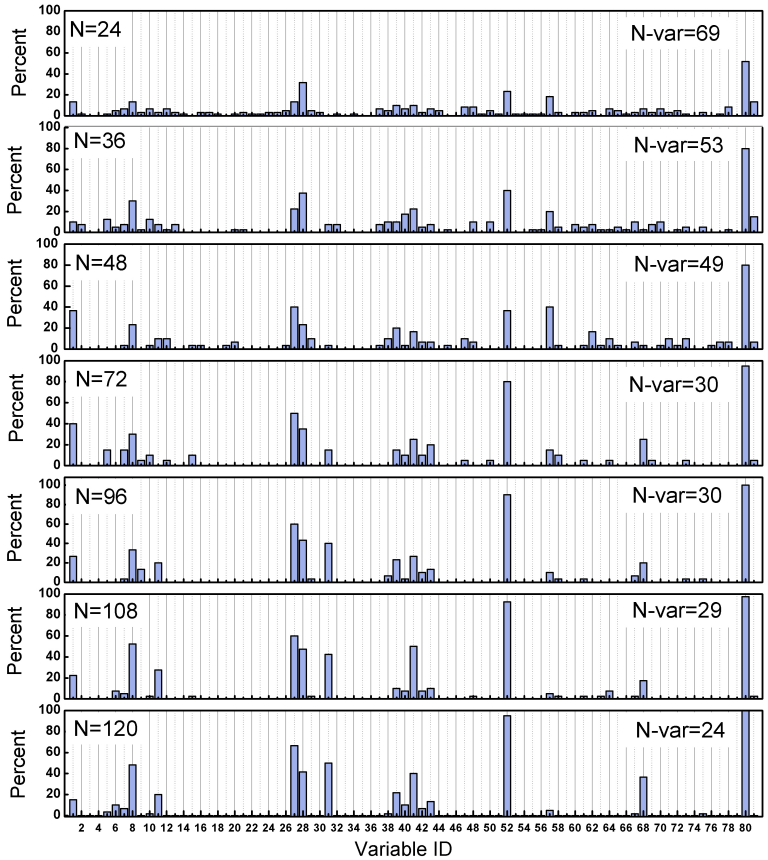


Figure S3 Percent of times a variable was selected into a final model. The corresponding variable names to the ID can be seen in table S3. The N-var shows the number of unique variables that were selected into models based on training sets of sizes increasing from 24 to 120.

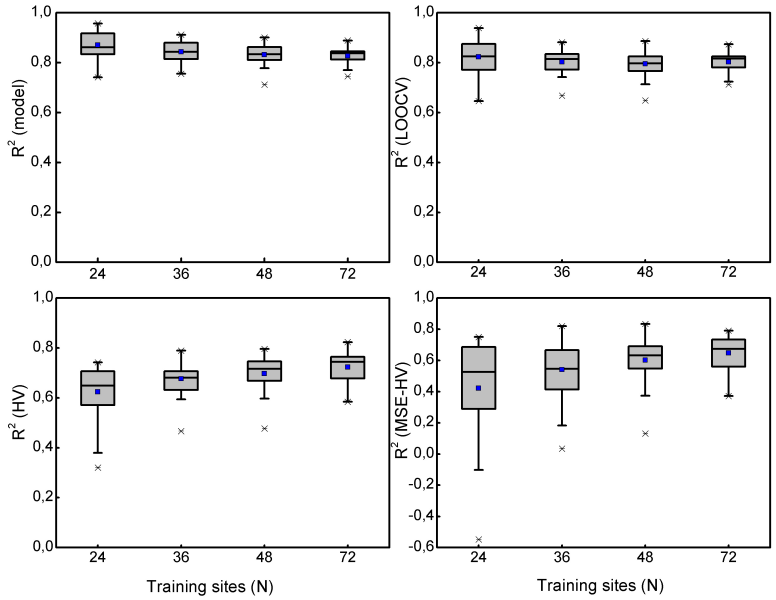


Figure S4 Box plots of model adjusted R^2 and LOOCV R^2 with increasing number of training sites. The HV and MSE-HV R^2 were based on fixed sets of 72 test sites which were randomly selected thirty times

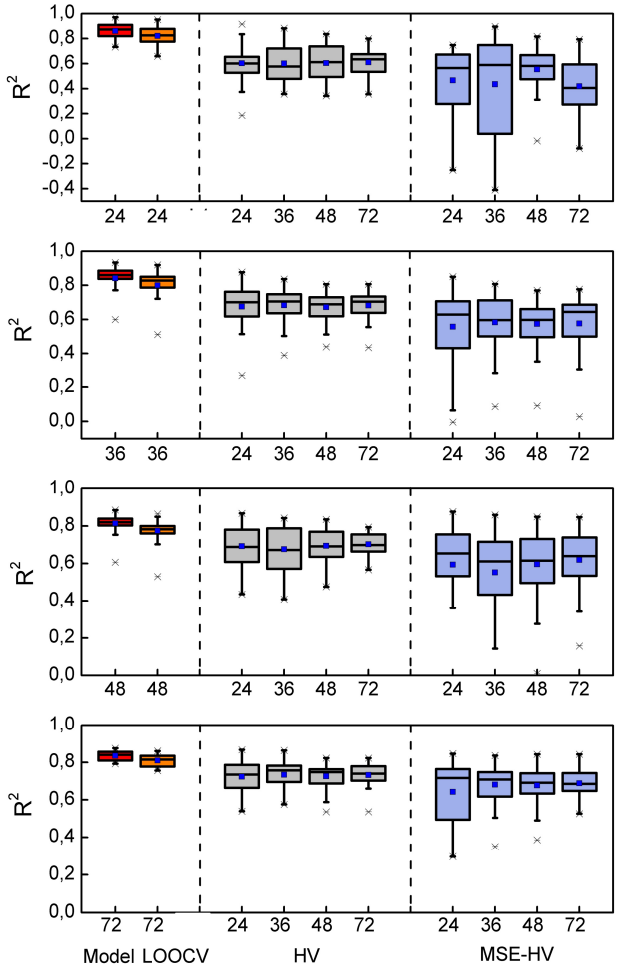


Figure S5 Box plots of model adjusted R^2 and LOOCV R^2 based on fixed number of training sites, and a varying number of test sites

Table S1 Potential predictor variables for LUR model development with specified buffer sizes and a priori defined directions of effect

GIS dataset	Predictor variables	Unit	Buffer size(m)	Direction
CORINE	Low density residential	m^2	100, 300, 500, 1000, 5000	+
CORINE	Industry	m^2	500, 1000, 5000	+
CORINE	Urban green+ Semi natural and forested areas	m^2	300, 500, 1000, 5000	-
Population density	Number of inhabitants	n	100, 300, 500, 1000, 5000	+
Household density	Number of households	n	100, 300, 500, 1000, 5000	+
-	Coordinates	m	N/A	N/A
Local road network	Traffic intensity on nearest road	veh/day	N/A	+
Local road network	Distance to the nearest road (inverse distance and inverse distance squared)	m^{-1},m^{-2}	N/A	+
Local road network	Product of traffic intensity on nearest road and inverse of distance to the nearest road and distance squared	veh/day/m veh/day/m^2	N/A	+
Local road network	Traffic intensity on nearest major road	veh/day	N/A	+
Local road network	Product of traffic intensity on nearest major road and inverse of distance to the nearest major road and distance squared	veh/day/m veh/day/m^2	N/A	+
Local road network	Total traffic load of major roads in a buffer (sum of (traffic intensity * length of all segments))	veh/day*m	25, 50, 100, 300, 500, 1000	+
Local road network	Total traffic load of all roads in a buffer (sum of (traffic intensity * length of all segments))	veh/day*m	25, 50, 100, 300, 500, 1000	+
Local road network	Heavy-duty traffic intensity on nearest road	veh/day	N/A	+
Local road network	Heavy-duty traffic intensity on nearest major road	veh/day	N/A	+
Local road network	Total heavy-duty traffic load of major roads in a buffer (sum of (heavy-duty traffic intensity * length of all segments))	veh/day*m	25, 50, 100, 300, 500, 1000	+
Local road network	Total heavy-duty traffic load of all roads in a buffer (sum of (heavy-duty traffic intensity * length of all segments))	veh/day*m	25, 50, 100, 300, 500, 1000	+
Central road network	Road length of all roads in a buffer	m	25, 50, 100, 300, 500, 1000	+
Central road network	Road length of major roads in a buffer	m	25, 50, 100, 300, 500, 1000	+

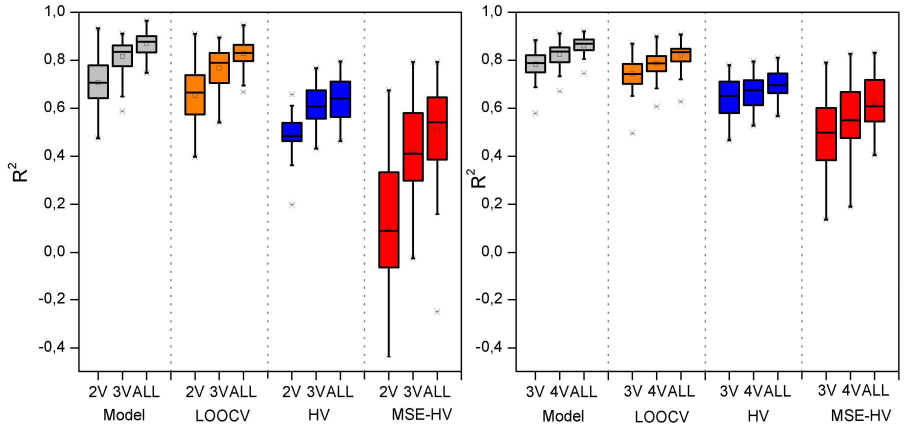


Figure S6 Comparison of the model adjusted R^2, the LOOCV, HV and MSE-HV R^2 based on fixed number of variables (N=24, N variable=2, 3; N=36, N variable=3, 4; ALL, no variable restriction)

Table S2 Summary statistics of models in relation to number of training sites and number of predictor variables included in the models. For instance, 2 models contain 2 variables respectively based on 24 training sites with average adjusted model R^2 and LOOCV, HV R^2 of 0.83, 0.81 and 0.56.

Training sites (N)	Variables (N)	Models (N)	Adjusted R^2	LOOCV R^2	HV R^2
24	2	2	0.83	0.81	0.56
	3	12	0.85	0.80	0.60
	4	25	0.85	0.81	0.60
	5	13	0.90	0.86	0.61
	6	8	0.92	0.88	0.59
36	3	1	0.77	0.74	0.70
	4	3	0.86	0.83	0.65
	5	23	0.86	0.82	0.68
	6	7	0.86	0.82	0.64
	7	6	0.87	0.83	0.66
48	4	4	0.84	0.81	0.61
	5	15	0.85	0.81	0.67
	6	8	0.86	0.82	0.68
	7	3	0.83	0.79	0.70
72	4	1	0.78	0.76	0.81
	5	9	0.85	0.82	0.69
	6	6	0.84	0.80	0.70
	7	3	0.83	0.80	0.73
	8	1	0.86	0.83	0.73
96	4	2	0.79	0.78	0.72
	5	6	0.81	0.79	0.70
	6	19	0.82	0.80	0.74
	7	2	0.83	0.81	0.78
	8	1	0.85	0.83	0.70
108	4	2	0.80	0.78	0.66
	5	11	0.82	0.80	0.71
	6	16	0.82	0.80	0.74
	7	8	0.83	0.81	0.72
	8	3	0.84	0.81	0.69
120	4	3	0.79	0.77	0.68
	5	15	0.81	0.79	0.73
	6	23	0.81	0.79	0.75
	7	17	0.82	0.80	0.74
	8	2	0.85	0.84	0.68

Table S3 Variable names in correspondence to variable ID

ID	Variable Names	ID	Variable Names
1	DISTINVMAJOR1	42	MAJORROADLENGTH_50
2	DISTINVMAJOR2	43	MAJORROADLENGTH_500
3	DISTINVNEAR1	44	NATURAL_GREEN_1000
4	DISTINVNEAR2	45	NATURAL_GREEN_300
5	HEAVYINTINVDIST	46	NATURAL_GREEN_500
6	HEAVYINTINVDIST2	47	NATURAL_GREEN_5000
7	HEAVYTRAFLOAD_100	48	POPEEA_100
8	HEAVYTRAFLOAD_1000	49	POPEEA_1000
9	HEAVYTRAFLOAD_25	50	POPEEA_300
10	HEAVYTRAFLOAD_300	51	POPEEA_500
11	HEAVYTRAFLOAD_50	52	POPEEA_5000
12	HEAVYTRAFLOAD_500	53	POP_100
13	HEAVYTRAFMAJOR	54	POP_1000
14	HEAVYTRAFMAJORLOAD_100	55	POP_300
15	HEAVYTRAFMAJORLOAD_1000	56	POP_500
16	HEAVYTRAFMAJORLOAD_25	57	POP_5000
17	HEAVYTRAFMAJORLOAD_300	58	ROADLENGTH_100
18	HEAVYTRAFMAJORLOAD_50	59	ROADLENGTH_1000
19	HEAVYTRAFMAJORLOAD_500	60	ROADLENGTH_25
20	HEAVYTRAFNEAR	61	ROADLENGTH_300
21	HHOLD_100	62	ROADLENGTH_50
22	HHOLD_1000	63	ROADLENGTH_500
23	HHOLD_300	64	TRAFLOAD_100
24	HHOLD_500	65	TRAFLOAD_1000
25	HHOLD_5000	66	TRAFLOAD_25
26	INDUSTRY_1000	67	TRAFLOAD_300
27	INDUSTRY_500	68	TRAFLOAD_50
28	INDUSTRY_5000	69	TRAFLOAD_500
29	INTINVDIST	70	TRAFMAJOR
30	INTINVDIST2	71	TRAFMAJORLOAD_100
31	INTMAJORINVDIST	72	TRAFMAJORLOAD_1000
32	INTMAJORINVDIST2	73	TRAFMAJORLOAD_25
33	LDRES_100	74	TRAFMAJORLOAD_300
34	LDRES_1000	75	TRAFMAJORLOAD_50
35	LDRES_300	76	TRAFMAJORLOAD_500
36	LDRES_500	77	TRAFNEAR
37	LDRES_5000	78	xcoord
38	MAJORROADLENGTH_100	79	xminusy
39	MAJORROADLENGTH_1000	80	xplusy
40	MAJORROADLENGTH_25	81	ycoord
41	MAJORROADLENGTH_300		

Table S4 Distributions of model adjusted R^2, the LOOCV, HV and MSE-HV R^2 in prediction for 72 test sites with thirty iterations.

Train(N)	TYPE	Min	P25	Median	P75	Max
24	Model	0.74	0.83	0.86	0.92	0.96
24	LOOCV	0.65	0.77	0.83	0.88	0.94
24	HV	0.32	0.57	0.65	0.71	0.74
24	MSE-HV	-0.55	0.29	0.53	0.69	0.75
36	Model	0.76	0.81	0.84	0.88	0.91
36	LOOCV	0.67	0.77	0.82	0.84	0.88
36	HV	0.47	0.63	0.68	0.71	0.79
36	MSE-HV	0.03	0.41	0.55	0.67	0.82
48	Model	0.71	0.81	0.83	0.86	0.90
48	LOOCV	0.65	0.77	0.80	0.83	0.89
48	HV	0.48	0.67	0.72	0.75	0.80
48	MSE-HV	0.13	0.55	0.64	0.69	0.83
72	Model	0.75	0.81	0.84	0.85	0.89
72	LOOCV	0.71	0.78	0.82	0.83	0.87
72	HV	0.58	0.68	0.74	0.76	0.82
72	MSE-HV	0.37	0.56	0.67	0.73	0.79

Table S5 Distribution of model adjusted R^2, the LOOCV, HV and MSE-HV R^2 based on fixed size training sets and a varying number of test sites with thirty iterations.

Train(N)	Test(N)	TYPE	Min	P25	Median	P75	Max
24	-	Model	0.73	0.82	0.87	0.91	0.97
24	-	LOOCV	0.66	0.78	0.83	0.88	0.95
24	24	HV	0.19	0.53	0.60	0.65	0.92
24	24	MSE-HV	-0.91	0.31	0.57	0.71	0.89
24	36	HV	0.36	0.48	0.58	0.72	0.88
24	36	MSE-HV	-0.41	0.26	0.51	0.71	0.90
24	48	HV	0.34	0.49	0.61	0.74	0.84
24	48	MSE-HV	-0.15	0.33	0.55	0.67	0.89
24	72	HV	0.36	0.53	0.64	0.68	0.80
24	72	MSE-HV	-0.25	0.32	0.55	0.62	0.84
36	-	Model	0.60	0.84	0.86	0.89	0.93
36	-	LOOCV	0.51	0.79	0.83	0.85	0.92
36	24	HV	0.27	0.62	0.70	0.76	0.88
36	24	MSE-HV	0.00	0.43	0.64	0.71	0.85
36	36	HV	0.39	0.64	0.71	0.75	0.84
36	36	MSE-HV	0.09	0.50	0.60	0.71	0.81
36	48	HV	0.44	0.62	0.69	0.73	0.81
36	48	MSE-HV	0.09	0.50	0.61	0.66	0.77
36	72	HV	0.43	0.64	0.71	0.73	0.81
36	72	MSE-HV	0.03	0.50	0.65	0.69	0.78
48	-	Model	0.61	0.80	0.82	0.84	0.89
48	-	LOOCV	0.53	0.76	0.78	0.80	0.86
48	24	HV	0.44	0.61	0.69	0.78	0.87
48	24	MSE-HV	-0.07	0.53	0.66	0.76	0.88
48	36	HV	0.41	0.57	0.67	0.79	0.84
48	36	MSE-HV	-0.15	0.43	0.62	0.72	0.86
48	48	HV	0.48	0.64	0.69	0.77	0.83
48	48	MSE-HV	0.01	0.50	0.62	0.73	0.85
48	72	HV	0.57	0.66	0.70	0.76	0.79
48	72	MSE-HV	0.16	0.53	0.64	0.74	0.85
72	-	Model	0.79	0.81	0.84	0.86	0.88
72	-	LOOCV	0.76	0.78	0.81	0.84	0.86
72	24	HV	0.54	0.66	0.73	0.79	0.87
72	24	MSE-HV	0.30	0.49	0.72	0.76	0.85
72	36	HV	0.58	0.70	0.76	0.78	0.86
72	36	MSE-HV	0.35	0.62	0.71	0.75	0.84
72	48	HV	0.54	0.69	0.75	0.76	0.82
72	48	MSE-HV	0.39	0.63	0.69	0.74	0.84
72	72	HV	0.54	0.70	0.74	0.78	0.82
72	72	MSE-HV	0.53	0.65	0.69	0.74	0.84

Table S6 Distribution of model adjusted R^2, the LOOCV, HV and MSE-HV R^2 based on fixed number of variables (N=24, N variable=2, 3; N=36, N variable=3, 4)

Train(N)	varnum	TYPE	Min	P25	Median	P75	Max
24	2	Model	0.48	0.64	0.71	0.78	0.93
24	3	Model	0.59	0.78	0.84	0.86	0.91
24	All	Model	0.75	0.84	0.88	0.90	0.96
24	2	LOOCV	0.40	0.57	0.67	0.74	0.91
24	3	LOOCV	0.54	0.71	0.79	0.83	0.89
24	All	LOOCV	0.67	0.80	0.83	0.87	0.95
24	2	HV	0.20	0.46	0.48	0.54	0.66
24	3	HV	0.43	0.56	0.61	0.67	0.77
24	All	HV	0.46	0.57	0.64	0.71	0.79
24	2	MSE-HV	-1.62	-0.06	0.09	0.33	0.68
24	3	MSE-HV	-0.02	0.30	0.41	0.58	0.79
24	All	MSE-HV	-0.25	0.39	0.54	0.65	0.79
36	3	Model	0.58	0.75	0.79	0.82	0.88
36	4	Model	0.67	0.80	0.84	0.85	0.91
36	All	Model	0.75	0.84	0.87	0.89	0.92
36	3	LOOCV	0.50	0.70	0.74	0.79	0.87
36	4	LOOCV	0.61	0.75	0.79	0.82	0.90
36	All	LOOCV	0.63	0.80	0.83	0.85	0.91
36	3	HV	0.47	0.58	0.65	0.71	0.78
36	4	HV	0.53	0.62	0.68	0.72	0.80
36	All	HV	0.57	0.67	0.70	0.75	0.81
36	3	MSE-HV	0.14	0.39	0.50	0.61	0.79
36	4	MSE-HV	0.19	0.48	0.55	0.67	0.83
36	All	MSE-HV	0.40	0.55	0.61	0.72	0.83

Chapter 3

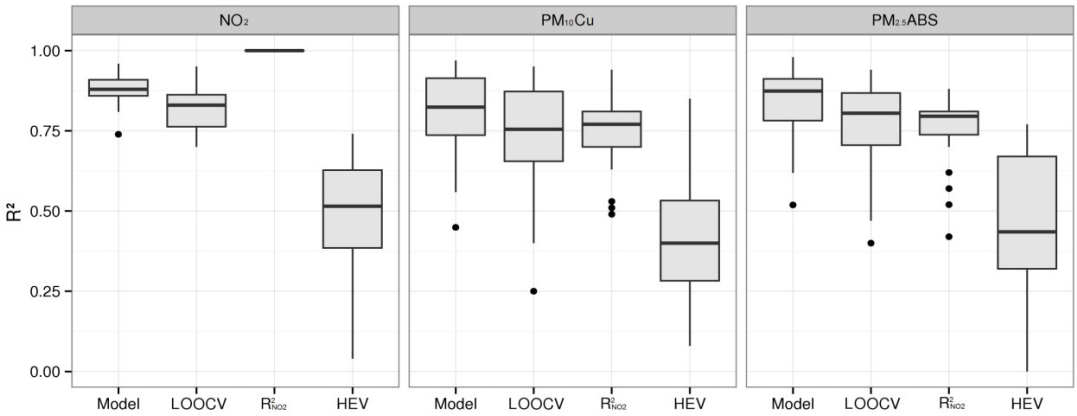
Evaluation of Land Use Regression Models for NO_2 and Particulate Matter in 20 European Study Areas: the ESCAPE Project

Meng Wang, Rob Beelen, Xavier Basagana, Thomas Becker, Giulia Cesaroni, Kees de Hoogh, Audrius Dedele, Christophe Declercq, Konstantina Dimakopoulou, Marloes Eeftens, Francesco Forastiere, Regina Grazuleviciene, Claudia Galassi, Barbara Hoffmann, Joachim Heinrich, Minas Iakovides, Nino Künzli, Michal Korek, Sarah Lindley, Anna Mölter, Gioia Mosler, Christian Madsen, Mark Nieuwenhuijsen, Harish Phuleria, Xanthi Pedeli, Ole Raaschou-Nielsen, Andrea Ranzi, Dorothee Sugiri, Morgane Stempfelet, Ming-Yi Tsai, Timo Lanki, Orsolya Udvardy, Mihaly J Varro, Kathrin Wolf, Gudrun Weinmay, Tarja Yli-Tuomi, Gerard Hoek, Bert Brunekreef

(Environmental Sciences & Technology 2013, 47, (9), 4357-4364)

Abstract

Land use regression models (LUR) frequently use leave-one-out-cross-validation (LOOCV) to assess model fit, but recent studies suggested that this may overestimate predictive ability in independent datasets. Our aim was to evaluate LUR models for nitrogen dioxide ($NO_{2)}$ and particulate matter (PM) components exploiting the high correlation between concentrations of PM metrics and NO_2. LUR models have been developed for NO_2, $PM_{2.5}$ absorbance and Copper (Cu) in PM_{10} based on 20 sites in each of the 20 study areas of the ESCAPE project. Models were evaluated with LOOCV and "hold-out evaluation (HEV)" using the correlation of predicted NO_2 or PM concentrations with measured NO_2 concentrations at the 20 additional NO_2 sites in each area. For NO_2, $PM_{2.5}$ absorbance and PM_{10} Cu, the median LOOCV R^2s were 0.83, 0.81 and 0.76 whereas the median HEV R^2 were 0.52, 0.44 and 0.40. There was a positive association between the LOOCV R^2 and HEV R^2 for $PM_{2.5}$ absorbance and PM_{10} Cu. Our results confirm that the predictive ability of LUR models based on relatively small training sets is overestimated by the LOOCV R^2s. Nevertheless, in most areas LUR models still explained a substantial fraction of the variation of concentrations measured at independent sites.



(Table of content)

Introduction

Epidemiological studies have suggested that long term exposure to air pollution is associated with adverse health effects[1-3]. Some of these studies have relied on estimating air pollution concentrations at the home addresses of study participants using Land Use Regression methods[4, 5]. Within the ESCAPE (European Study of Cohort for Air Pollution Effects) project, a comprehensive measurement program was conducted in 36 European study areas between 2008 and 2011. Substantial spatial variability of nitrogen oxide (NO_2, NO_x) and particulate matter (PM) was identified within and between these areas[6, 7]. To explain and predict within-area variability, land use regression (LUR) models were developed using a standardized approach[8].

Land use regression (LUR) modeling is a Geographic Information System (GIS) and statistics based method that exploits land use, geographic and traffic characteristics (e.g. traffic intensity, road length, population density) to explain spatial concentration variations at measured sites[9]. Within the ESCAPE project, PM and NO_2/NO_x models have been developed in 20 and 36 study areas respectively, using a standardized method[8, 10]. These models explained a large fraction of spatial variance in the measured pollution concentrations, as measured by R^2s ranging from 55~95% for NO_2 and for $PM_{2.5}$ absorbance.

Model evaluation is essential as the model R^2 may be artificially high[11]. Two common evaluation approaches are: the internal "leave-one-out-cross-validation (LOOCV)" and the external "hold-out-evaluation (HEV)" against independent measurements set aside for model evaluation. The HEV is preferable as it likely better reflects the predictive power of the model at locations where no measurements were taken, such as addresses of subjects in an epidemiological study, assuming that validation sites are representative of the distribution of subject's addresses. In a study with 144 NO_2 monitoring sites, we previously reported that the model adjusted R^2 decreased slightly from 0.87 to 0.82 with the increasing size of the training sets used for model development. In contrast, the HEV R^2 increased from 0.60 to 0.74 with training set size from 24 to 120[12]. This is likely due to some over-fitting[11]. Similar evaluations have been conducted in Girona, Spain and in Oslo, Norway with somewhat different results: in Girona, differences between LOOCV R^2 and HEV R^2 were larger than we found previously, in Oslo they were smaller[13, 14].

All these studies of LUR model performance evaluation were conducted for NO_2. Sampling of PM requires more effort and usually the number of sampling sites is not sufficient to allow for a separation into training and test dataset (for validation purpose) of sufficient size. To the best of our knowledge, no evaluations have been conducted for particulate matter LUR models.

Within the ESCAPE study area specific PM models were developed based on 20 training sites per area in most of the study areas[8]. In view of the recent model evaluation studies which were restricted to single areas, the goal of this paper is to evaluate model performance in all 20 ESCAPE study areas for spatial variation of PM and NO_2.

Methods

Study design

ESCAPE study areas included 20 sites with simultaneous measurements of both PM and NO_2, and 20 sites where only NO_2 was measured in each area. As we did not have PM concentration data available for sampling sites other than the 20 PM sites in each area, we made use of the high correlation between the annual average concentration of traffic-related PM metrics such as $PM_{2.5}$ absorbance, Copper in PM_{10} (PM_{10} Cu) and NO_2[7].We assessed the performances of LUR models developed using the PM/ NO_2 sites to predict the NO_2 concentrations at the sites where only NO_2 was measured. We used this as a surrogate for the true hold-out validation. In the paper we will refer to the PM / NO_2 sites as training sites and the NO_2 only sites as test sites.

Study areas and air pollution measurements

Details of the ESCAPE study design and the measurement campaign have been described previously[6, 7]. Briefly, an intensive monitoring campaign was conducted in 20 European study areas between October 2008 and May 2011. The abbreviations regarding to the study areas are shown in table S1. In each area, we chose sampling sites at street, urban background and regional background locations. These sites were selected to represent the spatial distribution of residential addresses of participants of cohort studies in these areas. Sampling of NO_2 was conducted at 40 sites, at half of which we also sampled PM. In the Netherlands/Belgium and Cataluña measurements were performed at 40 PM sites and 80 NO_2 sites. At each of the PM sites, NO_2 was measured simultaneously. The site selection procedure (http://www.escapeproject.eu/manuals/index.php) specified that the 20 PM sites had to be a random selection of the 40 sites in each area. This was not always achieved as it is easier to find monitoring locations for the passive NO_2 sampler than for the active PM samplers. We compared the distributions of NO_2 concentrations measured at the sites where only nitrogen oxides were measured, to those at the sites where both nitrogen oxides and PM were measured. Each selected site was measured in three two-week sampling periods in the cold, warm and intermediate seasons. Due to limited amount of samplers, five sites and the reference site were measured simultaneously. The measured values were adjusted for temporal variation using continuous measurements at a background location which was not influenced by local pollution and annual average concentrations for each site were calculated and were used for model development.

NO_2 was measured using Ogawa badges and following the Ogawa analysis protocol (Ogawa&Co V 3.98, USA, Inc.). $PM_{2.5}$ and PM_{10} samples were collected on pre-weighted filters using Harvard Impactors. These filters were then used to measure absorbance and detect elemental composition (e.g. Cu) by Energy Dispersive X-ray fluorescence (ED-XRF) at Cooper Environmental Services (Portland, USA). More detail is provided in a separate paper (de Hoogh, in preparation). Briefly, Forty-eight elements were measured. Quality assurance and control included analysis of NIST reference material(SRM 1128 and SRM987). All

analysis batches passed quality criteria of the laboratory. In each study area, about 20 field blanks and field duplicates were taken. We calculated the mean field blank and the detection limit.

Predictor variables for LUR model

We extracted values for the GIS predictor variables at the coordinates of sampling sites using ArcGIS (ESRI, Redlands, California). Details of the predictor variables have been described in previous papers[8, 10]. Briefly, the predictor variables were derived from both centrally available Europe-wide GIS database and locally collected GIS data from partners.

Central GIS predictor variables were comprised of road network, land use, population density and altitude data. High resolution digital road network was obtained from Eurostreets version 3.1(1:10,000 resolution) which were based on the TeleAtlas MultiNetTM dataset for the year 2008. For all roads and major roads, the total lengths of roads were calculated within a buffer size of 25, 50, 100, 300, 500, 1000 meters. Land use variables were derived from the CORINE (Coordination and INformation on the Environmental programme) database for the year 2000 for the buffer sizes of 100, 300, 500, 1000 and 5000 meters. Digital elevation data (SRTM 90m) were obtained through the Shuttle Radar Topographic Mission (http://srtm.csi.cgiar.org/).

Detailed road network with linked traffic intensity were available locally for most study areas. The accuracy should be at least 10m compared to the central road network. Data on traffic density were aggregated to annual means, as we were modeling annual mean concentrations. We did not obtain traffic counts for the exact monitoring hours as these traffic data were generally not available. Local land use, population density, altitude and other local variables were also extracted for modeling.

LUR model development

Models for $PM_{2.5}$ absorbance were developed by local partners supervised centrally while models for PM_{10} Cu were built centrally at IRAS (Institute for Risk Assessment Sciences, Utrecht University). Separate models were built for each area, we did not attempt to build a universal model to cover all study areas in view of differences between areas not sufficiently characterized by the available GIS data. For this paper we further developed models for NO_2, using only the data from the training sites. Detailed procedures of model development and results have been published elsewhere[8, 10]. LUR model results for elemental composition will be published later. A supervised stepwise regression was used to develop the LUR model. We first evaluated univariate regression of the corrected annual concentrations by entering all potential predictor variables. The variable producing the highest adjusted R^2 and having the a priori defined direction of effect (e.g. positive for traffic intensity) was selected as the first predictor. Secondly, the remaining variables were added separately and we assessed whether the variable with the highest increase in adjusted R^2 improved the model by at least 1%. This process continued until no more variable with the a priori specified sign could increase the model adjusted R^2 by at least 1%. In the final step, we excluded the

variables which had a p value >0.1. We checked whether the variance inflation factor (VIF) was lower than 3 in order to avoid multi-collinearity.

Model evaluation

As previously described[12], we performed two evaluation approaches:

1. Leave-one-out-cross-validation (LOOCV), which successively left out one site from the training data set and estimated models based on the remaining N-1 sites. In this procedure, the variables in the model were the same as identified using the full training data set; only the coefficients of the model changed.
2. Hold-out evaluation (HEV). For NO_2 this was straightforward as we compared NO_2 model predictions with measured NO_2 concentrations at test sites. True HEV for PM components was infeasible as training sets for PM were too small to split up for model building and validation. As an alternative, we evaluated PM models by investigating the correlations between the predicted values of PM metrics and the measured NO_2 at the test sets (HEV R^2).

A systematic check of the model evaluations was conducted in the following ways:

1. We restricted this analysis to PM components and areas with high correlations with measured NO_2 (squared Pearson correlation coefficient R^2>0.5).
2. We further evaluated whether the PM models could also fit NO_2 well by checking the correlations between predicted PM concentrations and measured NO_2 concentrations (R^2_{NO2}) at the training sets and included only areas where R^2_{NO2} was > 0.5.
3. Finally, we compared the variability and tested the distributions of NO_2 in the training and the test sets of each area by simple boxplot and t-tests to assess similarity of the two types of sites.

We compared the model performances of the PM metrics with the model performance of the NO_2 models, the latter reflecting true HEV. We evaluated the accuracy of the HEV only for the NO_2 model by calculating the root mean squared error (RMSE) and the mean difference between predictions and observations (MD) as the HEV for $PM_{2.5}$ absorbance and PM_{10} Cu was indirect. As a check of our approaching using correlation with NO_2 as surrogate for HEV, we made use of two larger areas (the Netherlands & Belgium and Cataluña) with 40 PM sites. Ten datasets were randomly generated for model development (n=20) and evaluation (n=20) for $PM_{2.5}$ absorbance and PM_{10} Cu. We compared the indirect HEV R^2 (based on correlation with NO_2) with true HEV in these two areas.

We calculated the HEV R^2 by truncating the values of predictors in the test data sets that were outside the range of the values observed in the data set for model development. This is standard procedure within ESCAPE for exposure assignment and was done to prevent unrealistic predictions based on model extrapolations. Our previous study showed that with a small amount of locations for model building, the range of the variables for the model development may not cover the whole range when they were extended to larger numbers of independent test sites. Therefore, the predicted values may strongly deviate from the observations, especially when non-linear functions are used such as 1/(distance to road). [12] We

explored the impact of truncation on HEV R^2. Analyses were conducted with SAS 9.2.

Results

Table 1 shows the squared Pearson correlation coefficients between NO_2 and selected PM components. Median correlations were high for both $PM_{2.5}$ absorbance and PM_{10} Cu. Substantial variability of correlations was found between study areas. For $PM_{2.5}$ absorbance, the R^2 with NO_2 in all the ESCAPE study areas were higher than 0.5. For PM_{10}Cu, Gyor was the only area with low correlation with NO_2. The highest correlations between NO_2 and PM components were frequently observed in big cities e.g. Munich (Germany), London/Oxford (United Kingdom), Barcelona (Spain) and Paris (France) with large spatial concentration contrasts compared with relatively small cities with smaller spatial contrast e.g. Gyor (Hungry) and Kaunas (Lithuania)[7].

Table 1 squared Pearson correlation coefficients (R^2) between measured NO_2 and $PM_{2.5}$ absorbance and PM_{10} Cu in 20 European study areas.

Study areas	$PM_{2.5}$ absorbance	PM_{10} Cu
Oslo, Norway	0.75	0.73
Stockholm, Sweden	0.86	0.64
Helsinki, Turku, Finland	0.81	0.91
Copenhagen, Denmark	0.86	0.84
Kaunas, Lithuania	0.55	0.69
Manchester, UK	0.74	0.76
London, Oxford, UK	0.88	0.89
Netherlands & Belgium	0.86	0.83
Ruhr area, Germany	0.89	0.91
Munich, Germany	0.87	0.94
Vorarlberg, Austria	0.59	0.70
Paris, France	0.90	0.89
Györ, Hungary	0.65	0.25
Lugano, Switzerland	0.64	0.85
Turin, Italy	0.87	0.81
Rome, Italy	0.89	0.77
Barcelona, Spain	0.91	0.87
Catalunya, Spain	0.89	0.83
Athens, Greece	0.85	0.78
Herakion, Greece	0.63	0.66
Median	0.86	0.82
Interquartile range	0.19	0.17

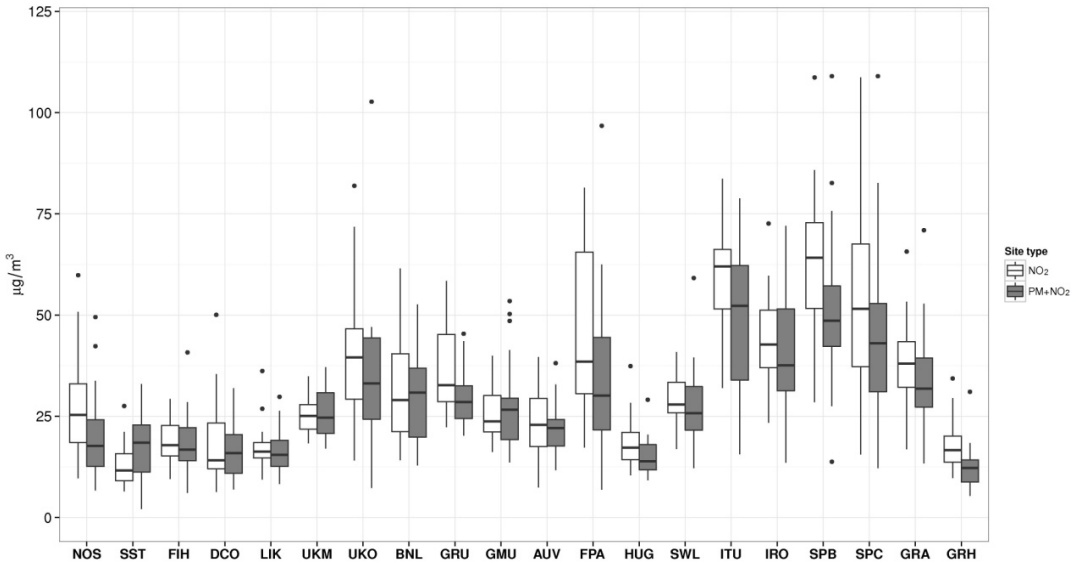


Figure 1 Boxplot of NO_2 concentrations at PM/NO_2 (training) sites and NO_2-only (test) sites in 20 ESCAPE study areas. The upper, middle and bottom layers of the box show the 75, 50, 25th percentiles of the dataset. NOS: Oslo, Norway; SST: Stockholm, Sweden; FIH: Helsinki, Finland; DCO: Copenhagen, Denmark; LIK: Kaunas, Lithuania; UKM: Manchester, UK; UKO: London/Oxford, UK; BNL:Netherlands and Belgium; GRU: Ruhr area, Germany; GMU: Munich, Germany; AUV: Vorarlberg, Austria; FPA: Paris, France; HUG: Gyor, Hungry; SWL: Lugano, Switzerland; ITU: Turin, Italy; IRO: Rome, Italy; SPB: Barcelona, Spain; SPC: Cataluña, Spain; GRA: Athens, Greece; GRH: Heraklion, Greece

Table 2 Comparison between model R^2 and LOOCV R^2 for NO_2 and PM components (training sites), R^2 between predicted concentrations and measured NO_2 at training sites (R^2_{NO2}) and R^2 between predicted concentrations and measured NO_2 at test sites (HEV R^2) in 20 European study areas.

Modeled Pollutant	Model R^{2a}		LOOCV R^{2b}		R^2_{NO2}[c]		HEV R^{2d}	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR
NO_2	0.88	0.05	0.83	0.10	1.00	0.00	0.52	0.24
$PM_{2.5}$ absorbance	0.87	0.13	0.81	0.16	0.80	0.07	0.44	0.35
$PM_{10}Cu$	0.82	0.18	0.76	0.22	0.77	0.11	0.40	0.25

[a]Model R^2: Model adjusted R^2; [b]LOOCV R^2: Leave-One-Out-Cross-validation R^2
[c]R^2_{NO2} shows the correlations between predicted NO_2 or PM components concentrations with measured NO_2 concentrations at the training sites, being the NO_2/PM sites.
[d]HEV R^2 is hold-out evaluation R^2, approximated by the correlation of model predictions with measured NO_2 at test sites, which is NO_2-only sites

The variability of NO_2 concentrations was similar for the training sites and the test sites for most areas (Figure 1). The mean NO_2 concentration did not differ significantly between the training and the test sites, with the exception of the study areas of Paris, Heraklion, Turin, Ruhr area, Oslo and Stockholm county ($p<0.05$). Table 2 shows the distributions of model R^2 and LOOCV R^2 for NO_2, $PM_{2.5}$ absorbance and PM_{10}Cu and R^2 between predicted concentrations and measured NO_2 at the training sites (R^2_{NO2}) and test sites (HEV R^2).

Figure 2 and online supplement tables S2-S4 show the model performance and structure for all individual study areas, including the predictor variables in the identified LUR models. Vorarlberg and Gyor were excluded from $PM_{2.5}$ absorbance and PM_{10} Cu respectively due to lower Correlation R^2 with measured NO_2 or R^2_{NO2} than 0.5. High median model R^2s were observed as 0.82 for PM_{10} Cu, 0.87 for $PM_{2.5}$ absorbance and 0.88 for NO_2. The median LOOCV R^2s were 5~6% lower than the model R^2s. The median correlations (R^2_{NO2}) of the PM model predictions with the measured NO_2 concentrations in the training data sets were as high as the squared correlations (Pearson R^2) between observations (Table 1), ranging from 0.77 for PM_{10} Cu to 0.80 for $PM_{2.5}$ absorbance. In contrast, the models explained substantially less variation in the independent test data sets. The NO_2 models developed on the 20 training sites had the best prediction ability (median HEV R^2=0.52). The RMSE and MD ranged from 3.18 to 18.57μg/m^3 (median: 6.53 μg/m^3) and from -8.64 to 2.71 (median: -2.38 μg/m^3) respectively. The $PM_{2.5}$ absorbance and PM_{10} Cu models explained only a slightly smaller fraction of the measured NO_2 concentration than the NO_2 models (median HEV R^2=0.44 and 0.40 respectively). The IQR of R^2s of each pollutant was higher for hold-out evaluations than for cross-validation and model development, indicating substantial variability of HEV R^2s across study areas.

Table 3 Comparison between model R^2 and LOOCV R^2 for PM components and indirect and direct hold out evaluation in the Netherlands& Belgium and Cataluña (median (IQR))

Pollutants	Areas	Model R^{2a}	LOOCV R^{2b}	R^2_{NO2}[c]	HEVR^2 (NO_2)[d]	HEV R^2(PM)[e]
$PM_{2.5}$ absorbance	BNL	0.90(0.06)	0.87(0.08)	0.83(0.04)	0.68(0.11)	0.76(0.13)
	SPC	0.85(0.10)	0.81(0.14)	0.82(0.10)	0.56(0.15)	0.51(0.17)
PM_{10}Cu	BNL	0.84(0.04)	0.79(0.11)	0.83(0.07)	0.57(0.12)	0.56(0.09)
	SPC	0.82(0.08)	0.77(0.09)	0.71(0.10)	0.45(0.32)	0.45(0.36)

[a]Model R^2: Model adjusted R^2; [b]LOOCV R^2: Leave-One-Out-Cross-validation R^2

[c]R^2_{NO2} shows the correlations between predicted NO_2 or PM components concentrations with measured NO_2 concentrations at the training sites, being the NO_2/PM sites.

[d]Indirect HEV R^2(NO_2): correlations between predicted PM components and measured NO_2 at the 20 test sites.

[e]Direct HEV R^2 (PM): correlations between predicted and measured PM components at the 20 test sites. BNL: Belgium & the Netherlands; SPC: Cataluña, Spain. The 40 sites were randomly divided in test and training sets 10 times

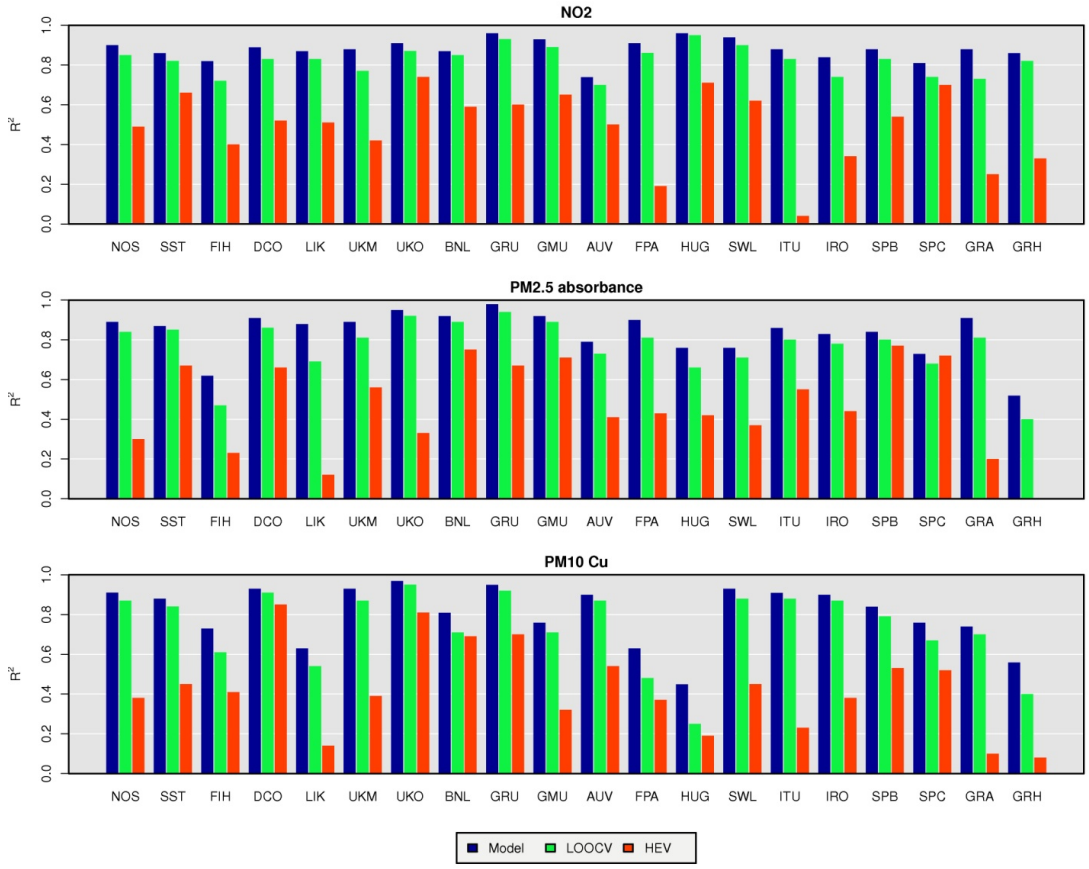


Figure 2 Model R^2; LOOCV R^2 (at NO_2/PM sites) for NO_2 & PM components; and R^2 of model predictions with NO_2 measurements (at NO_2 -only sites. HEV R^2) in 20 European countries. See Figure 1 for coding of the locations.

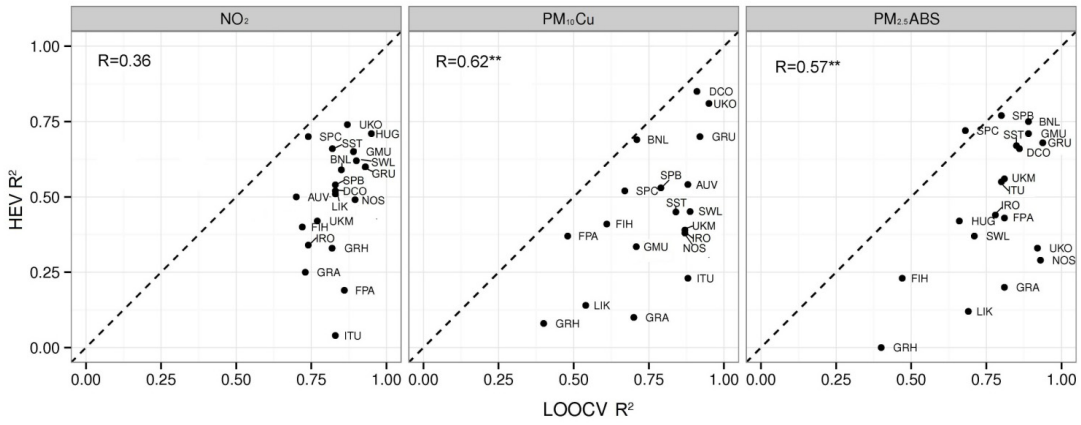


Figure 3 LOOCV R^2 (X-axis) versus HEV R^2 (Y-axis) in study areas. The codes corresponding to the areas are shown in Figure 1 and Table S1. R^2:*p<0.1, **p<0.05

In the sensitivity analysis with 10 sets of random selected 20 training and test PM sites in the Netherlands and Belgium and in Cataluña, the HEV R^2 validated by the same PM metric did not significantly deviate from the HEV R^2 validated by NO_2 for $PM_{2.5}$ absorbance and PM_{10} Cu (paired t-test, $p>0.1$). This supports our approach of using of NO_2 as proxy to evaluate the PM models (Table 3 and Figure S2). Similar differences were found between model R^2 and HEV R^2 for NO_2 in these two areas as in the analysis comprising all study areas.

The HEV was calculated with truncated predictors. We saw that by restricting the predictors in the test sets to the range of values that were obtained in the training sets, improved the median HEV R^2s by 8%, 5% and 8% for NO_2, $PM_{2.5}$ absorbance, and PM_{10} Cu respectively (Table S5).

Figure 3 presents scatterplots of R^2 of LOOCV versus R^2 of HEV in individual areas. In general, there were positive associations between LOOCV R^2 and HEV R^2, indicating that better models as judged from LOOCV were on average better in HEV as well. The correlations were significant ($p<0.1$) for PM_{10} Cu and $PM_{2.5}$ absorbance, but not for NO_2 ($p\geq0.1$). There was however a wide scatter. In some areas, models that exhibited very stable performances in cross-validation reflected much lower HEV R^2s than the model R^2. For instance, the models of NO_2 and PM_{10} Cu in Turin have both high model R^2 (>0.87) and high LOOCV R^2 (>0.82) whereas the HEV R^2 dropped dramatically by over 66% from model R^2 (Table S2, 4). This also applies to the models in a few other areas e.g. Paris, Kaunas, Heraklion and Athens (Table S2-4). For NO_2, the five areas with the lowest HEV R^2 (<0.40) were predominantly in southern Europe (Turin, Paris, Athens, Heraklion and Rome). For absorbance, the lowest HEV R^2 (<0.30) were found more spread, specifically in Oslo, Helsinki, Kaunas, Athens and Heraklion. For Cu, the lowest HEV R^2 (<0.30) were found more spread, specifically in Kaunas, Gyor, Turin, Athens and Heraklion.

Discussion

This study shows that for a wide range of study areas and pollutants including NO_2, $PM_{2.5}$ absorbance and PM_{10} Cu, model and LOOCV R^2 from land use regression models based on relatively small training sets overestimate predictive ability in independent test sets. Despite this overestimation, in most areas LUR models still explained a substantial fraction of the spatial variation measured at independent sites. The predictions were better for the areas e.g. Western Europe with more detailed predictor variables.

Evaluations of LUR predictive power and the effects of varying the number of sampling sites have been recently reported in four studies conducted in single areas for the pollutant NO_2[12-15]. The conclusions of these studies were variable, ranging from negligibly (LOOCV R^2: 0.67, HEV R^2: 0.64, N=20)[14] to seriously inflated R^2 of model and LOOCV R^2s compared to HEV R^2 (LOOCV R^2: 0.72, HEV R^2: 0.22, N=20)[13]. Our results for NO_2 can be directly compared with these studies. Our models based on a large multicenter study showed similar patterns as observed in our recent work in the Netherlands only[12], whereas the studies by Basagana et al. (2012)[13] and Johnson et al. (2010) [16] showed larger gaps between HEV R^2 and model or LOOCV R^2. In our current study the median HEV R^2 was still

52%, indicating that a substantial fraction of the measured variation was explained by the LUR models based upon 20 sites. In our previous work[12], we found a HEV R^2 of 63% for models based upon 24 sites.

The differences between model R^2 and HEV R^2 for PM absorbance and Cu were evaluated with the NO_2 concentration at the test sites, because independent PM data were not available. The difference between model R^2 and HEV R^2 for $PM_{2.5}$ absorbance and PM_{10} Cu was only slightly larger than for NO_2. For these PM metrics some of the gap is due to the use of NO_2 for the evaluation. To test this impact, we divided the HEV R^2 by the R^2_{NO2} in table 3, which can be interpreted as the highest possible squared correlation for PM metrics. This resulted in median HEV R^2 of 62% and 52% for $PM_{2.5}$ absorbance and PM_{10} Cu, respectively. These adjusted HEVs are still much larger than the LOOCV. These PM metrics have strong relations to tailpipe and non-tailpipe traffic emissions[16, 17]. We restricted the evaluation to the areas with high correlation of the measured concentrations with NO_2 (Table 1) and high correlations of PM model predictions with NO_2 at the sites used for model development (Table S2, 4) (R^2>0.5). Our sensitivity analysis indicated that use of NO_2 proxy for HEV showed no significant difference as compared to use of the same PM metrics for true HEV in the Netherlands & Belgium and Cataluña, suggesting that it was reasonable to use NO_2 to evaluate the prediction ability of $PM_{2.5}$ absorbance and PM_{10} Cu models in this study. A limitation of the use of NO_2 for PM metrics evaluation is that we can only evaluate the correlation and not the accuracy of the model. The evaluation of the NO_2 models suggested that the predictions may slightly underestimate concentrations in most of the study areas.

The differences between model R^2 and HEV R^2 were recognized as a phenomenon of some over-fitting, in combination with incomplete representation of relevant area characteristics in small training sets[11-13, 18]. The model R^2 and LOOCV R^2 may be inflated when models are based on small number of training sites and when many candidate predictors are available. In the ESCAPE study, we used a supervised approach with a priori defined directions of effects and restricted the potential predictors to limit the risk of over-fitting. Our results showed that despite substantial variability of LOOCV R^2 and HEV R^2 in study areas, the areas with higher LOOCV R^2 tended to produce better predictions for the independent data, therefore, suggested more robust performances of models in predicting values at the cohort addresses in some areas.

We also noted that in a few areas, LOOCV R^2 was much lower than HEV R^2. This is likely explained at least in part by simple random variation (associations might have been different in these areas with other training and/or test sets in these same areas). However, the scatterplots in Figure 3 show that LOOCV R^2 and HEV R^2 were positively associated, suggesting that models in some areas were truly more predictive than in other areas. This is supported by Figures S1 which shows that the HEV R^2 is positively associated with the correlation between NO_2 and PM component measurements. The level of the HEV R^2 could be related to complexity of study areas and quality of measurements and predictor variables. With more detailed predictor variables, the models in the Western European centers generally performed better than the models in other areas. This suggests a sensitivity

analysis in the epidemiological analysis using HEV R^2 rather than LOOCV R^2. Previous studies displayed a slight reduction of NO_2 model R^2s and LOOCV R^2s as a function of increasing number of training sites[12, 13, 15]. Our results supported this variation in model performances for a large number of areas using a standardized sampling and modeling method. We compared performances between NO_2 models which were centrally built for testing by IRAS based on 20 sites (40 for Netherlands & Belgium and Cataluña) and models which were optimized by local partners based on a full set of 40 sites (80 for Netherlands & Belgium and Cataluña). The median R^2s of model and LOOCV cross validation decreased from 0.88 to 0.81 and from 0.83 to 0.73 respectively (Fig.S3). The effect of restricting the out-of-range predictor values to the range of the training sets has been discussed elsewhere[12, 13]. Our results support that the range truncation approach increases the HEV R^2 of our LUR models in most study areas. It is therefore important that the selected sites cover the variability of predictor variables and pollutant concentrations in the study area well[8].

As the PM models will be applied to the epidemiological studies in all the ESCAPE study areas, the quality of estimated exposure of cohorts will largely depend on the prediction ability of models to the independent dataset, i.e. the HEV R^2. Although we cannot directly estimate absolute errors of PM metrics in the test sets, the HEV R^2 with measured NO_2 can still be informative to the health studies. We will, for instance, include model performance in meta-regressions of the cohort-specific effect estimates which are currently being developed. In summary, we found model R^2 and LOOCV R^2 to be substantially higher than HEV R^2 in LUR models developed for $PM_{2.5}$ absorbance and PM_{10} copper in 20 study areas across Europe. Despite this overestimation, in most areas LUR models still explained a substantial fraction of the variation measured at independent sites.

Acknowledgements

We thank all those who participated and were responsible for air pollution measurements, data management, model building and project supervision in all the ESCAPE study areas. The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2011) under grant agreement number: 211250.

References

1. Brunekreef, B.; Holgate, S. T., Air pollution and health. *Lancet* **2002,** *360*, (9341), 1233-42.
2. Pope, C. A.; Burnett, R. T.; Thun, M. J.; Calle, E. E.; Krewski, D.; Ito, K.; et al., Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama-J Am Med Assoc* **2002,** *287*, (9), 1132-1141.
3. Dockery, D. W.; Pope, C. A.; Xu, X. P.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; et al., An Association between Air-Pollution and Mortality in 6 United-States Cities. *New Engl J Med* **1993,** *329*, (24), 1753-1759.
4. Beelen, R.; Hoek, G.; van den Brandt, P. A.; Goldbohm, R. A.; Fischer, P.; Schouten, L. J.; et al., Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environmental Health Perspectives* **2008,** *116*, (2), 196-202.
5. Dijkema, M. B. A.; Mallant, S. F.; Gehring, U.; van den Hurk, K.; Alssema, M.; van Strien, R. T.; et al., Long-term Exposure to Traffic-related Air Pollution and Type 2 Diabetes Prevalence in a

Cross-sectional Screening-study in the Netherlands. *Environ Health-Glob* **2011,** *10*.
6. Cyrys, J.; Eeftens, M.; Heinrich, J.; Ampe, C.; Armengaud, A.; Beelen, R.; et al., Variation of NO_2 and NO_x concentrations between and within 36 European study areas: results from the ESCAPE study. *Atmos Environ* **2012,** *62*, 374 - 390.
7. Eeftens, M.; Tsai, M.; Ampe, C.; Anwander, B.; Beelen, R.; Cesaroni, G.; et al., Spatial variation of $PM_{2.5}$, PM_{10}, $PM_{2.5}$ absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO_2 - Results of the ESCAPE project. *Atmospheric Environment* **2012,** *62*, 303 - 317.
8. Eeftens, M.; Beelen, R.; de Hoogh, K.; Bellander, T.; Cesaroni, G.; Cirach, M.; et al., Development of land use regression models for $PM_{2.5}$, $PM_{2.5}$ absorbance, PM_{10} and PMcoarse in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* **2012**, 11195-11205.
9. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; et al., A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* **2008,** *42*, (33), 7561-7578.
10. Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; et al., Development of NO_2 and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe - the ESCAPE project. *Atmospheric Environment* **2013,** *72, 10-23*.
11. Babyak, M. A., What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine* **2004,** *66*, 411 - 421.
12. Wang, M.; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B., Systematic Evaluation of Land Use Regression Models for NO_2. *Environ Sci Technol* **2012,** *46*, (8), 4481-4489.
13. Basagana, X.; Rivera, M.; Aguilera, I.; Agis, D.; Bouso, L.; Elosua, R.; et al., Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment* **2012,** *54*, 634-642.
14. Madsen, C.; Gehring, U.; Haberg, S. E.; Nafstad, P.; Meliefste, K.; Nystad, W.; et al., Comparison of land-use regression models for predicting spatial NO(x) contrasts over a three year period in Oslo, Norway. *Atmospheric Environment* **2011,** *45*, (21), 3576-3583.
15. Johnson, M.; Isakov, V.; Touma, J. S.; Mukerjee, S.; Ozkaynak, H., Evaluation of land-use regression models used to predict air quality concentrations in an urban area. *Atmospheric Environment* **2010,** *44*, (30), 3660-3668.
16. He, L. Y.; Hu, M.; Zhang, Y. H.; Huang, X. F.; Yao, T. T., Fine particle emissions from on-road vehicles in the Zhujiang Tunnel, China. *Environ Sci Technol* **2008,** *42*, (12), 4461-4466.
17. Bukowiecki, N.; Lienemann, P.; Hill, M.; Figi, R.; Richard, A.; Furger, M.; et al., Real-World Emission Factors for Antimony and Other Brake Wear Related Trace Elements: Size-Segregated Values for Light and Heavy Duty Vehicles. *Environmental Science & Technology* **2009,** *43*, (21), 8072-8078.
18. Craig, M. H.; Sharp, B. L.; Mabaso, M. L. H.; Kleinschmidt, I., Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure. *Int J Health Geogr* **2007,** *6*.

Chapter 3

Appendix

(Online: http://pubs.acs.org/doi/suppl/10.1021/es305129t)

Table S1 Coding list of study areas and number of sites per area

Code	Study Area	N (PM/NO_2)	N (NO_2 only)
NOS	Oslo, Norway	19	20
SST	Stockholm, Sweden	19	20
FIH	Helsinki, Turku, Finland	20	20
DCO	Copenhagen, Denmark	20	20
LIK	Kaunas, Lithuania	20	20
UKM	Manchester, UK	19	20
UKO	London, Oxford, UK	20	20
BNL	Netherlands &Belgium	40	40
GRU	Ruhr area, Germany	20	20
GMU	Munich, Germany	20	20
AUV	Vorarlberg, Austria	20	20
FPA	Paris, France	20	20
HUG	Györ, Hungary	20	20
SWL	Lugano, Switzerland	20	20
ITU	Turin, Italy	20	20
IRO	Rome, Italy	20	20
SPB	Barcelona, Spain	20	20
SPC	Catalunya, Spain	40	40
GRA	Athens, Greece	20	20
GRH	Herakion, Greece	20	20

Code is country followed by area. Deviations from 20 due to exclusion of non-representative sites (Eeftens et al. EST&t 2012;11195-11205)

Table S2 Model R^2, LOOCV R^2 for NO_2 (NO_2/PM sites) and R^2,root mean square errors (RMSE) and MD between predicted NO_2 and measured NO_2 at NO_2 only sites in 20 European countries, including predictor variables.

Code	Model R^2	LOOCV R^2	HEV[a] R^2	RMSE $\mu g/m^3$	MD[b] $\mu g/m^3$	Variables
NOS	0.85	0.81	0.49	8.95	-8.64	HDRES5000
SST	0.88	0.85	0.68	3.18	1.84	TRAFNEAR TRAFLOAD_50
FIH	0.81	0.72	0.40	4.56	1.47	ROADLENGTH_500 ROADLENGTH_50
DCO	0.88	0.83	0.52	8.10	-2.54	POP_1000L TRAFLOAD_50 MAJORROADLENGTH_100 INDUSTRY_5000
LIK	0.86	0.83	0.51	4.08	-1.46	POP_100 POP1000
UKM	0.87	0.77	0.42	3.92	-0.19	INTMAJORINVDIST LDRES_1000 MAJORROADLENGTH_100 SQRALT HDRES_300 ROADLENGTH_100
UKO	0.90	0.87	0.74	7.79	-3.25	MAJORROADLENGTH_1000 HLDRES5000
BNL	0.84	0.80	0.75	5.46	-4.58	INTMAJORINVDIST ROADLENGTH500 TRAFMAJORLOAD_50 HEAVYTRAFMAJORLOAD _25_1000 PORT_5000 HEAVYTRAFMAJORLOAD_25

Code	Model R^2	LOOCV R^2	HEV[a] R^2	RMSE µg/m^3	MD[b] µg/m^3	Variables ***(Continued)***
GRU	0.95	0.93	0.60	7.11	-5.44	LDRES_100 TRAFLOAD_50 INDUSTRY_5000 MAJORROADLENGTH_1000 PM10_2008_20TO40
GMU	0.92	0.89	0.65	3.88	2.71	HD_LD_RES_LVA1000 ROADLENGTH50 ROADLENGTH_300 TRAFLOAD50
AUV	0.73	0.70	0.50	5.64	-2.07	ROADLENGTH300 RES5000 TRAFLOAD300
FPA	0.90	0.86	0.19	18.57	-1.24	NATURAL_1000 TRAFMAJORLOAD_300
HUG	0.95	0.95	0.71	3.57	-3.27	TRAFLOAD_100 HDRES_500 INDUSTRY_5000 MAJORROADLENGTH_500
SWL	0.93	0.90	0.62	3.84	-1.09	SQRALT TRAFMAJOR TRAFLOAD500 TRAFLOAD50
ITU	0.87	0.83	0.04	12.37	-3.08	LDRES300 HDRES5000 ROADLENGHT100 NATURAL5000
IRO	0.83	0.74	0.34	9.43	-1.64	ROADLENGTH_1000 MAJORROADLENGTH_25 DISTINVMAJOR1
SPB	0.87	0.83	0.54	12.64	-5.97	INTINVDIST2 TRAFLOAD_100 HHOLD_5000 NATURAL_5000
SPC	0.80	0.74	0.70	12.25	-8.08	HDRES_5000 TRAFMAJORLOAD_500 NATURAL_5000 INTMAJORINVDIST2
GRA	0.87	0.73	0.25	9.70	-2.21	ROADLENGHT_1000 URBGREEN_5000 INDUSTRY_1000 HHOLD_500
GRH	0.85	0.82	0.33	5.96	-4.00	ROADLENGTH_100 TRAFLOAD_300 TRAFMAJORLOAD_25 POP_500 AIRPORT2_1000 DISTINVMAJORC1

[a]HEV R^2 with variable range restriction;
[b]MD: the mean of difference between predictions and observations per area.
.

Table S3 Model R^2, LOOCV R^2 for $PM_{2.5}$ absorbance (NO_2/PM sites) and R^2 between predicted $PM_{2.5}$ABS and measured NO_2 at PM/NO_2 and NO_2 only sites in 20 European countries, including predictor variables.

Code	Model R^2	LOOCV R^2	R^2_{NO2}[a]	HEV R^2[b]	Variables
NOS	0.95	0.93	0.79	0.29	NATURAL300 ROADLENGTH25 SQRALT MAJORROADLENGTH50
SST	0.86	0.85	0.87	0.67	TRAFLOAD1000 roadlength_500 heavytrafload_50 water_5000
FIH	0.61	0.47	0.57	0.23	TRAFLOAD_50 URBNATURAL_500L
DCO	0.9	0.86	0.8	0.66	MAJORROADLENGTH_300 HHOLD_5000 Trafnear
LIK	0.87	0.69	0.62	0.12	Industry_5000 TRAFLOAD50 LDRES300 POP100
UKM	0.88	0.81	0.8	0.56	TRAFMAJOR MAJORROADLENGTH_100 NATURAL_5000 URBGREEN_1000
UKO	0.94	0.92	0.78	0.33	ROADLENGTH_100 HEAVYTRAFLOAD500 HLDRES5000
BNL	0.91	0.89	0.81	0.75	DISTINVMAJORC2 TRAFLOAD_500 MAJORROADLENGTH50 REG_EST_PM25abs
GRU	0.97	0.94	0.88	0.67	HDLDRES_5000 HEAVYTRAFLOAD_50 heavytrafload_100 heavytrafload_100_1000
GMU	0.91	0.89	0.8	0.71	industry_5000 pop_1000 TRAFLOAD50 ROADLENGTH50
FPA	0.89	0.81	0.84	0.43	ROADLENGTHKSN3150_300 INDUSTRY_5000 HDRES_1000
HUG	0.75	0.66	0.81	0.42	TRAFMAJORLOAD_100 MAJORROADLENGTH_300 ROADLENGTH_500 TRAFLOAD_100
SWL	0.75	0.71	0.79	0.37	DISTINVMAJOR_C2 ROADLENGTH300 HEAVYTRAFLOAD50 SQRALT

Code	Model R^2	LOOCV R^2	R^2_{NO2}[a]	HEV R^{2b}	Variables ***(Continued)***
ITU	0.85	0.8	0.85	0.55	res1000
					majorroadlength_100
					natural5000
					res100
					trafload1000
					natural1000
IRO	0.82	0.78	0.75	0.44	Intmajorinvdist
					pop01_5000
SPB	0.83	0.8	0.8	0.77	hdres_300
					intinvdist2
					trafload_50
SPC	0.72	0.68	0.7	0.72	intmajorinvdist1
					roadlenght_1000
					natural_5000
					distinvmajorc1
GRA	0.9	0.81	0.76	0.2	TRAFMAJORLOAD_25
					ROADLENGTH_300
					HDRES_5000
					TRAFLOAD_500
					MAJORROADLENGTH_50
GRH	0.51	0.4	0.52	0	POP_100
					MAJORROADLENGTH_300

[a]$R^2_{(NO2)}$ is correlations between predicted $PM_{2.5}$ absorbance and measured NO_2 at PM/NO_2 sites
[b]HEV R^2 with variable range restriction;

Table S4 Model R^2, LOOCV R^2 for $PM_{10}Cu$ (NO_2 sites) and R^2 between predicted $PM_{10}Cu$ and measured NO_2 at PM/NO_2 and NO_2 only sites in 20 European countries, including predictor variables.

Code	Model R^2	LOOCV R^2	R^2_{NO2}[a]	HEV R^{2b}	Variables
NOS	0.78	0.72	0.81	0.3	SQRALT
					DISTINVMAJOR1
					TRAFLOAD1000
SST	0.87	0.84	0.53	0.45	HEAVYTRAFLOAD_500
					DISTINVMAJOR1
					HEAVYTRAFMAJOR
FIH	0.72	0.61	0.83	0.41	ROADLENGTH_50
					POP_1000L
					TRAFMAJORLOAD_50
DCO	0.92	0.91	0.77	0.85	TRAFLOAD_300
LIK	0.62	0.54	0.77	0.14	DISTINVMAJOR1
					TRAFLOAD500
UKM	0.92	0.87	0.8	0.39	ROADLENGTH_25
					HDRES_300
					MAJORROADLENGTH_100
					MAJORROADLENGTH_100_1000
UKO	0.96	0.95	0.88	0.81	ROADLENGTH1000
					TRAFLOAD25

Code	Model R^2	LOOCV R^2	R^2_{NO2}[a]	HEV R^{2b}	Variables ***(Continued)***
BNL	0.8	0.71	0.77	0.69	TRAFMAJORLOAD_50
					ROADLENGTH25
					PORT_5000
					HEAVYTRAFMAJORLOAD_1000
					TRAFNEAR
GRU	0.94	0.92	0.94	0.7	MAJORROADLENGTH_1000
					TRAFLOAD_100
					INDUSTRY_5000
GMU	0.75	0.71	0.9	0.32	INDUSTRY5000
					TRAFNEAR
					ROADLENGTH_50
AUV	0.89	0.87	0.63	0.54	ROADLENGTH300
					RES5000
					TRAFLOAD25
FPA	0.62	0.48	0.73	0.37	INDUSTRY_1000
					TRAFLOAD_300
SWL	0.92	0.88	0.79	0.45	ROADLENGTH300
					TRAFLOAD25
ITU	0.9	0.88	0.72	0.23	NATURAL5000
					TRAFMAJORLOAD50
					POP_WEIGHTED500
IRO	0.89	0.87	0.75	0.38	MAJORROADLENGTH_50
					TRAFLOAD_1001
					TRAFMAJORLOAD_25
SPB	0.83	0.79	0.81	0.53	NATURAL_5000
					HHOLD_5000
					DISTINVMAJOR1
SPC	0.75	0.67	0.76	0.52	INDUSTRY_1000
					URBGREEN_1000
					TRAFMAJORLOAD_50
					DISTINVMAJOR1
					GREEN_300
					HDRES_5000
GRA	0.73	0.7	0.51	0.1	ROADLENGTH_100
					TRAFMAJORLOAD_300
					TRAFLOAD_25
GRH	0.55	0.4	0.64	0.08	MAJORROADLENGTH_25
					POP_100

[a]$R^2_{(NO2)}$ is correlations between predicted $PM_{10}Cu$ and measured NO_2 at PM/NO_2 sites
[b]HEV R^2 with variable range restriction;

Table S5 HEV R^2 before and after variable truncation.

Pollutants	HEV* NO_2		HEV NO_2	
	Median	IQR	Median	IQR
NO_2	0.44	0.30	0.52	0.24
$PM_{2.5}$ absorbance	0.39	0.34	0.44	0.35
$PM_{10}CU$	0.34	0.31	0.40	0.25

HEV* is without variable truncation (restriction of predictor variable values to those observed at the training sites). HEV = hold out evaluation. IQR= interquartile rang

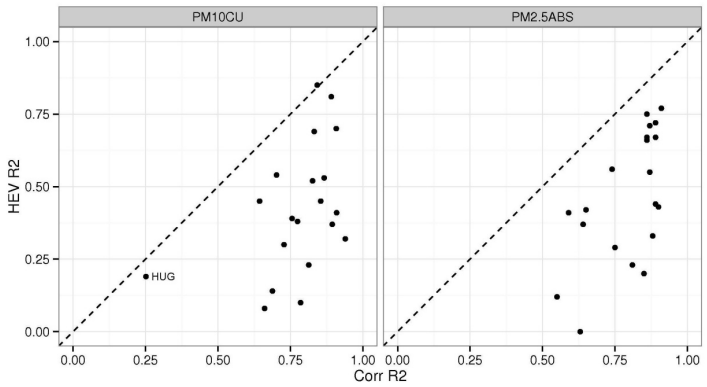


Figure S1 Relationship of the correlation between measured NO_2 and PM components (Corr R^2) and Hold Out evaluation against NO_2 at test sites (HEV R^2) in 20 study areas for PM_{10}Cu and $PM_{2.5}$ absorbance ($PM_{2.5}$ABS)

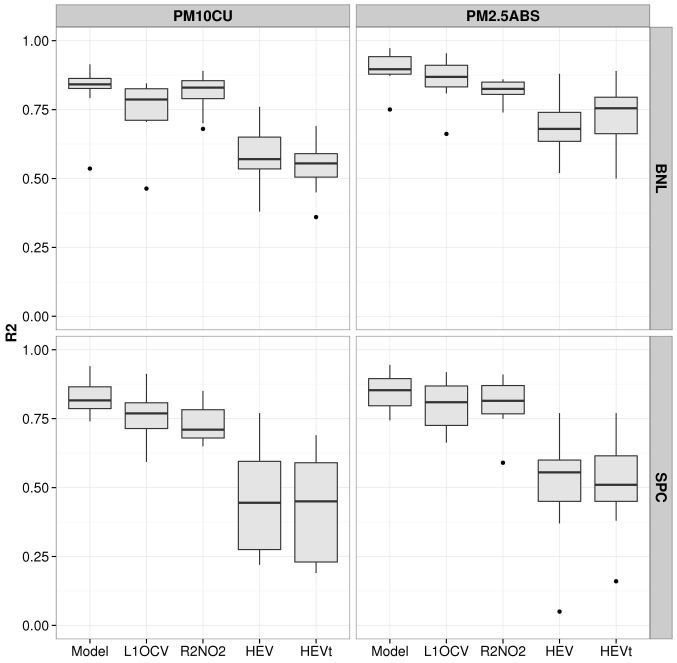


Figure S2 The model performances of PM_{10}Cu and $PM_{2.5}$ absorbance in the Netherland/Belgium and Cataluña based on 10 times random selection of training sites(N=20) from 40 PM sites. The HEV (correlation of predicted PM component with measured NO_2 at test sites) and HEVt (correlation of predicted PM component with measured PM component at test sites) are indirect and direct evaluations.

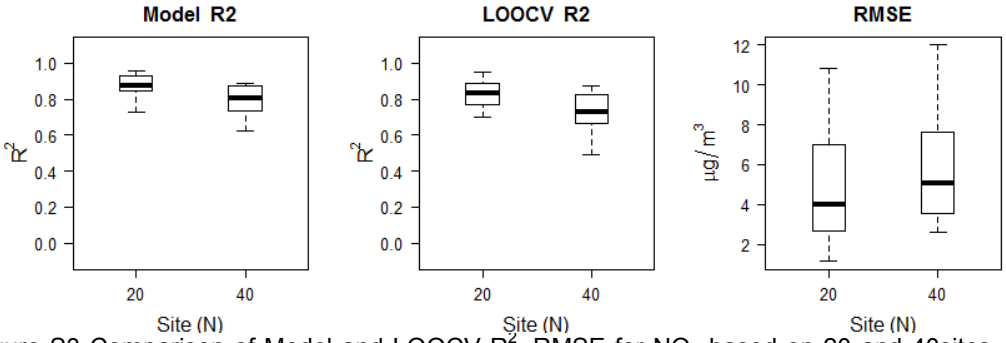


Figure S3 Comparison of Model and LOOCV R^2, RMSE for NO_2 based on 20 and 40sites. In the large Netherlands and Catalunya areas 40 versus 80 sites.

Chapter 4

Performance of Multi-Cities Land Use Regression Models for Nitrogen Dioxide and Fine Particles

Meng Wang, Rob Beelen, Tom Bellander, Matthias Birk, Giulia Cesaroni, Marta Cirach, Josef Cyrys, Kees de Hoogh, Christophe Declercq, Konstantina Dimakopoulou, Marloes Eeftens, Kirsten T Eriksen, Francesco Forastiere, Claudia Galassi, George Grivas, Joachim Heinrich, Barbara Hoffmann, Alex Ineichen, Michal Korek, Timo Lanki, Sarah Lindley, Lars Modig, Anna Mölter, Per Nafstad, Mark Nieuwenhuijsen, Wenche Nystad, David Olsson, Ole Raaschou-Nielsen, Martina Ragettli, Andrea Ranzi, Morgane Stempfelet, Dorothea Sugiri, Ming-Yi Tsai, Orsolya Udvardy, Mihaly J. Varró, Danielle Vienneau, Gudrun Weinmayr, Kathrin Wolf, Tarja Yli-Tuomi, Gerard Hoek, Bert Brunekreef

(Environmental Health Perspectives, under revision)

Abstract

Background: Land use regression (LUR) models have mostly been developed to explain intra-urban variations in air pollution based on often small local monitoring campaigns. Transferability of LUR models from city to city has been investigated, but little is known about the performance of models based on large numbers of monitoring sites covering a large area.
Objectives: To develop European and regional LUR models and to examine their transferability to areas not used for model development.
Methods: We evaluated LUR models for nitrogen dioxide (NO_2) and Particulate Matter ($PM_{2.5}$, $PM_{2.5}$ absorbance) by combining standardized measurement data from 17 (PM) and 23 (NO_2) ESCAPE study areas across 14 European countries for PM and NO_2. Models were evaluated with cross validation (CV) and hold-out validation (HV). We investigated the transferability of the models by successively excluding each study area from model building.
Results: The European model explained 56% of the concentration variability across all sites for NO_2, 86% for $PM_{2.5}$ and 70% for $PM_{2.5}$ absorbance. The HV R^2s were only slightly lower than the model R^2 (NO_2: 54%, $PM_{2.5}$: 80%, absorbance: 70%). The European NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance models explained a median of 59%, 48% and 70% of within-area variability in individual areas. The transferred models predicted a modest to large fraction of variability in areas which were excluded from model building (median R^2: 59% NO_2; 42% $PM_{2.5}$; 67% $PM_{2.5}$ absorbance).
Conclusions: Using a large dataset from 23 European study areas, we were able to develop LUR models for NO_2 and PM metrics that predicted measurements made at independent sites and areas reasonably well. This finding is useful for assessing exposure in health studies conducted in areas where no measurements were conducted.

Introduction

Many studies have documented adverse health effects associated with long-term exposure to air pollutants [1]. With the improvement of the accuracy of geographical data, air pollution models incorporating data from geographical information system (GIS) are of increasing interest in exposure assessment [2-4]. Land use regression (LUR) modeling is a popular method used for exposure assessment in health studies [5-7]. LUR is a GIS and statistics based method that exploits land use, geographic and traffic characteristics (e.g. traffic intensity, road length, population density) to explain spatial concentration variations at monitoring sites.

Land use regression models were mostly constructed and utilized to predict concentrations within metropolitan areas [8-11] or small regions[12, 13]. Often, models have been based on measurements made at a relatively small number of sampling sites (20~80 sites). Our recent study showed a positive association between the number of sampling sites and the prediction capability of models for NO_2 based on 144 sites in the Netherlands [14], in agreement with observations for Girona, Spain [15]. At least for some of the reported studies, there is still room to improve the model performances if more sampling sites were selected [2]. Several studies have reported the possibilities of building models in large areas in Europe, United States and Canada [16-19]. With a large number of sites, these models explained large fractions of NO_2 variability (61%~90%) and modest fraction variability of PM (40%~50%) across all sites. However, the large-area studies were all based upon routine monitoring. Routine monitoring networks are often not optimally designed to detect small-scale spatial variation within urban areas. There is generally only a small number of sites within individual areas with often a lack of high exposure settings. Because of the small number of routine monitoring sites within individual cities, it is typically not possible to evaluate how well a large-area model explains within-city variability. This is relevant for epidemiological studies based in individual cities. A study in Switzerland based upon study-specific monitoring suggested that a country wide model did not perform well within six of the eight geographically diverse study areas [20].

The applicability of LUR models can be increased by transferring them to adjacent areas with similar geography and GIS databases where no or few measurements were conducted. The transferability of models has been investigated for local and national models [21-23]. Most of the earlier studies recommended using the locally built models, even though transferred models explained variations in concentrations fairly well. This is because all the transferred models were city-city or country- country transfers for which local specific variables were not available and there was no advantage in the number of sampling sites as compared to the locally developed models.

So far, few studies attempted to explore the performance of LUR models with combined geographical areas in terms of prediction ability and transferability at independent sites and areas mainly due to lack of sufficient, comparable measurement data.

In the context of the ESCAPE project (European Study of Cohorts for Air

Pollution Effects), we applied a standardized approach for measurements, GIS variable collection and model development for NO_2 and Particulate Matter (PM) in 36 study areas in Europe [24-27]. We recently published LUR models developed within individual study areas for nitrogen dioxide (NO_2) and particulate matter (Beelen, et al. 2013; Eeftens et al. 2012a). The ESCAPE database provides a unique opportunity to address important questions regarding application of LUR models developed for even large areas. Therefore, the aims of this study are 1) to develop LUR models for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance based on combining the ESCAPE study areas across Europe and across four regions of Europe; 2) to evaluate the model performances systematically in terms of the model fitting and prediction ability; and 3) to investigate the transferability of the regional and European models to monitoring sites and areas not included in the model building.

Methods

Study areas and air pollution measurements

Details of the ESCAPE study design and the measurement campaign have been described previously [24, 25]. Briefly, an intensive monitoring campaign was conducted in 36 European study areas between October 2008 and May 2011. ESCAPE included 20 areas with simultaneous measurements of both PM and NO_2 at 20 sites per area, and 20 sites where only NO_2 was measured. In an additional 16 areas only NO_2 measurements were conducted at 40 sites per area. The number of measurement sites was doubled in the large study area of the Netherlands & Belgium. In each area, we chose sampling sites at street, urban background and regional background locations. These sites were selected to represent the spatial distribution of air pollution and residential addresses of participants of cohort studies in these areas. Annual averaged concentrations were calculated from three samples in different seasons and were adjusted for temporal trends with data from continuous reference sites in each area. In this paper, we selected the 23 areas (Figure1) in which traffic intensity variables were available for LUR model building in line with the importance of traffic intensity variables in model development [27]. This included 17 of the 20 PM/ NO_2 areas and 6 of the 16 NO_2 only areas. We allocated the areas to 4 regions according to the geographical location, the characteristics of the climate, the traffic intensity levels and the configuration of the cities/country. These regions included five areas in north Europe (Oslo, Stockholm, Copenhagen, Helsinki, Umea), seven in the west (Netherlands and Belgium, London, Manchester, Bradford, Ruhr area, Erfurt, Paris), six in the center (Munich, Vorarlberg, Györ, Lugano, Grenoble, Lyon) and five in the south (Turin, Rome, Athens, Barcelona, Marseille) (Figure1 and Table S1).

For this study we selected NO_2, $PM_{2.5}$ absorbance and $PM_{2.5}$ to represent traffic-related and a more complex mixture of sources, respectively. NO_2 was measured using Ogawa badges following the Ogawa analysis protocol (Ogawa&Co V 3.98, USA, Inc.). $PM_{2.5}$ samples were collected on pre-weighted filters using Harvard Impactors, and were then used to measure absorbance[24,25].

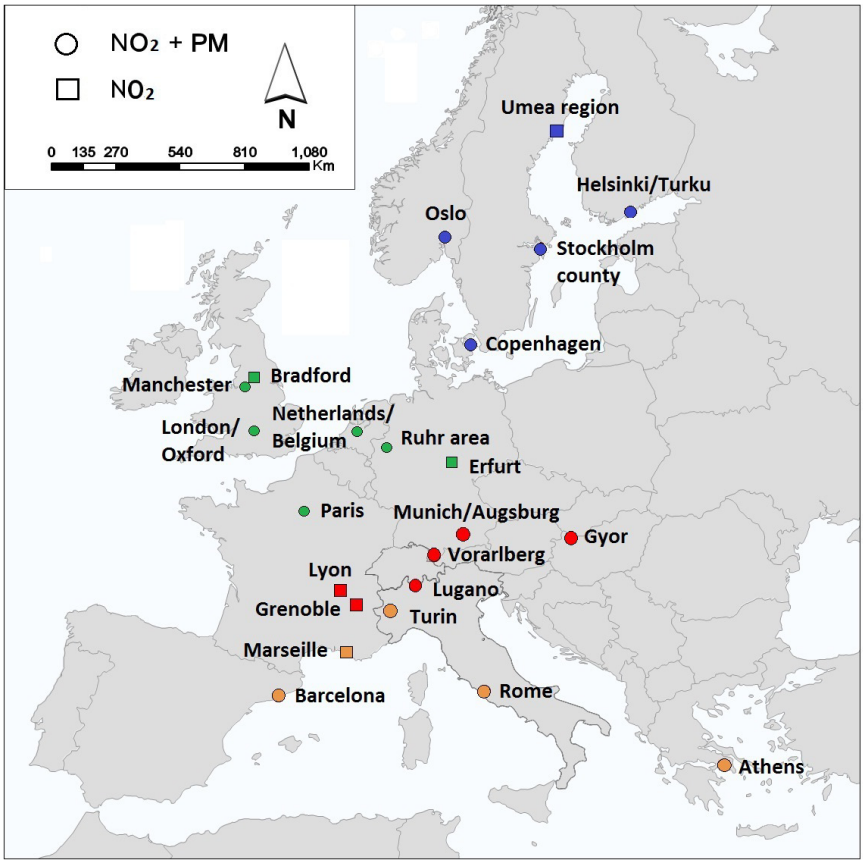


Figure 1 Map of study areas including region indication (blue: north Europe; green: west Europe; red: central Europe; orange: south Europe)

Predictor variables

We extracted values for the GIS predictor variables at the coordinates of sampling sites using ArcGIS (ESRI, Redlands, California). Details of the predictor variables have been described in previous papers [26, 27]. Briefly, the predictor variables were derived from both centrally available Europe-wide GIS databases and locally collected GIS data by the local centers using standard definitions.

Central GIS predictor variables included road network, land use, population density and altitude data. The digital road network was obtained from Eurostreets version 3.1 for the year 2008. For all roads and major roads, the total lengths of roads were calculated within a buffer size of 25, 50, 100, 300, 500, 1000 meters. Traffic intensity data were not available for this road network. Land use variables were derived from the CORINE (Coordination and INformation on the Environmental programme) database for the year 2000 for the buffer sizes of 100, 300, 500, 1000 and 5000 meters. Digital elevation data (SRTM 90m) were obtained through the Shuttle Radar Topographic Mission (http://srtm.csi.cgiar.org/). Detailed road network with linked traffic intensity were obtained from local sources for all 23 areas. Local land use, population density, altitude and other local variables were also locally extracted for modeling.

For the regional and European models, we pooled the data by including all the central GIS predictors and the local traffic variables with traffic intensity. We

combined the centrally available land use variables high and low residence density, and the natural and urban green variables as not all the areas contained them separately. We made efforts to incorporate more local common variables for specific regions to capture regional variations. We included regional background concentrations of NO_2, PM absorbance and $PM_{2.5}$ as the mean of the measured concentrations at (1 to 3) ESCAPE regional background sites in each local study area to characterize the spatial differences between study areas. In total, 49 variables were evaluated at the European level and 54, 53, 54, 64 variables in the north, west, middle and south regions (see Supplemental Material, Table S2).

Model development

A total of 960 NO_2 sites and 356 PM sites were available for modeling from 23 and 17 study areas respectively. Detailed procedures of the NO_2 and PM model development have been published elsewhere [26, 27]. The regional and European models were developed using the same strictly standardized approaches. Briefly, a supervised stepwise regression was used to develop the LUR model. We first evaluated univariate regression of the annual concentrations by entering all potential predictor variables. We forced the regional background concentration variable in the first step (for the European and regional models). Then the variable which produced the highest adjusted R^2 and which had the a priori defined direction of effect (e.g. positive for traffic intensity) was selected as the second predictor. Secondly, the remaining variables were added separately and we assessed whether the variable with the highest increase in adjusted R^2 improved the model by at least 1%. This process continued until no more variable with the a priori specified sign could increase the model adjusted R^2 by at least 1%. In the final step, we excluded the variables which had a p value >0.1. We checked whether the variance inflation factor (VIF) was lower than 3 in order to avoid multi-collinearity.

Model evaluations

We used three approaches for model evaluation:

1. We investigated the model fit at individual study areas by applying the European/regional model to the sites of each area that were used for modeling. The $Model_{intra}$ R^2 shows the within area variations explained by the European/regional models which is directly comparable with the R^2 of area-specific models. The $Model_{intra}$ R^2 is important for studies conducted within individual cities that use the European/regional model. The overall R^2 is relevant for multi-city studies that exploit both within and between city-variability of air pollution contrasts.
2. We conducted cross validations: 1) leave-one-out-cross-validation (LOOCV), which successively left out one site from the N observations (N=960 for NO_2 and 356 for PM) that was used for model development and refitted developed models based on the remaining N-1 sites. This was iterated N times and the R^2 was obtained from the regression model of the observations against the predictions. LOOCV R^2 may not sufficiently reflect the stability of model fit if the number of sites for modeling is very large. Therefore, we also used 2)

leave-one-area-out-cross-validation (LOAOCV) by leaving out all observations from a complete area, refitting the model as in LOOCV, and investigating the agreement between predicted and observed concentrations for each area that was left out. It reflects the heterogeneity of model fit due to regional variations between study areas.

3. The hold-out validation (HV) is an evaluation of model predictive power to independent sites not used for model building. It reflects the prediction ability of models to the cohort addresses within the areas on which the models had been established. As a test, we divided the full set into two parts, the training sets were used for modeling and the remaining test sets were used for external evaluation. For NO_2, we developed models using the PM/ NO_2 sites with 20-40 sites per area (480 sites in total) as training sets and the remaining 480 NO_2-only sites as test sets. For $PM_{2.5}$ and $PM_{2.5}$ absorbance, a randomly selected 25% of the PM sites stratified by study area were used for validation purpose as we had fewer sites available for PM model building than for NO_2 model building. The HV R^2 is the squared Pearson correlation between predictions and observations at the independent sites throughout the whole study area. We calculated the HV R^2 by truncating the values of predictors in the test data sets that were outside the range of the values observed in the data set for model development, to prevent unrealistic predictions based on model extrapolations [14]. Prediction errors were estimated by root mean squared error (RMSE).

Transferability of LUR models

To evaluate the prediction abilities of the regional/ European models to independent individual study areas, we developed the regional and European models by excluding one area at a time and applied the transferred models directly to the sites of the area that was left out. Therefore, 23 NO_2 models and 17 PM models were built respectively until each of the study areas had been excluded once from model building.

The $TRANS_{intra}$ R^2 is the squared Pearson correlation between observed and predicted values in each of the remaining area that was excluded from modeling. The TRANS R^2_{intra} is different from the $Model_{intra}$ R^2 as the measurements conducted in the respective validation areas were completely left out from model development.

Results

NO_2 and PM concentrations

Table 1 shows the concentration distributions of NO_2 and PM metrics across the study areas by site types. Substantial spatial variations were found for all the pollutants across Europe. The variability was larger for NO_2 than for $PM_{2.5}$. The spatial variability for $PM_{2.5}$ absorbance was intermediate between $PM_{2.5}$ and NO_2. Concentration contrasts were larger at the street sites for NO_2 and $PM_{2.5}$ absorbance

Table 1 Distributions of measured annual average NO_2 and PM concentrations across Europe

Pollutants	Type[a]	N[b]	Min	P25	Median	P75	Max
NO_2	S	454	11.80	25.48	33.98	49.90	109.00
($\mu g/m^3$)	UB	414	3.03	15.38	22.88	30.67	57.63
	RB	92	1.53	9.56	15.48	17.98	32.87
$PM_{2.5}$	S	166	7.87	12.03	17.18	21.17	36.30
($\mu g/m^3$)	UB	144	5.62	10.97	15.87	18.62	32.59
	RB	47	4.42	11.20	13.86	16.64	23.24
$PM_{2.5}$absorbance	S	166	0.78	1.63	2.16	2.81	5.09
($10^{-5}m^{-1}$)	UB	144	0.53	1.23	1.67	2.01	3.03
	RB	47	0.33	0.92	1.16	1.45	2.37

[a]site types: S-street sites; UB-urban background; RB-regional background;
[b]total number of sites in the study areas

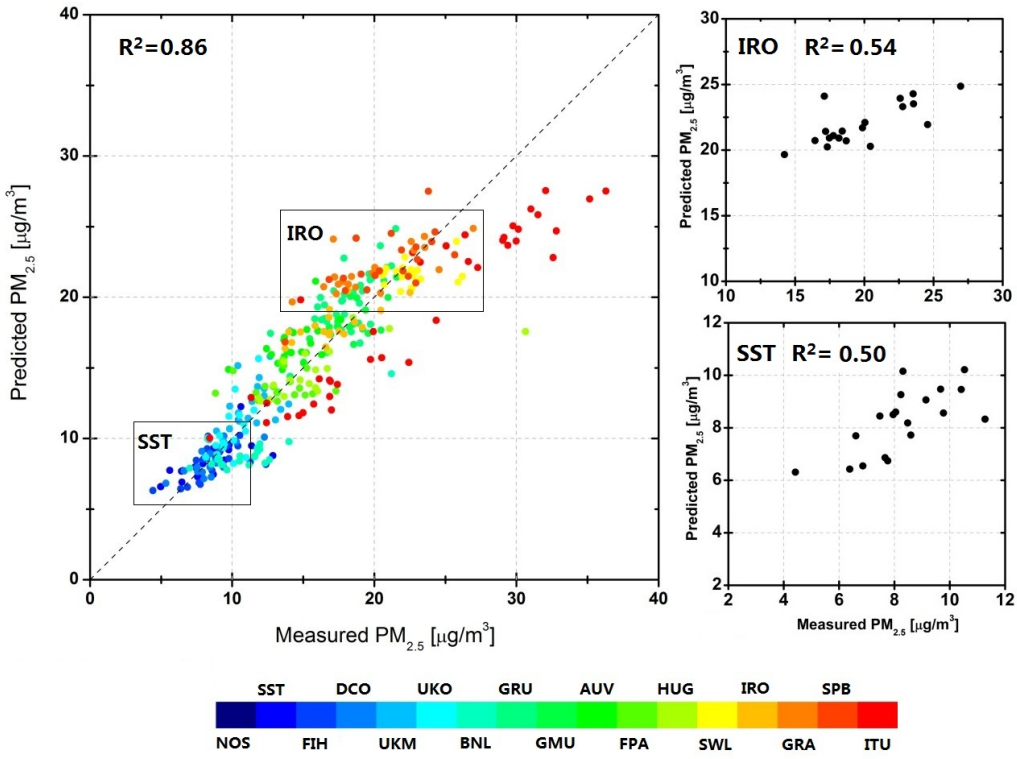


Figure 2 Scatterplot of predicted and measured $PM_{2.5}$ with study areas color coded and two area-specific examples, Rome (IRO) and Stockholm (SST). NOS: Oslo, Norway; DCO: Copenhagen, Denmark; FIH: Helsinki and Turku, Finland; SST: Stockholm, Sweden; SUM: Umea, Sweden; BNL: Netherlands and Belgium; FPA: Paris, France; GRU: Ruhr area, Germany; UKM: Manchester, UK; UKO: London/Oxford, UK; GRE: Erfurt, Germany; UKB: Bradford, UK; AUV: Vorarlberg, Austria; GMU: Munich/Augsburg, Germany; HUG: Gyor, Hungry; SWL: Lugano, Switzerland; FGR: Grenoble, France ; FLY: Lyon, France; GRA: Athens, Greece; ITU: Turin, Italy; IRO: Rome, Italy; SPB: Barcelona, Spain; FMA: Marseille, France.

than at the urban and rural background sites. Concentration contrasts for $PM_{2.5}$ were more similar at all the site types, suggesting an influence of multiple sources in addition to traffic.

Models in combined areas

Table 2 shows the model details of NO_2, PM and $PM_{2.5}$ absorbance combining all the European study areas. The NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance models explained 56%, 86% and 70%, respectively, of the variation across all sites, which includes both within and between area variations. The LOOCV R^2 was 1% lower than or identical to the model R^2. The LOAOCV R^2 was 5% and 6% lower than the model R^2 for NO_2 and $PM_{2.5}$, and was identical to the $PM_{2.5}$ absorbance model R^2. The hold-out validation R^2s (50% training vs. 50% test sites for NO_2, 75% training vs. 25% test sites for PM metrics) were nearly identical to the model R^2s, explaining 54%, 80% and 70% for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance at the independent validation sites respectively (see Supplemental Material, Table S3). The HV RMSE values were close to the values of LOOCV and LOAOCV RMSE for NO_2 and PM metrics. The RMSE values were relatively small compared to the range of measurements as shown in Table S3 (see Supplemental Material).

All the models in Table 2 included traffic intensity variables. The regional background concentration explained a large fraction (71%) of variation in $PM_{2.5}$ documenting the importance of between-area differences for $PM_{2.5}$ as compared to that for the more traffic-related pollutants NO_2 and $PM_{2.5}$ absorbance.

The regional models performed equally well as the European models in all regions except Southern Europe, where none of the models performed well in terms of the predictions to the independent sites (HV R^2: 0-0.23) (see Supplemental Material, Table S4).

As shown in Table 2, the median within-area variability ($Model_{intra}$ R^2) explained by the European model for NO_2 and absorbance at individual study areas was similar to the overall model R^2, suggesting predominant sources of local emissions. For $PM_{2.5}$, the median $model_{intra}$ R^2 was much lower than the overall model R^2 (0.48 vs. 0.86). Figure 2 and Figure S1 present the correlation between predicted and measured $PM_{2.5}$, $PM_{2.5}$ absorbance and NO_2 by study areas. As the figures show, the variation of $PM_{2.5}$ between areas was substantial compared to the within areas variation. On the contrary, for NO_2 and $PM_{2.5}$ absorbance, variation within areas was substantial compared to the variation between areas. The observations are more under-predicted within individual areas for PM metrics (median regression slope: 0.47 $PM_{2.5}$; 0.57 $PM_{2.5}$ absorbance; 0.56 NO_2) than across the whole European study areas (regression slope: 0.85 $PM_{2.5}$; 0.70 $PM_{2.5}$ absorbance; 0.57 NO_2).

Transferability

Table 3, shows the performance of the models which used all monitoring data excluding one area at the time. These models explained on average 57%, 84% and 69% variability of NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance respectively. The model structures and R^2s were similar to the models in Table 2 which were based on all

Table 2 European models for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance

Predictors	Partial R^2	Beta[b]	$Model_{intra}$[c] R^2/IQR	LOOCV[d] R^2/RMSE	LOAOCV[e] R^2/RMSE
NO_2 (N^a=960, final model R^2=0.56)					
Regional background concentration;	0.08	2.63E-01	0.59/0.19	0.56/11.25	0.50/8.49
Traffic load in 50m;	0.35	2.44E-06		($\mu g/m^3$)	($\mu g/m^3$)
Road length in 1000m;	0.50	2.74E-04			
Natural and Green in 5000m;	0.55	-2.84E-07			
Traffic intensity on the nearest road;	0.56	2.21E-04			
Intercept		1.38E+01			
$PM_{2.5}$ (N^a=356, final model R^2=0.86)					
Regional background concentration	0.71	9.73E-01	0.48/0.16	0.85/2.37	0.81/2.38
Traffic load between 50m and 1000m;	0.81	4.75E-09		($\mu g/m^3$)	($\mu g/m^3$)
Traffic load in 50m;	0.84	5.28E-07			
Road length in 100m;	0.86	2.12E-03			
Intercept		3.06E-01			
$PM_{2.5}$ absorbance (N^a=356, final model R^2=0.70)					
Regional background concentration;	0.28	9.06E-01	0.70/0.19	0.69/0.46	0.70/0.45
Traffic load in 50m;	0.58	2.07E-07		($10^{-5}m^{-1}$)	($10^{-5}m^{-1}$)
Road length in 500m;	0.67	2.90E-05			
Natural and Green in 5000m;	0.69	-9.63E-09			
Traffic load between 50m and 1000m;	0.70	4.20E-10			
Intercept		2.95E-01			

[a]N: number of monitored sites available for model building; [b]Beta: coefficients of predictor variables in the models;
[c]The $Model_{intra}$ R^2s show the median and Inter Quartile Range of the within-area variability explained by the European model in individual areas; [d]LOOCV: Leave-One-Out- Cross-Validation; [e]LOAOCV: Leave-One-Area-Out-Cross-Validation

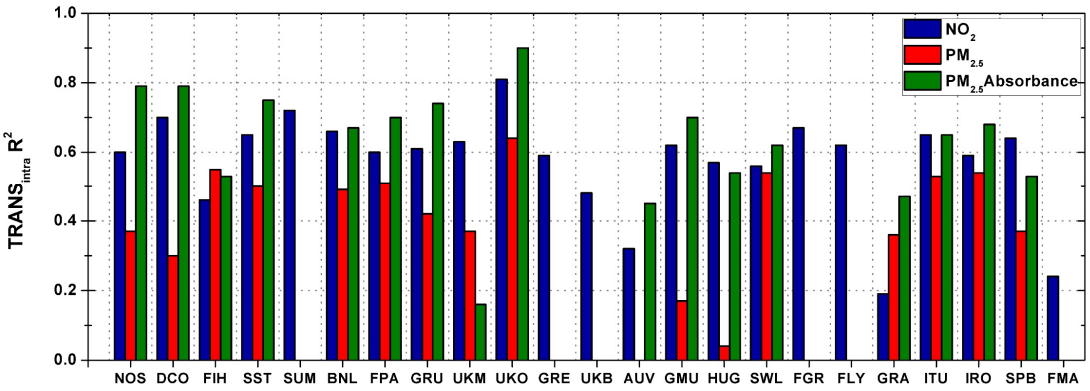


Figure 3 Transferability (TRANS$_{intra}$ R^2) of the European models for NO_2 and PM in the 23 study areas. See Figure 2 for coding reference.

Table 3 Transferability of European models to areas which were not used for model building for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance (median (IQR)).

	Model	Model$_{intra}$[a]	TRANS$_{intra}$[b]	
	R^2	R^2	R^2	RMSE
NO_2	0.57(0.01)	0.59(0.19)	0.59(0.09)	5.58(2.28)
$PM_{2.5}$	0.84(0.01)	0.48(0.16)	0.42(0.17)	1.14(0.58)
Absorbance	0.69(0.01)	0.70(0.19)	0.67(0.21)	0.23(0.07)

[a]Model$_{intra}$ R^2: R^2 of within-area variation explained by European model, the same data as in table 2

[b]TRANS$_{intra}$: squared correlations and RMSE between the predictions and observations at independent areas.

study areas. They included the same variable categories but with to some extent different buffer sizes. The models predicted the spatial variations of NO_2 and $PM_{2.5}$ absorbance well in the areas not used for model building with median TRANS$_{intra}$ R^2s of 0.59 for NO_2 and 0.67 for $PM_{2.5}$ absorbance. Transferability was less for $PM_{2.5}$ with a median R^2 of 0.42. The same pattern was found for the model R^2 focusing on within-area variability only (Model$_{intra}$). The variation in prediction R^2s was relatively small for NO_2 with an IQR of 0.09, but larger for $PM_{2.5}$ (IQR 0.17) and $PM_{2.5}$ absorbance (IQR 0.21) showing that predictions were less comparable for the two PM metrics. The variation is shown in figure 3 and Figure S2. Interestingly, this did not depend so much on area as on the specific combination of area and component. For example, the areas in Hungary (Gyor), Germany (Munich) and Austria (Vorarlberg) showed decent model fit and predictability for NO_2 and $PM_{2.5}$ absorbance, but almost no model fit and predictability for $PM_{2.5}$. The transferred regional models showed similar characteristics as those of the European models, while the median TRANS$_{intra}$ R^2 was slightly lower (see Supplemental Material, Table S5).

Discussion

In this study we developed LUR models for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance with combined measurement data from 23 study areas across Europe. For NO_2 and $PM_{2.5}$ absorbance, these models predicted spatial variations in areas not used for model building well. For $PM_{2.5}$, prediction R^2s were moderate for intra-urban variation.

Comparisons with other studies

Our European models performed comparable or even better in predictions of NO_2 and $PM_{2.5}$ than other published study results [16, 17, 19]; Beckerman et al. 2013). For $PM_{2.5}$ absorbance, this is the first report of LUR models in such a large geographical area. We observed no heterogeneity of model fit across study areas in the European model (LOAOCV R^2s were close to the model R^2). Our European and regional models have several strengths compared to local models in previous studies: 1) our sampling sites were deliberately chosen to represent real population exposures as compared to most other studies using routine monitoring stations [16-18]; 2) we incorporated local traffic intensity data not available in Europe-wide databases. All the models included traffic intensity variables, improving model fit over models not having local traffic intensity data.

Our $PM_{2.5}$ European model explained a median of 48% within-area variations as compared to the overall model R^2 of 86% which was largely explained by substantial differences in regional background concentrations. This was consistent with the R^2s of the Canadian and American $PM_{2.5}$ model (46% and 63%) of which the satellite data alone explained 41% and 52% of the variability respectively [17]. $PM_{2.5}$ is well known to be a regional pollutant with a large fraction of secondary aerosol, not explained well by the local GIS and traffic variables typically available for LUR model building.

Comparison with ESCAPE area-specific models

The ESCAPE area-specific models explained a median of 82%, 71% and 89% of the concentration variations for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance [26, 27]. This is higher than the R^2 of within-area variability explained by the European models in Table 2 ($Model_{intra}$ R^2: 59%, 48%, 70% respectively) However, we and others have shown that NO_2 and PM models based on small training sets and a large number of variables overestimate predictive ability in independent test sets [28]. The Hold Out Validation R^2 in these analyses were in the same order of magnitude as shown in this paper, documenting that with large numbers of training sites, model R^2 are actually very close to Hold Out validation R^2s with average differences of the model R^2s versus HV R^2s of just 2%, 6% and 0% for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance.

Most of the combined models included traffic variables in both large (≥500m) and small buffers (≤50m), representing general area characteristics as well as localized influences. In contrast to the study-area specific ESCAPE models [26, 27] , none of our European models included population/residence density but instead selected road length in large buffers which likely also represents urban-rural difference in terms of population distributions [29].

Transferability of combined models

Previous studies on the transferability of LUR models were mainly focusing on city-to-city or country-to-country transferability. [30] concluded that the SAVIAH models could be applied to other UK cities after calibrating with data from a few monitoring sites. Poplawski et al. (2009) and Allen et al.(2011) observed that local calibration may improve the predictions of the Canadian city specific models to a few other comparable cities in Canada and the United States. Vienneau et al. (2010) found reasonable transferability of British and the Dutch models between these two countries. Only one study in Canada investigated the transferability of national NO_2 models to explain spatial variations of NO_2 in seven specific areas (Edmonton, Montreal, Sarnia, Toronto, Victoria, Vancouver and Winnipeg) with mean $TRANS_{intra}$ R^2 of 0.43 [17]. All the previous studies concluded that the performances of the transferred models were worse than those of the local source models. Our results show prediction capabilities for the traffic related pollutants NO_2 and $PM_{2.5}$ absorbance which are at a par with those documented, in terms of Hold Out validation R^2s, with previous local exercises [14, 15]. This is likely due to the fact that the ESCAPE study used highly standardized monitoring and GIS data for model building across all areas. This suggests that our combined models can be carefully applied to other areas in Europe with common predictors, similar geographies and availability of regional background concentration of this area. Traffic related pollutant models (NO_2 and $PM_{2.5}$ absorbance) showed stronger prediction power than those of models representing mixed sources ($PM_{2.5}$). In individual areas of central Europe, the multi-cities models performed poorly however, probably due to the lack of an important local predictor variable, e.g. residential density in Munich and Vorarlberg, industry in Hungary or altitude in Lugano. Therefore, caution is need when transferring the European models to cities with other specifications beyond the explanatory variables (e.g. industrial city, harbor city).

Implications for epidemiological studies

So far, epidemiological studies have mostly used locally developed exposure models. This was not different in ESCAPE where the health findings based on these local exposure models are currently being published. A logical next step is to compare the estimates and epidemiological results with those obtained using predictions from a European model as input, because for exposure assessment with LUR models, the most efforts are mainly in sampling campaign and GIS data collection. Hence, there is also the perspective to include new study populations from areas where local measurements were never conducted but relevant predictor variables are available.

Conclusions

European LUR models for NO_2, $PM_{2.5}$ and $PM_{2.5}$ absorbance were found to have reasonable power to predict spatial variations of these components in areas not used for model building.

References

1. Brunekreef, B.; Holgate, S. T., Air pollution and health. *Lancet* **2002,** *360*, (9341), 1233-42.
2. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.;et al., A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* **2008,** *42*, (33), 7561-7578.
3. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; et al., A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology* **2005,** *15*, (2), 185-204.
4. Bellander, T.; Berglind, N.; Gustavsson, P.; Jonson, T.; Nyberg, F.; Pershagen, G.; et al., Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environmental Health Perspectives* **2001,** *109*, (6), 633-639.
5. Gehring, U.; Wijga, A. H.; Fischer, P.; de Jongste, J. C.; Kerkhof, M.; Koppelman, G. H.; et al., Traffic-related air pollution, preterm birth and term birth weight in the PIAMA birth cohort study. *Environmental Research* **2011,** *111*, (1), 125-135.
6. Cesaroni, G.; Badaloni, C.; Gariazzo, C.; Stafoggia, M.; Sozzi, R.; Davoli, M.; et al., Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ Health Perspect* **2013,** *121*, (3), 324-31.
7. Estarlich, M.; Ballester, F.; Aguilera, I.; Fernandez-Somoano, A.; Lertxundi, A.; Llop, S.; et al., Residential exposure to outdoor air pollution during pregnancy and anthropometric measures at birth in a multicenter cohort in Spain. *Environ Health Perspect* **2011,** *119*, (9), 1333-8.
8. Madsen, C.; Carlsen, K. C. L.; Hoek, G.; Oftedal, B.; Nafstad, P.; Meliefste, K.; et al., Modeling the intra-urban variability of outdoor traffic pollution in Oslo, Norway - A GA(2)LEN project. *Atmospheric Environment* **2007,** *41*, (35), 7500-7511.
9. Marshall, J. D.; Nethery, E.; Brauer, M., Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmospheric Environment* **2008,** *42*, (6), 1359-1369.
10. Hoek, G.; Beelen, R.; Kos, G.; Dijkema, M.; van der Zee, S. C.; Fischer, P. H.; et al., Land Use Regression Model for Ultrafine Particles in Amsterdam. *Environmental Science & Technology* **2011,** *45*, (2), 622-628.
11. Su, J. G.; Jerrett, M.; Beckerman, B.; Wilhelm, M.; Ghosh, J. K.; Ritz, B., Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. *Environmental Research* **2009,** *109*, (6), 657-670.
12. Henderson, S. B.; Beckerman, B.; Jerrett, M.; Brauer, M., Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology* **2007,** *41*, (7), 2422-2428.
13. Brauer, M.; Hoek, G.; van Vliet, P.; Meliefste, K.; Fischer, P.; Gehring, U.; et al., Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* **2003,** *14*, (2), 228-39.
14. Wang, M.; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B., Systematic Evaluation of Land Use Regression Models for NO_2. *Environ Sci Technol* **2012,** *46*, (8), 4481-4489.
15. Basagana, X.; Rivera, M.; Aguilera, I.; Agis, D.; Bouso, L.; Elosua, R.; et al., Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment* **2012,** *54*, 634-642.
16. Vienneau, D.; de Hoogh, K.; Briggs, D., A GIS-based method for modelling air pollution exposures across Europe. *Sci Total Environ* **2009,** *408*, (2), 255-266.
17. Hystad, P.; Setton, E.; Cervantes, A.; Poplawski, K.; Deschenes, S.; Brauer, M.; et al., Creating National Air Pollution Models for Population Exposure Assessment in Canada. *Environmental Health Perspectives* **2011,** *119*, (8), 1123-1129.
18. Hart, J. E.; Yanosky, J. D.; Puett, R. C.; Ryan, L.; Dockery, D. W.; Smith, T. J.; et al., Spatial Modeling of PM_{10} and NO_2 in the Continental United States, 1985-2000. *Environmental Health Perspectives* **2009,** *117*, (11), 1690-1696.
19. Beelen, R.; Hoek, G.; Pebesma, E.; Vienneau, D.; de Hoogh, K.; Briggs, D. J., Mapping of background air pollution at a fine spatial scale across the European Union. *Sci Total Environ*

2009, *407*, (6), 1852-1867.
20. Liu, L. J. S.; Tsai, M. Y.; Keidel, D.; Gemperli, A.; Ineichen, A.; Hazenkamp-von Arx, M.; et al., Long-term exposure models for traffic related NO_2 across geographically diverse areas over separate years. *Atmospheric Environment* **2012,** *46*, 460-471.
21. Poplawski, K.; Gould, T.; Setton, E.; Allen, R.; Su, J.; Larson, T.; et al., Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. *J Expo Sci Env Epid* **2009,** *19*, (1), 107-117.
22. Allen, R. W.; Amram, O.; Wheeler, A. J.; Brauer, M., The transferability of NO and NO_2 land use regression models between cities and pollutants. *Atmospheric Environment* **2011,** *45*, (2), 369-378.
23. Vienneau, D.; de Hoogh, K.; Beelen, R.; Fischer, P.; Hoek, G.; Briggs, D., Comparison of land-use regression models between Great Britain and the Netherlands. *Atmos Environ* **2010,** *44*, (5), 688-696.
24. Cyrys, J.; Eeftens, M.; Heinrich, J.; Ampe, C.; Armengaud, A.; Beelen, R.; et al., Variation of NO_2 and NO_x concentrations between and within 36 European study areas: results from the ESCAPE study. *Atmos Environ* **2012,** *62*, 374 - 390.
25. Eeftens, M.; Tsai, M.; Ampe, C.; Anwander, B.; Beelen, R.; Cesaroni, G.; et al., Spatial variation of $PM_{2.5}$, PM_{10}, $PM_{2.5}$ absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO_2 - Results of the ESCAPE project. *Atmospheric Environment* **2012,** *62*, 303 - 317.
26. Eeftens, M.; Beelen, R.; de Hoogh, K.; Bellander, T.; Cesaroni, G.; Cirach, M.; et al., Development of land use regression models for $PM_{2.5}$, $PM_{2.5}$ absorbance, PM_{10} and PMcoarse in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* **2012**, 11195-11205.
27. Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; et al., Development of NO_2 and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe - the ESCAPE project. *Atmos Environ* **2013,** *72*, 10 - 23.
28. Wang, M.; Beelen, R.; Basagana, X.; Becker, T.; Cesaroni, G.; de Hoogh, K.; et al., B., Evaluation of Land Use Regression Models for NO_2 and Particulate Matter in 20 European Study Areas: The ESCAPE Project. *Environ Sci Technol* **2013,** *47*, (9), 4357-64.
29. de Hoogh, K.; Wang, M.; Adam, M.; Badaloni, C.; Beelen, R.; Birk, M.; et al., Development of land use regression models for particle composition in 20 study areas in Europe. *Environ Sci Technol* **2013**.
30. Briggs, D. J.; de Hoogh, C.; Guiliver, J.; Wills, J.; Elliott, P.; Kingham, S.; et al., A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Sci Total Environ* **2000,** *253*, (1-3), 151-167.

Chapter 4

Appendix

Table S1 Coding of the study areas and regional background concentration levels

Code	Type	Region	Study Area
NOS	PM/NO_2	North	Oslo, Norway
SST	PM/NO_2	North	Stockholm, Sweden
FIH	PM/NO_2	North	Helsinki, Turku, Finland
DCO	PM/NO_2	North	Copenhagen, Denmark
SUM	NO_2	North	Umea,Sweden
UKM	PM/NO_2	West	Manchester, UK
UKO	PM/NO_2	West	London, Oxford, UK
BNL	PM/NO_2	West	Netherlands &Belgium
GRU	PM/NO_2	West	Ruhr area, Germany
GRE	NO_2	West	Erfurt, Germany
UKB	NO_2	West	Bradford, UK
FPA	PM/NO_2	West	Paris, France
GMU	PM/NO_2	Central	Munich, Germany
AUV	PM/NO_2	Central	Vorarlberg, Austria
FLY	NO_2	Central	Lyon, France
HUG	PM/NO_2	Central	Györ, Hungary
SWL	PM/NO_2	Central	Lugano, Switzerland
FGR	NO_2	Central	Grenoble, France
ITU	PM/NO_2	South	Turin, Italy
IRO	PM/NO_2	South	Rome, Italy
SPB	PM/NO_2	South	Barcelona, Spain
FMA	NO_2	South	Marseille, France
GRA	PM/NO_2	South	Athens, Greece

Table S2 List of predictor variables for model development, buffer sizes and a priori defined direction of effect

Region[a]	Variable	Buffer size (m)	Direction
All	High and low residential density	100, 300, 500, 1000, 5000	+
All	Port	300, 500, 1000, 5000	+
All	Industry	300, 500, 1000, 5000	+
All	Urban green and natural areas	100, 300, 500, 1000, 5000	-
All	Squared root of altitude	-	-
All	Road length	50, 100, 300, 500, 1000	+
All	Major road length	50, 100, 300, 500, 1000	
All	Traffic intensity in the nearest road	NA	+
All	(Squared) Inverse distance to the nearest road	NA	+
All	(Squared) Invest distance to the nearest road*traffic intensity in the nearest road	NA	+
All	Traffic intensity in the major road	NA	+
All	(Squared) Inverse distance to the nearest major road	NA	+
All	(Squared) Invest distance to the major road *traffic intensity in the major road	NA	+
All	Total traffic load of roads in a buffer (sum of (traffic intensity * length of all segments))	50, 100, 300, 500, 1000	+
All	Total traffic load of major roads in a buffer (sum of (traffic intensity * length of all segments))	50, 100, 300, 500, 1000	+
NE,WE, SE	Population	100, 300, 500, 1000, 5000	+
CE,SE	Urban green	100, 300, 500, 1000, 5000	-
CE,SE	Natural areas	100, 300, 500, 1000, 5000	-
SE	High residential density	100, 300, 500, 1000, 5000	+
SE	Low residential density	100, 300, 500, 1000, 5000	+

[a]All: all study areas; NE: north Europe; WE: west Europe; CE: central Europe; SE: south Europe

Table S3 Descriptive of European model performances for NO_2 and PM metrics using 50% NO_2 training sets and 75% PM training sets for modeling and the remaining 50% and 25% test sets for hold-out validation

Model	N[a]	Determinants	Partial R^2	Beta	LOOCV R^2/RMSE	HV[c] R^2/RMSE
NO_2	480	Regional background concentration ;	0.08	3.36E-01	0.55/10.35	0.54/11.20
		Traffic load in 50m;	0.37	2.60E-06	($\mu g/m^3$)	($\mu g/m^3$)
		Road length in 1000m;	0.52	2.65E-04		
		Natural and Green in 5000m;	0.55	-2.19E-07		
		Traffic intensity on the nearest road;	0.57	1.90E-04		
		Intercept		1.10E+01		
$PM_{2.5}$	270	Regional background concentration;	0.71	9.63E-01	0.85/2.27	0.80/2.78
		Traffic load between 50m and 1000m;	0.82	5.37E-09	($\mu g/m^3$)	($\mu g/m^3$)
		Road length in 50m;	0.84	6.89E-03		
		Traffic load in 50m;	0.86	4.94E-07		
		Intercept		4.72E-01		
$PM_{2.5}$ ABS[b]	270	Regional background concentration;	0.29	9.58E-01	0.68/0.47	0.70/0.45
		Traffic load in 50m;	0.56	2.13E-07	($10^{-5}m^{-1}$)	($10^{-5}m^{-1}$)
		Road length in 500m;	0.66	3.53E-05		
		Industry in 5000m;	0.68	2.50E-08		
		Natural and Green in 5000m;	0.69	-8.65E-09		
		Intercept		1.11E-01		

[a]N: number of training sites for modeling; $PM_{2.5}$ ABS: $PM_{2.5}$ absorbance

[c]The HV R^2s represent the correlation between predicted and measured concentrations at validation monitoring sites not used for model building (50% for NO_2, 25% for PM metrics, see methods section).

Table S4 Descriptive of model performances at regional scales using full number of sites

Regi-on[a]	Determinants	Partial R^2	Beta	$Model_{intra}$ R^2/IQR[c]	LOOCV R^2	LAOCV R^2	HV[d] R^2
NE	**NO_2 (N^b=200, final model R^2=0.61)**						
	Regional background concentration;	0.20	9.75E-01	0.63/ 0.15	0.63	0.52	0.57
	Traffic load between 50 and 300m	0.48	8.45E-08				
	Traffic load in 50m	0.55	2.64E-06				
	Road length in 1000m;	0.60	1.19E-04				
	Traffic load in 300 and 1000m;	0.61	2.06E-08				
	Intercept		2.34E-01				
	$PM_{2.5}$(N^b=78, final model R^2=0.70)						
	Regional background concentration;	0.28	5.39E-01	0.68/0.25	0.66	0.59	0.60
	Natural and Green in 1000m;	0.64	-1.03E-06				
	Traffic density*inverse distance to the	0.67	2.04E-04				
	nearest road;	0.69	1.28E-04				
	Road length between 50 and 500m;	0.70	9.17E-03				
	Major road length in 50m;		4.26E+00				
	Intercept						
	$PM_{2.5}$absorbance(N^b=78, final model R^2=0.69)						
	Regional background concentration;	0.12	6.77E-01	0.80/0.11	0.62	0.02	0.69
	Traffic load in 50m;	0.50	1.12E-07				
	Road length in 500m;	0.59	2.26E-05				
	Natural and green in 5000m;	0.64	-1.00E-08				
	Inverse distance to major road;	0.69	1.49E+00				
	Intercept		5.57E-01				

Table S4 Descriptive of model performances at regional scales using full number of site***(Continued)***

Region[a]	Determinants	Partial R^2	Beta	$Model_{intra}$ R^2/IQR[c]	LOOCV R^2	LAOCV R^2	HV[d] R^2
WE	**NO_2 (N^b=320, final model R^2=0.64)**						
	Regional background concentration;	0.00	-2.55E-02	0.65/0.29	0.61	0.54	0.64
	Traffic load in 50m;	0.41	4.89E-06				
	Population in 1000m;	0.58	2.88E-04				
	Squared altitude;	0.62	-6.02E-01				
	Major road length in 500m;	0.64	1.37E-03				
	Intercept		2.37E+01				
	$PM_{2.5}$ (N^b=119, final model R^2=0.80)						
	Regional background concentration;	0.68	7.35E-01	0.48/0.13	0.78	0.71	0.71
	Major road length in 50m;	0.79	1.47E-02				
	Industry in 5000m;	0.80	1.07E-07				
	Intercept		4.42E+00				
	$PM_{2.5}$absorbance(N^b=119, final model R^2=0.75)						
	Regional background concentration;	0.01	6.51E-02	0.80/0.10	0.70	0.68	0.74
	Traffic load in 50m;	0.56	2.78E-07				
	Major road length in 1000m;	0.69	1.47E-05				
	Population in 1000m;	0.73	8.33E-06				
	Traffic load in major roads in 500m;	0.75	2.06E-09				
	Intercept		1.03E+00				
CE	**NO_2 (N^b=240, final model R^2=0.63)**						
	Traffic load in 1000m;	0.54	5.63E-08	0.57/0.10	0.61	0.36	0.56
	Traffic intensity to the nearest road;	0.60	2.74E-04				
	Road length in 50m;	0.63	2.02E-02				
	$PM_{2.5}$ (N^b=79, final model R^2=0.82)						
	Regional background concentration;	0.72	1.17E+00	0.25/0.48	0.79	0.34	0.84
	Road length in 50m;	0.81	8.44E-03				
	Traffic load in 100m;	0.82	1.76E-07				
	Intercept		-2.61E+00				
	$PM_{2.5}$absorbance(N^b=79, final model R^2=0.61)						
	Regional background concentration;	0.00	8.70E-01	0.63/0.06	0.55	0.55	0.15
	Traffic load in major roads in 50m;	0.38	1.82E-07				
	Road length in 300m;	0.53	1.05E-04				
	Natural and Green in 5000m	0.61	-1.62E-08				
	Intercept		4.19E-01				
SE	**NO_2 (N^b=200, final model R^2=0.75)**						
	Regional background concentration;	0.00	-1.22E+00	0.63/0.25	0.70	0.12	0.23
	Low residual density in 5000m;	0.53	5.42E-07				
	Population in 1000m;	0.65	1.85E-04				
	Traffic intensity to the major road;	0.70	3.00E-04				
	Road length in 50m;	0.75	2.90E-02				
	Intercept		1.53E+01				
	$PM_{2.5}$ (N^b=80, final model R^2=0.23)						
	Road length in 100m;	0.10	3.91E-03	0.50/0.13	0.17	0.00	0.00
	Traffic density in nearest road;	0.23	1.56E-04				
	Intercept		1.69E+01				
	$PM_{2.5}$absorbance(N^b=80, final model R^2=0.59)						
	Regional background concentration;	0.01	9.23E-04	0.67/0.08	0.53	0.42	0.16
	Traffic density in nearest road;	0.42	2.15E-05				
	Natural in 5000m;	0.53	-3.46E-08				
	Major road length in 50m;	0.59	3.50E-03				
	Intercept		2.59E+00				

[a]NE: north Europe; WE: west Europe; CE: central Europe; SE: south Europe; [b]N: number of training sites for modeling;

[c]The $Model_{intra}$ R^2s show the median and Inter Quartile Range of the within-area variability explained by the Regional model in individual areas

[d]The HV R^2s represent the correlation between predicted and measured concentrations at validation monitoring sites not used for model building (50% for NO_2, 25% for PM metrics, see methods section).

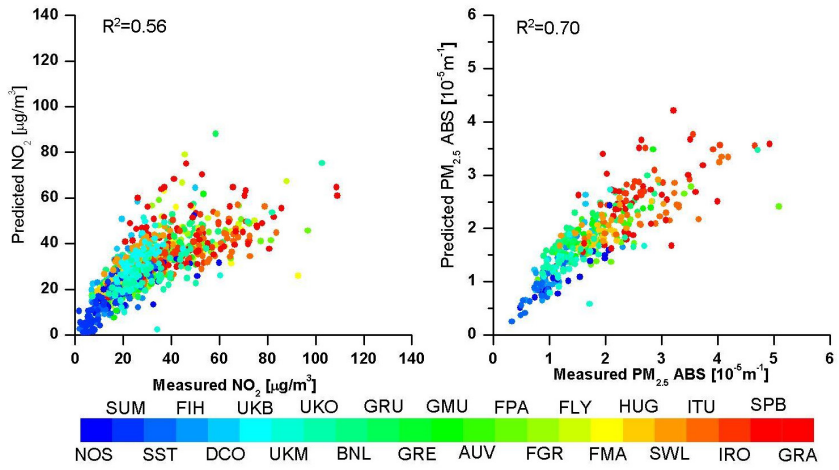


Figure S1 Scatterplot of predicted and measured of NO_2 and $PM_{2.5}$ absorbance with study areas color coded. Coding of areas please see table S1.

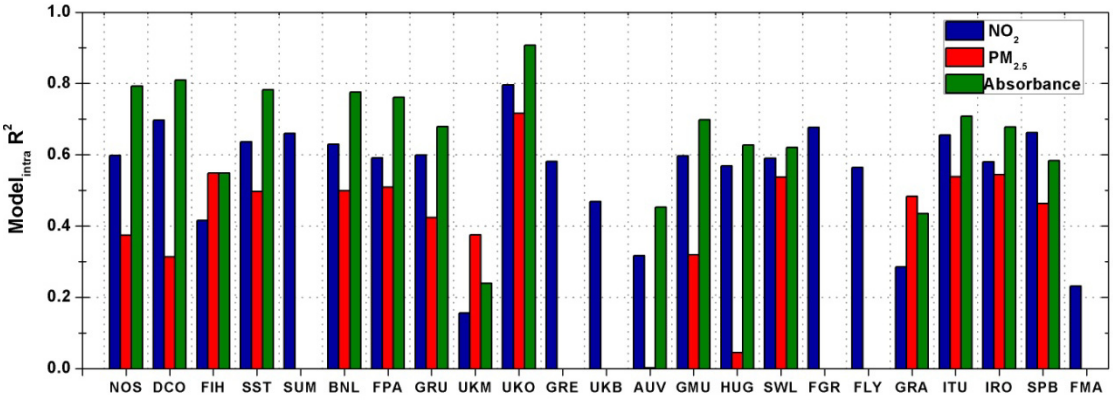


Figure S2 $Model_{intra}$ R^2) of the European models for NO_2 and PM in the 23 study areas. Coding of areas please see table S1

Table S5 Transferability of the regional models to the independent areas not used for model building (Median(IQR))

	Region	Regional model	
		Model(R^2)	$TRANS_{intra}$ (R^2) [a]
NO_2	North	0.67(0.00)	0.71(0.42)
	West	0.68(0.00)	0.69(0.16)
	Central	0.68(0.00)	0.54(0.25)
	South	0.65(0.00)	0.43(0.25)
	All[b]	0.68(0.01)	0.58(0.32)
$PM_{2.5}$	North	0.69(0.04)	0.36(0.35)
	West	0.82(0.01)	0.40(0.19)
	Central	0.86(0.07)	0.12(0.21)
	South	0.71(0.22)	0.31(0.22)
	All[b]	0.77(0.17)	0.32(0.28)
$PM_{2.5}$ absorbance	North	0.69(0.00)	0.55(0.41)
	West	0.75(0.00)	0.77(0.30)
	Central	0.61(0.00)	0.52(0.19)
	South	0.59(0.00)	0.40(0.18)
	All[b]	0.69(0.14)	0.49(0.39)

[a]$TRANS_{intra}$: squared correlations between the predictions and observations at independent areas.
[b]All: Median and interquartile range of regional model R^2s and $TRANS_{intra}$ R^2s in all the study area

Chapter 5

Development of Land Use Regression Models for Particle Composition in 20 Study Areas in Europe

Kees de Hoogh, Meng Wang, Martin Adam, Chiara Badaloni, Rob Beelen, Matthias Birk, Giulia Cesaroni, Marta Cirach, Christophe Declercq, Audrius Dėdelė, Evi Dons, Audrey de Nazelle, Marloes Eeftens, Kirsten Eriksen, Charlotte Eriksson, Paul Fischer, Regina Grazuleviciene, Alexandros Gryparis, Barbara Hoffmann, Michael Jerrett, Klea Katsouyanni, Minas Iakovides, Timo Lanki, Sarah Lindley, Christian Madsen, Anna Mölter, Gioia Mosler, Gizella Nádor, Mark Nieuwenhuijsen, Göran Pershagen, Annette Peters, Harisch Phuleria, Nicole Probst-Hensch, Ole Raaschou-Nielsen, Ulrich Quass, Andrea Ranzi, Euripides Stephanou, Dorothea Sugiri, Per Schwarze, Ming-Yi Tsai, Tarja Yli-Tuomi, Mihály J Varró, Danielle Vienneau, Gudrun Weinmayr, Bert Brunekreef, Gerard Hoek

(Environmental Sciences & Technology 2013, 47, (11), 5778-5786)

Abstract

Land Use Regression (LUR) models have been used to describe/model spatial variability of annual mean concentrations of traffic related pollutants like nitrogen dioxide (NO_2), nitrogen oxides (NO_x) and particulate matter (PM). No models have yet been published of elemental composition. As part of the ESCAPE project, we measured the elemental composition in both the PM_{10} and $PM_{2.5}$ fraction sizes at 20 sites in each of 20 study areas across Europe. LUR models for eight a priori selected elements; copper (Cu), iron (Fe), potassium (K), nickel (Ni), sulphur (S), silicon (Si), vanadium (V) and zinc (Zn) were developed. Good models were developed for Cu, Fe and Zn in both fractions (PM_{10} and $PM_{2.5}$) explaining on average between 67 and 79% of the concentration variance (R^2) with a large variability between areas. Traffic variables were the dominant predictors, reflecting non-tailpipe emissions. Models for V and S in the PM_{10} and $PM_{2.5}$ fractions and Si, Ni and K in the PM_{10} fraction performed moderately with R^2 ranging from 50 to 61%. Si, NI and K models for $PM_{2.5}$ performed poorest with R^2 under 50%. The LUR models are used to estimate exposures to elemental composition in the health studies involved in ESCAPE.

Table of content

Introduction

Associations between long term air pollution and health effects have been widely reported.[1-3] The influence of road traffic related emissions on cardio-respiratory morbidity and mortality is well documented.[4] Most epidemiological studies have reported associations of mortality and hospital admissions due to respiratory and cardiovascular disease with particulate matter (PM) characterized as the mass concentration of particles <10μm or <2.5μm (PM_{10} and $PM_{2.5}$, respectively). Particulate matter is a complex mixture varying spatially and temporally in chemical composition and particle size related to the sources from which they originate. A major uncertainty is which components within the PM mixtures are responsible for the health effects.[5,6] Components for which associations have been reported include (transition) metals, elemental carbon, inorganic secondary aerosols (sulfate, nitrate) and organic components but the evidence is not consistent.[5,6] Most studies that have assessed health effects related to elemental composition have been short-term exposure studies using e.g. data for the US Chemical Speciation network from a few sites[7,8] or specially designed short-term campaigns.[9] Very few studies have assessed health effects related to long-term exposure to elemental composition.[10] Lack of spatially resolved elemental composition measurement data and a lack of models for elemental composition have contributed to this gap. So far the emphasis of epidemiological research into traffic-related air pollution and health has focused on NO_2, the soot content of PM and ultrafines, all reflecting exhaust emissions.[4] There is increasing concern about non-tailpipe emissions including brake and tyre wear which may result in high concentrations of transition metals such as Cu and Fe.[5]

Land use regression has been used to model the spatial variation of the annual mean concentrations for a range of traffic-related pollutants including NO_2, NO_x, PM_{10}, $PM_{2.5}$, the soot (EC) content of $PM_{2.5}$ and VOCs.[11] More recently LUR has been used to predict the spatial variability of polycyclic aromatic hydrocarbons (PAHs)[12] but to our knowledge not for elemental composition.

Within the Framework of the European Study of Cohorts for Air Pollution Effects (ESCAPE), we measured intra-urban spatial variation of NO_2, PM_{10}, $PM_{2.5}$ and its elemental composition in 20 study areas across Europe. LUR models were developed based on these measurements. LUR models for NO_2, PM_{10}, $PM_{2.5}$ and $PM_{2.5}$ absorbance have been published.[13, 14] ESCAPE is an EU wide study investigating the relationship between traffic-related pollution and health, using existing cohorts to which a harmonized exposure assessment is applied. The exposures derived from these LUR models will be used to explore associations between these elements, alone or in combination, and specific health outcomes.

The aim of this paper is to describe the development and performance of LUR models in 20 study areas across Europe, for 8 a priori selected trace elements: copper (Cu), iron (Fe), potassium (K), nickel (Ni), vanadium (V), sulphur (S), silicon (Si) and zinc (Zn) in both the $PM_{2.5}$ and PM_{10} size fractions.

Methods

In ESCAPE we a priori selected 8 from the 48 measured elements for further epidemiological evaluation. The set of elements was selected based upon evidence for health effects (toxicity), representation of major anthropogenic sources, a high percentage of detected samples (>75%) and good precision of measurements. We selected Cu, Fe and Zn mainly for (non-tailpipe) traffic emissions; S for long-range transport; Ni and V for mixed oil burning / industry; Si for crustal material and K for biomass burning.[15] Elements may have multiple sources, so they do not necessarily represent single sources.

Sampling and analysis

A particulate matter monitoring campaign was conducted in 20 study areas across Europe between October 2008 and April 2011 (see table of content). The monitoring campaign has been described in detail by Eeftens et al,[16] including a description of the study areas and the monitoring equipment used. Briefly, measurements for each area were conducted over a one year period, obtained during three 2-weekly periods across a year to capture seasonal variations. The spatial variation of PM concentrations across the study areas was measured with 20 monitoring sites. These were positioned close to roads (traffic sites), in urban areas away from roads (urban background) and in rural settings (rural background) with on average 12, 6 and 2 of these site types respectively. A common sampling protocol was used for the monitoring site selection. Measurements were performed simultaneously at five sites. One reference site, located in an urban or rural background location depending on the study area, was established to measure continuously for 2-week periods during the year to adjust for temporal variation. This site was used in the calculation of the temporally adjusted annual average concentrations for each of the 20 monitoring sites

Monitoring was performed using the Harvard impactor, which collects particle matter smaller than 2.5 µm ($PM_{2.5}$) and 10 µm (PM_{10}) on separate filters using an air flow of appr. 10 l/min. All PM_{10} and $PM_{2.5}$ samples were analysed for elemental composition using Energy Dispersive X-ray fluorescence (XRF). Analyses were performed in Cooper Environmental Services, Portland USA. Filters were analysed between December 2010 and July 2011, after a first series of analyses on June 2010 to test adequate detection. Forty-eight elements were measured. Quality assurance and control included analysis of NIST reference material (SRM 1228 and SRM 987), repeated analysis of a multi-elemental quality control standard (Multi 30585) and replicate analysis of about 10% of the samples. All analysis batches passed the quality criteria of the laboratory. In addition, about 20 field blanks and field duplicates were taken in each study area. From the field blanks we calculated the mean field blank and the detection limit (DL). Field duplicates were used to calculate the precision of measurements expressed as coefficient of variation.[16]

Concentrations were calculated by multiplying the reported concentration of an element (µg/cm^2) with the exposed filter area (7.8 cm^2), subtracting the study area-specific mean field blank and dividing by the sample volume. Limits of

detection per study area were calculated as three times the standard deviation of field blanks divided by the nominal sample volume of 25.2 m^3. Concentration values of individual samples below the DL were retained and not replaced with a standard value. In these calculations we removed 6 of the about 400 blanks, because of extreme values.

Annual averages were calculated after adjusting for temporal variation measured at the continuous reference site. For each of the three sampling periods, the absolute difference of the concentration measured in that period from the overall annual mean at the reference site was used as adjustment, following procedures for PM_{10}, $PM_{2.5}$ and soot.[13,16] In a few cases extreme concentrations measured at the reference site resulted in negative adjusted average concentrations for individual elements. Log-transformation of the concentrations did not resolve this problem. We decided to exclude these extreme sampling periods from the analysis per element. Extreme was defined as an elemental concentration at the reference site higher than four times the interquartile range above the 75th percentile of the reference site measurements or below the 25th percentile. Outliers were detected for 31 of in total 8320 (20 areas x 26 periods x 16 elements) sampling periods. In 11 study areas, PM sampling occasionally failed at the reference site. Elemental concentrations were estimated using routine monitoring sites,[16] provided that the squared correlation between the measured element and the routinely measured component was higher than 0.50. With the exception of the London study area, no elemental composition was available from routine networks and we therefore used PM_{10}, $PM_{2.5}$, NO_2 and NO_x. Particularly the non-traffic elements could not be well predicted and were left missing. Site-specific averages were only calculated if two or three valid adjusted samples were obtained. These procedures resulted in 92 missing values for all elements, sites and areas (7040 in total).

GIS predictor data

The geographical location of the monitoring sites was determined using a combination of GPS readings at the site and manual corrections using detailed local road maps to ensure the correct position of the site in relation to the road network.

Potential predictor variables for LUR model development covering a range of potential emission sources were extracted for all study areas, at the monitoring site locations, using the geographical information system (GIS) ArcGIS 9.3 & 10 (ESRI, Redlands, CA, USA). A detailed description of the geographical input data used can be found in Eeftens et al.[13] In brief, geographical data was split into centrally and locally sourced data. Central datasets were obtained at the European level for all study areas including information on roads (EuroStreets version 3.1), land use (CORINE land cover 2000), altitude (SRTM 90m Digital Elevation Data) and population (enhanced EEA population density data using CORINE land cover 2000). Where available, local data was obtained by the individual centres and included data on traffic flows, more detailed land use and emission data.

Road and traffic intensity variables were extracted in circular buffers of 25, 50, 100, 300, 500 and 1000m reflecting the local influence of these sources on air pollution levels. Land cover variables were extracted in buffer distances of 100

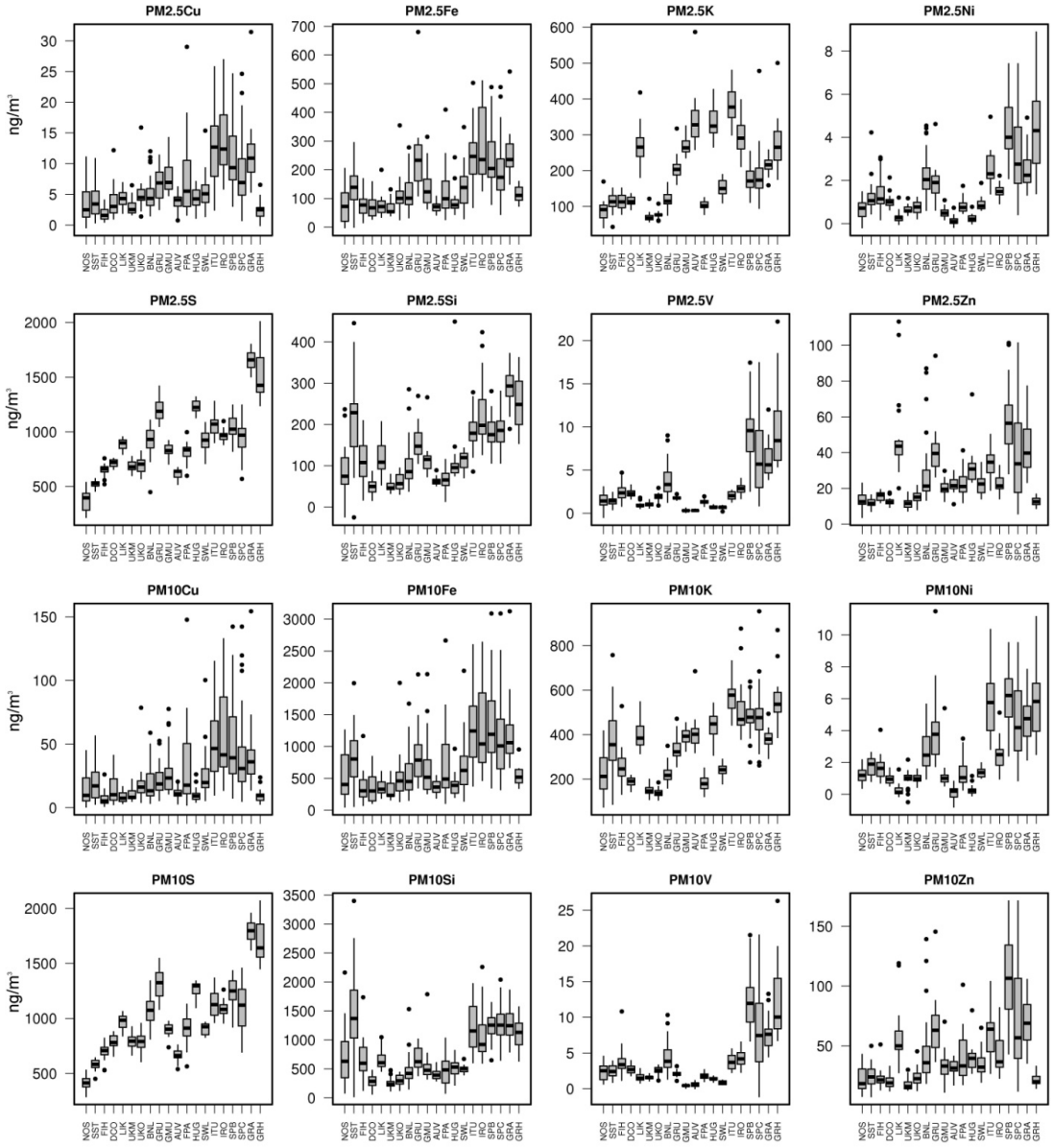


Figure 1 Annual mean concentrations (ng/m^3) for 8 elements in $PM_{2.5}$ (top) and PM_{10} (bottom) fractions – study areas[a] are shown from north to south (note the different scales on the y-axis for these elements between $PM_{2.5}$ and PM_{10} fraction).

[a]NOS: Oslo (Norway); SST: Stockholm County (Sweden); FIH: Helsinki/Turku (Finland); DCO: Copenhagen (Denmark); LIK: Kaunas (Lithuania); UKM: Manchester (United Kingdom); UKO: London/Oxford (United Kingdom); BNL: Netherlands/ Belgium; GRU: Ruhr Area (Germany); GMU: Munich/Augsburg (Germany); AUV: Vorarlberg (Austria); FPA: Paris (France); HUG: Gyor (Hungary); SWL: Lugano (Switzerland); ITU: Turin (Italy); IRO: Rome (Italy); SPB: Barcelona (Spain); SPC: Catalunya (Spain); GRA: Athens (Greece); GRH: Heraklion (Greece)

, 300, 500, 1000 and 5000m. Buffer distances of 1000 and more were included to reflect regional influences, not picked up by the smaller buffers. In addition, indicators variables such as distance to the nearest (main) road and traffic intensity on the nearest road were computed and included in the database. A descriptive table of all variables can be found in the Supporting Information Table S1.

LUR model development

Land use regression (LUR) models for the 8 elements in both fractions (PM_{10} and $PM_{2.5}$) were developed centrally at IRAS, NL for the individual 20 study areas using a standardised approach programmed in SAS 9.2. Eeftens et al. describes in detail the supervised stepwise selection procedure used to develop the linear regression models.[13] In brief, predictors giving the highest adjusted R^2 were subsequently added to the model if they conformed to the direction of effect defined a priori and added more than 1% to the adjusted R^2. The final models were checked for *p*-value (removed when *p*-value > 0.10), co-linearity (variables with Variance Inflation Factor (VIF) > 3 were removed and model rerun) and influential observations (models with Cook's D > 1 were further examined). The final models were evaluated by leave-one-out cross validation (LOOCV) and Morans's I to indicate possible spatial autocorrelation in the residuals.

We assessed the effect of seasonality by calculating the variability of the ratio of PM_{10}Cu concentrations between traffic sites and urban background sites across seasons and by evaluating the correlation between the PM_{10}Cu measurements in the different seasons. To closer investigate differences between the PM models we applied the best PM_{10} model structure to the $PM_{2.5}$ data and compared the model performances with the best $PM_{2.5}$ models. We also evaluated the correlation between the predicted elemental concentrations and the previous developed PM and absorbance models at the 20 (and in Belgium/Netherlands and Barcelona 40) NO_2/NO_x sites. For this purpose we developed new NO_2 and NO_x models using the PM sites.

Results

Measurements

The eight selected elements for the development of land use regression models were detected in the large majority (>80%) of the samples. Precision was best for S, Cu and Fe but poorer for Ni and V especially in study areas with low concentration levels (Tables S3 and S4).

Box plots of the $PM_{2.5}$ and PM_{10} fractions of Cu, Fe, K, Ni, S, Si, V and Zn annual mean concentrations measured in the 20 study areas are shown in figure 1 and pollutant ranges are shown in Table S2. A more detailed interpretation of the measured concentrations will be published separately. In brief, substantial variations of annual mean concentrations were observed within and between the majority of study areas and elements. The largest within-area contrasts were found for Cu, Fe, Si and Zn, with the largest contrasts generally found in South European study areas. The exceptions were Si and K where the largest within area contrasts were found in Oslo, Stockholm and Helsinki/Turku.The smallest within-area

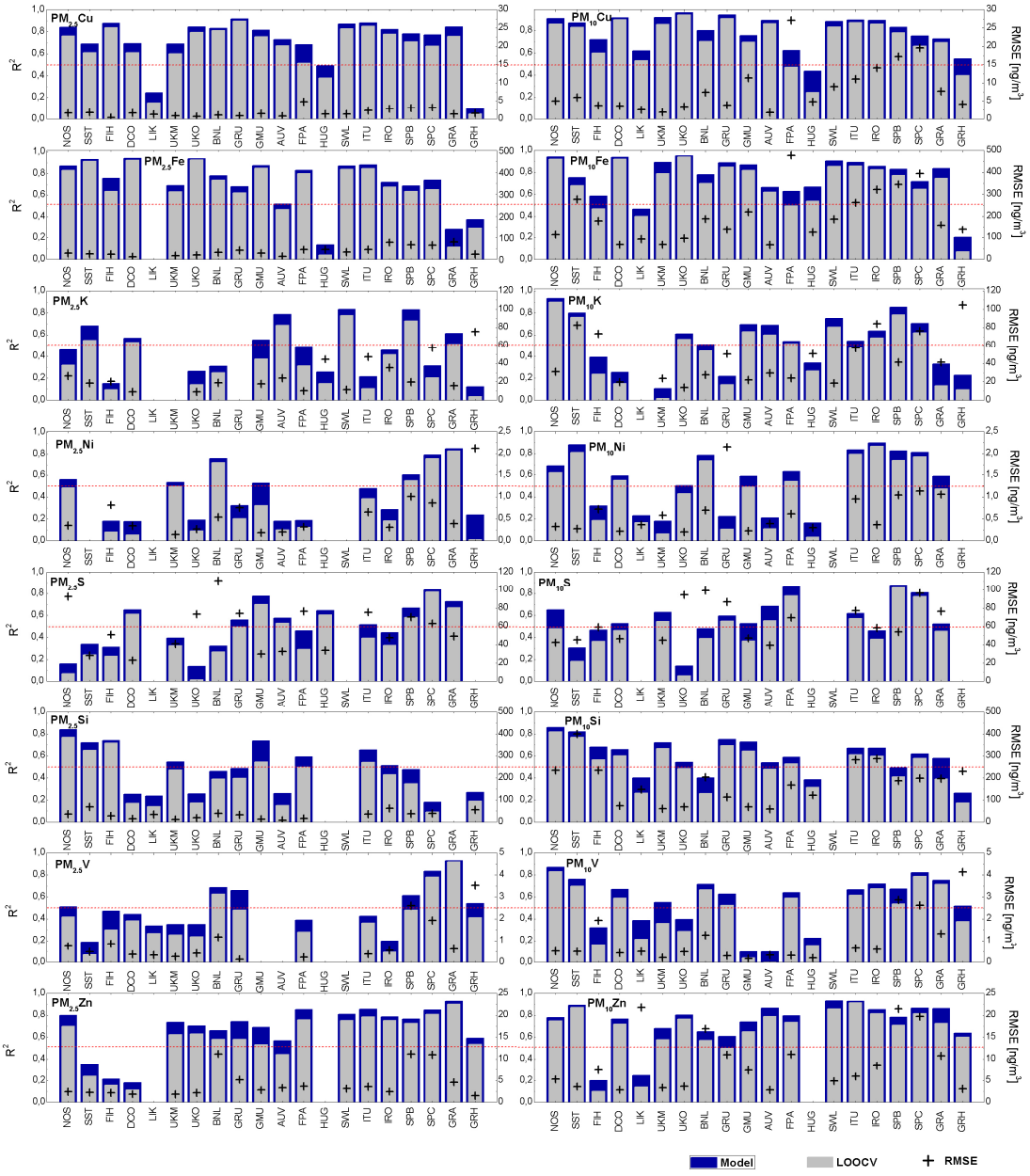


Figure 2 LUR model performance and evaluation model building R^2, LOOCV R^2 and RMSE for $PM_{2.5}$ (left) and PM_{10} (right) – study areas are shown from north to south.

contrasts were found for S. In addition, a positive north south gradient was observed with higher concentrations of most elements measured in southern study areas. Measured annual mean S in $PM_{2.5}$ concentrations, for example, show a steady increasing north-south gradient with averages from 369 ng/m^3 measured in Oslo, Norway to 1657 ng/m^3 in Athens, Greece. The median $PM_{2.5}/PM_{10}$ fraction for all study areas were 0.30, 0.21, 0.56, 0.62, 0.91, 0.18, 0.72 and 0.64 for Cu, Fe, K, Ni, S, Si, V and Zn respectively. This is also illustrated in figure 1; larger amounts of Cu, Fe and Si are found in the PM_{10} fraction, whereas S and V were predominantly found in the fine fraction, which is in line with published results.[15, 17]

Land Use Regression Modeling

Land use regression models could be developed for the majority of the study areas and elements. From a potential 320 models (20 study areas x 8 elements x 2 size fractions), 292 models were developed. For 99% of these models, no evidence of spatial autocorrelation in the residuals was detected (Moran's I: $p > 0.05$). For 29 element-study area combinations, 19 in the $PM_{2.5}$ fraction and 10 in PM_{10}, no models could be developed. The majority of the unsuccessful modelling attempts (models without any significant predictor) were in Lugano (Switzerland), Kaunas (Lithuania) and Gyor (Hungary) with 8, 7 and 6 respectively. The large number of missing models in Lugano was due to missing data in reference site periods which meant that several sites (4) could not be adjusted and therefore not used. The elements with the largest fraction of missing models were S, Ni and V (7, 6 and 5 respectively). The lack of any model was probably related to small within-area variability (e.g. S) and poor precision of the measurements in areas with low concentrations (Ni and V).

Figure 2 shows the model R^2, LOOCV R^2 and RMSE by study area and element, with full model details shown in Table S4. Table 1 presents the mean and range of R^2 of the models. To explore which predictor variables contribute to particular elements in each fraction, the predictor variables were grouped into five categories (population, industry, ports, natural and traffic). Figure 3 shows a count of the number of study areas having at least 1 predictor within each of the defined categories.

Table 1 Average model R^2 and LOOCV R^2 for each element by fraction

Pollutant	PM_{10} Number of models	PM_{10} Average of model R^2	PM_{10} Average of LOOCV R^2	$PM_{2.5}$ Number of models	$PM_{2.5}$ Average of model R^2	$PM_{2.5}$ Average of LOOCV R^2
Copper (Cu)	20	0.79	0.73	20	0.72	0.65
Iron (Fe)	20	0.76	0.70	19	0.71	0.65
Potassium (K)	19	0.53	0.45	18	0.45	0.35
Nickel (Ni)	18	0.55	0.47	16	0.43	0.34
Sulphur (S)	16	0.57	0.48	17	0.50	0.42
Silicon (Si)	19	0.61	0.52	17	0.48	0.39
Vanadium (V)	19	0.55	0.46	16	0.50	0.40
Zinc (Zn)	19	0.73	0.66	18	0.67	0.59

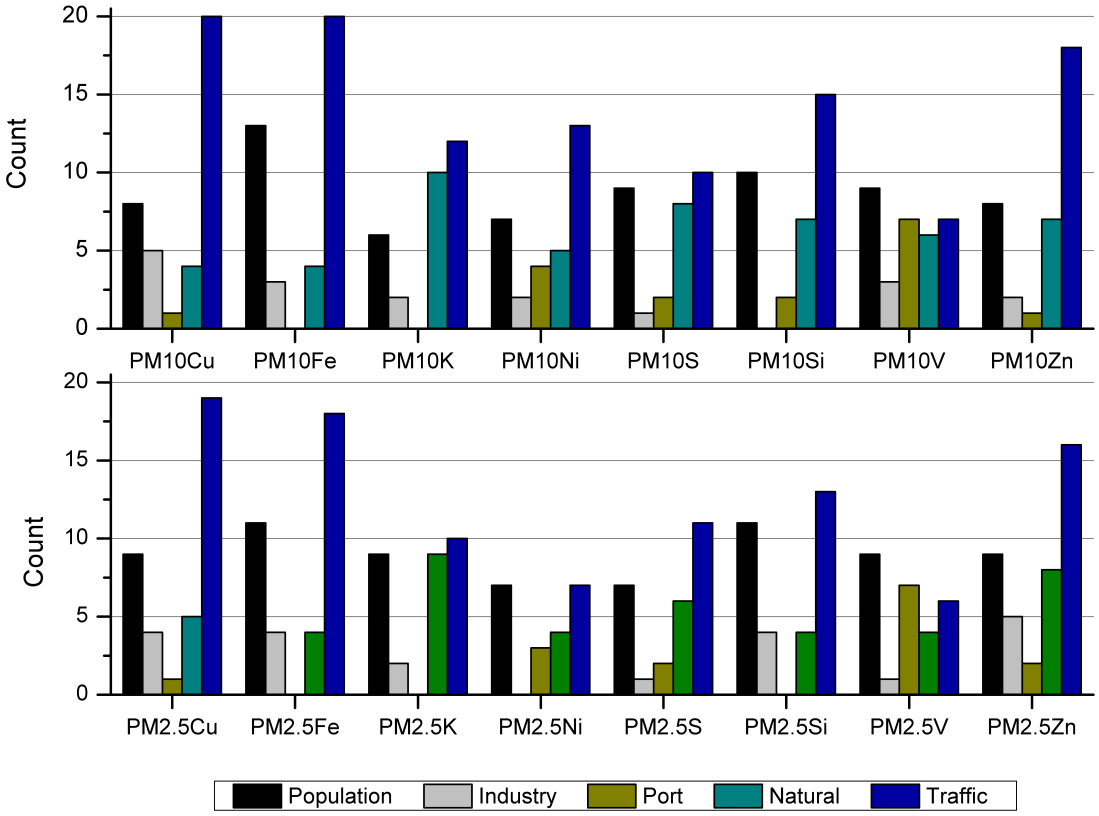


Figure 3 Counts of study areas with one or more predictor variables categorised by predictor variable categories: population; industry; ports/airports; natural; traffic for PM_{10} (top) and $PM_{2.5}$ (bottom).

Copper (Cu) models

LUR models for all study areas were developed with most (17 for both PM_{10} and $PM_{2.5}$ fractions) producing a R^2 and LOOCV R^2 greater than 50%. The low R^2s in some study areas are likely to be due to the small variation observed in the measurement data. Poor models in Heraklion can be explained by the lack of local traffic flow data. Generally, the predictor variables explaining the variance were dominated by traffic and road related variables, occurring in 20 PM_{10}Cu and 19 $PM_{2.5}$Cu models (see Figure 3). Traffic intensity was a predictor variable in all 20 $PM_{2.5}$Cu models and in 18 PM_{10}Cu models, whereas road length came in as a predictor in 12 $PM_{2.5}$Cu and 14 PM_{10}Cu models (see Table S4).

Iron (Fe) models

Similar to the LUR models developed for copper, LUR models for iron with R^2 and LOOCV R^2 greater than 0.50 were developed in most study areas (17 PM_{10}Fe and 15 $PM_{2.5}$Fe). Low model and LOOCV R^2s can, in most other study areas, be explained by small variations in measured Fe concentrations. This is not the case for $PM_{2.5}$Fe in Athens, where significant variation was measured, but a good LUR could not be developed. Traffic intensity and road length were predictor variables in the majority of models; respectively 20 in PM_{10}Fe models and 18 in $PM_{2.5}$Fe models. Population and/or residential areas also added to the explained variance in

the majority of models (13 in PM_{10}Fe and 11 in $PM_{2.5}$Fe).

Potassium (K) models

Land use regression models developed for potassium (19 for PM_{10}K and 18 for $PM_{2.5}$K) had a lower model R^2 than the Cu/Fe/Zn models. The average model R^2s were 0.53 in the PM_{10} fraction and 0.45 for $PM_{2.5}$, with only 12 and 7 models respectively with a model R^2 more than 0.50. Predictor variables associated with natural land (9 and 10 in $PM_{2.5}$K and PM_{10}K respectively) and road and/or traffic variables (12 in PM_{10}K and 10 in $PM_{2.5}$K) were present in the majority of models.

Nickel (Ni) models

As shown in Figure 1, largest variations in nickel measurements were found in southern Europe and this was reflected in relatively good models in Italy and Spain for PM_{10}Ni and in Spain and Athens for $PM_{2.5}$Ni. Belgium/Netherlands was the only other study area with a good model (LOOCV > 50%) in both fractions. Apart from traffic and road occurring in 13 PM_{10} models, no clear predictor variable category was dominant with the number of variables evenly distributed across the different categories (Figure 3). Port land use and distance to the harbour were predictors in the models for some study areas with major harbours (Netherlands Rotterdam, Barcelona, Athens and Heraklion). In the Netherlands, the Ni model further included a highly significant geographic coordinate, describing a decrease of concentrations with increasing distance to the sea.

Sulphur (S) models

Sulphur models were developed with mixed success; the average LOOCV R^2 was 0.48 for PM_{10}S and 0.42 for $PM_{2.5}$S. Measured sulphur concentrations were progressively higher going southwards, combined with a small increase in contrast within the study area observed in the same direction. This larger contrast again led to better models for Spain (both fractions) and Athens ($PM_{2.5}$S). The dominant predictor variable was traffic, which came into 10 PM_{10}S and 11 $PM_{2.5}$S models.

Silicon (Si) models

LUR model R^2s for silicon ranged considerably with high model R^2s in Oslo, Stockholm and Helsinki/Turku for $PM_{2.5}$Si and in Oslo and Stockholm for PM_{10}Si, coinciding with highest contrast of measured silicon in northern Europe. As shown in Figure 3, similar as with sulphur, the dominant predictor variable in the LUR models was traffic (15 in PM_{10} and 13 in $PM_{2.5}$)

Vanadium (V) models

Models developed for vanadium varied with the best models in Oslo, Belgium/Netherlands and Southern Europe for both fractions. Like nickel, variations of measured vanadium concentrations were highest in the southern study areas of Spain and Greece, probably explaining high performing models in these areas. In addition, PM_{10}V and $PM_{2.5}$V models for areas located near ports (Oslo, Helsinki/Turku, Copenhagen, Belgium/Netherlands, Ruhr area (with major inland river ports), Barcelona ($PM_{2.5}$V only), Catalunya (PM_{10}V only), Athens and Heraklion) generally included a large buffer for the port variable (5000m). V models

contained port more often than Ni models. Population variables entered the models relatively frequently.

Zinc (Zn) models

Zinc was successfully modelled in most study areas with average model LOOCV R^2 of 0.66 and 0.59 respectively for the PM_{10} and $PM_{2.5}$ fraction. Figure 1 shows high levels and contrast in measured zinc concentrations in Southern European study areas which resulted in slightly better models going southwards, especially for $PM_{2.5}$Zi. Compared to the other predictor variables, traffic predictor variables were dominant.

Sensitivity analyses

Correlation of model predictions with standard NO_2/PM models

The median R^2 of the Cu, Fe and PM_{10} Zn model predictions with the NO_2 and $PM_{2.5}$ absorbance model predictions was about 0.60, with large variability between the study areas (Table S5). PM_{10} Fe model predictions were higher correlated with especially absorbance (Median R^2=0.72). Model predictions of all components with PM_{10} or $PM_{2.5}$ models were generally moderate (R^2<0.4), with the exception of Si and PM_{10} and PM coarse (R^2=0.57 and 0.67).

Seasonality

The ratio between PM_{10}Cu concentrations measured at traffic and urban background sites was on average 2.4, 1.8, 2.1 and 1.9 in spring, summer, autumn and winter respectively, with northern European countries showing higher ratios compared to southern European countries. Correlations between PM_{10}Cu measurements for all sites between the seasons were similar with a median R of 0.88.

Model Transfer

When applying the PM_{10} model structure to $PM_{2.5}$ data, on average 10% less variability was explained when compared to the best $PM_{2.5}$ models, with Cu models showing the largest difference (15% on average) and Ni showing the smallest (6%).

Discussion

To our knowledge, no study has attempted to develop land use regression models for the spatial variation of long term ambient concentrations of elemental composition in the PM_{10} and $PM_{2.5}$ fractions. Land use regression models for 8 a priori selected elements were developed with success. Good LUR models were developed for Cu, Fe and Zn in both PM_{10} and $PM_{2.5}$ fractions explaining, on average, between 67 and 79% of the variance. Moderate models were developed for S, Si, Ni, V, K for PM_{10} and V and S for $PM_{2.5}$ with model R^2 ranging from 50 to 61%. Models for the elements K, Ni and Si in the $PM_{2.5}$ fraction performed poorest with both average model and LOOCV R^2s below 50%. For all eight elements, PM_{10} models explained, on average, more variance when compared to $PM_{2.5}$ models; 64% compared to 56%.

Performance of LUR models between elements

The explained variability for the components Cu, Zn and Fe was similar to the reported values in previous studies for the more commonly modeled components NO_2 and soot.[11] For the other five elements, explained variance was in the low end of the range of R^2 reported. For the same ESCAPE study areas we recently reported median model R^2 of 71%, 89% and 77% for $PM_{2.5}$, $PM_{2.5}$ absorbance and PM_{10} respectively.[13] For NO_2, the median R^2 was 83%.[14] Compared to the more commonly modeled components, there was more variability in performance between areas for the elements, with low R^2 in several areas and occasionally no models possible. In comparison, for $PM_{2.5}$, $PM_{2.5}$ absorbance and $PM_{10,}$ models were possible for all areas.

The difference in explained variability between the pollutants and between the study areas is likely related to differences in contrast in the measured concentrations and availability of predictor variables representing the major source of a component. Viana et al. provided an overview of results of studies in Europe investigating source apportionment of particulate matter.[15] The eight trace elements selected in this study were found to be mainly associated with the following main sources types; vehicle source (Fe, Zn, and Cu), crustal source (Si, Fe) and mixed industrial/fuel oil combustion (V, Ni). Within these source types, Fe was mostly associated with road dust and brake abrasion, Zn and Cu with tyre and brake abrasion and Si with re-suspended road dust. Both V and Ni were linked to crude oil and derived mainly from shipping emissions. The distinction between road dust and crustal source was often difficult because of overlapping profiles.[15]

Good LUR models were developed for *Cu, Fe and Zn*, which are all associated with non-tailpipe emissions from road traffic. Consistently, traffic related predictor variables dominated the other source categories. Previous land use regression studies have documented the importance of traffic variables for models of NO_2 and consequently developed the specification of traffic variables as predictor variables.[11] Based upon these previous studies, the ESCAPE study put a lot of effort into collecting predictor variables describing traffic sources. Furthermore traffic sites were overrepresented in the monitoring campaigns. Previous studies on traffic have exclusively focussed on exhaust emissions, represented by NO_2 and occasionally soot or EC. Our study adds that we can also model components of the non-exhaust emissions well. Model R^2s were slightly lower than for $PM_{2.5}$ absorbance, possibly because the emission is less well represented by traffic intensity variables. Finally, other difficulties to capture sources of these components (e.g. soil) may complicate modelling. When models are applied to addresses of the subjects participating in the cohort studies, an important issue is whether the predictions of the Cu, Fe and Zn models can be distinguished from the absorbance models (Table S5). The predictions of the Cu and Fe and to a lesser extent Zn models on average had a moderately high correlation with the predictions of the NO_2, NO_x and $PM_{2.5}$ absorbance models. This suggests that in the epidemiological analysis, it may be possible to evaluate the independent effects of non-tailpipe and tailpipe markers of traffic. However, in some areas, the correlation is too high to separate the two markers.

Models for elements which are predominantly emitted by other sources than

traffic, which are not represented by our GIS data, were more difficult to develop with the set of available predictor variables. No specific predictor variables, for example, were available which could pick up sources of secondary inorganic aerosols (S), soil material (Si) and biomass burning (K).

The *Ni and V* models were overall poorer, related mostly to very low concentrations in areas without the main sources present. In areas with higher concentrations, models with a good R^2 could be developed. Ports are important sources, represented in the predictor variables by CORINE land cover, a proxy for ship emissions. In contrast to road traffic, shipping emissions do not affect all study areas, e.g. most inland areas are not affected and consequently show very low concentrations (e.g. Munich, Lugano). While the port variable occurs frequently in the vanadium models it is not often in the models for nickel. Industry land use was rarely included in models, although specific industrial process may emit Ni and V.[15] We, however, did not have specific information on type of industry. In the Netherlands, the port variable may actually represent both shipping emissions and emissions from the oil refineries located near the Rotterdam harbour. Some models contain traffic variables, e.g. the Ruhr area Ni model includes only road length in a 1000m buffer. These variables likely do not represent direct traffic emissions, but account for the observed urban-rural differences in concentrations.

Given the small within-study area variability, the performance of the *S* models was reasonable. S is affected by large scale regional sources and less affected by local sources. Models accounted for the modestly increased urban rural differences and the sulphur emissions of diesel vehicles. In a few models industry and port was included, consistent with the identification of the Ni/V/SO_4^{2-} cluster as a source.[15]
Si models contained fewer traffic variables compared to the Cu, Fe and Zn models, suggesting that our models were able to provide some distinction between soil and road dust. The Si models often included the same predictor variables as the PM coarse model published previously,[13] accounting for especially urban – rural differences.

K models were relatively poor because we did not have specific GIS predictor variables for biomass burning, except in Helsinki/Turku. A study in Vancouver, modeled wood smoke levels, but this study was based upon a detailed emission inventory.[18]
The increasing north-south gradient observed in measured concentrations of some elements was also found by Viana et al[15] who attributed this to the influence of African dust.

Differences across study areas

With regards to study area, as shown in Table S6, the fewest number of models were developed in Lugano and Kaunas (respectively 8 and 10 out of 16 elements), however, the 8 LUR models created in Lugano showed on average the highest variance (86%). On the contrary, LUR models in Kaunas, together with Heraklion and Gyor, explained the least variance (33, 36 and 37% respectively). In most of the study areas models were developed for 15 or 16 elements. Of those, Barcelona (72%), Cataluña (71%) and Athens (69%) were high performing study areas with consistent high model R^2s. Differences between study areas can be explained by

concentration level and related precision of the measurements (Ni, V), within study area concentration contrast, complexity of the study area and the availability of GIS data. We previously reported that lack of traffic intensity data resulted in poorer models.[13]

PM_{10} and $PM_{2.5}$ models

On average, elemental PM_{10} models in most study areas explained a higher variance than elemental $PM_{2.5}$ models. The exceptions were the study areas Munich, Athens, Ruhr area and Gyor were $PM_{2.5}$ models yielded higher model R^2s. The difference is probably explained by the higher concentrations measured in the PM_{10} fraction. Furthermore, especially Cu, Zn, Fe and Si are concentrated in the coarse fraction. The $PM_{2.5}$ measurement reflects the tail of this coarse fraction and represents the source less well. We preferred to model the PM_{10} concentration and not the coarse (PM_{10} – $PM_{2.5}$) fraction, because $PM_{2.5}$ reflects the tail of the coarse particle fraction. Additional analysis on applying the PM_{10} model to $PM_{2.5}$ data revealed a drop in explained variability, justifying the development of separate models for both fractions.

Seasonality

We tested seasonality in the PM_{10}Cu fraction and, on the basis of our monitoring data, found no evidence of significant differences between the seasons.

Limitations

One of the main limitations of this study is the limited number of monitoring sites, 20 per study area, the LUR models are based on, as also discussed in Eeftens et al.[13] Despite the small number of sites, LUR models were developed that differed substantially between elements broadly reflecting known sources. We refrained from combining study areas as we anticipated that the available GIS predictors might lack specificity for the specific elements, which would be aggravated for larger areas, e.g. for the general industry land use variable. The limited number of monitoring sites also affects the validation of the LUR models. The validation method selected for ESCAPE, and which indeed is used in many other studies, was leave-one-out-cross-validation (LOOCV). Recent studies, however, suggested that this may significantly overestimate predictive ability in independent datasets.[19, 20] Wang et al. evaluated this by using the high correlation between PM metrics and NO_2/NO_x concentrations to estimate elemental concentration of Cu at NO_2/NO_x only monitoring sites.[20] External hold-out-evaluation (HEV) was then used to test the LUR models for Cu on the independent set of 20 monitoring sites. Even though the results confirmed that LOOCV R^2s are overestimated when using small sample set, the LUR models still explained a large fraction of the spatial variation within a study area (appr 50% for PM_{10}Cu).

A further limitation is the lack of specific GIS variables for especially industry and wood smoke. More specific GIS data for these sources are difficult to obtain. We evaluated source-specific emission data in some areas, but these rarely explained variation over the available GIS variables, probably because of a too large spatial scale.[21]

We also recognise that due care must be taken if using some of the poorer LUR models in health studies, for example by incorporating the cross validation R^2s in a sensitivity analysis.

Acknowledgements

The research leading to these results was funded by the European Community's Seventh Framework Program (FP7/2007-2011) projects ESCAPE (grant agreement number: 211250) and TRANSPHORM (ENV.2009.1.2.2.1). We thank all those who have been involved in exposure monitoring and modeling: Christophe Ampe, Erwan Gremaud, Sally Liu, Alex Ineichen, Kees Meliefste and Marjan Tewis.

References

1. Brunekreef, B.; Holgate, S. T. Air pollution and health. Lancet 2002, 360, 1233–1242.
2. World Health Organization. Systematic review of air pollution, a global update. Geneva, World Health Organization, 2006
3. Pope, C. A.,3rd; Dockery, D. W. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc. 2006, 56, 709-742.
4. HEI. Report 17: Traffic-related Air Pollution: a Critical Review of the Literature on Emissions, Exposure, and Health Effects. Health Effects Institute, 2010.
5. Kelly, F. J.; Fussell, J. C. Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. Atmos. Environ. 2012, 60, 504-526.
6. Stanek, L. W.; Sacks, J. D.; Dutton, S. J.; Dubois, J. B. Attributing health effects to apportioned components and sources of particulate matter: An evaluation of collective results. Atmos. Environ. 2011, 45, 655- 5663.
7. Bell, M. L.; Ebisu, K.; Peng, R. D.; Samet, J. M.; Dominici, F. Hospital admissions and chemical composition of fine particle air pollution. Am. J. Respir. Crit. Care Med. 2009, 179, 1115-1120.
8. Mostofsky, E.; Schwartz, J.; Coull, B.A.; Koutrakis, P.; Wellenius, G.A.; Suh, H.H.; et al., Modeling the Association Between Particle Constituents of Air Pollution and Health Outcomes. Am. J. Epidemiol. 2012, 176, 317-326.
9. Strak, M.; Janssen, N. A.; Godri, K. J.; Gosens, I.; Mudway, I. S.; Cassee, F. R.; et al., Respiratory Health Effects of Airborne Particulate Matter: The Role of Particle Size, Composition, and Oxidative Potential-The RAPTES Project. Environ. Health Perspect. 2012, 120, 1183-1189
10. Ostro, B.; Lipsett, M.; Reynolds, P.; Goldberg, D.; Hertz, A.; Garcia, C.; et al., Long-term exposure to constituents of fine particulate air pollution and mortality: Results from the California Teachers Study. Environ. Health Perspect. 2010, 118, 363-369. Revised estimates used Environ. Health Perspect. 2011, 119, A242-243.
11. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; et al., A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 2008, 42, 7561−7578.
12. Noth, E. M.; Hammond, S. K.; Biging, G. S.; Tager, I. B. A spatial-temporal regression model to predict daily outdoor residential PAH concentrations in an epidemiologic study in Fresno, CA. Atmos. Environ. 2011, 45, 2394-2403.
13. Eeftens, M.; Beelen, R.; de Hoogh, K.; Bellander, T.; Cesaroni, G.; Cirach, M.; et al., Development of NO_2 and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe – the ESCAPE project. Atmos. Envir. 2013, 72, 10-23.
15. Viana, M.; Kuhlbusch, T. A. J.; Querol, X.; Alastuey, A.; Harrison, R. M.; Hopke, P. K.; et al., Source apportionment of particulate matter in Europe: A review of methods and results. J. of Aerosol Sci. 2008, 39, 827-849.
16. Eeftens, M.; Tsai, M.; Ampe, C.; Anwander, B.; Beelen, R.; Bellander, T.; et al., Spatial variation of $PM_{2.5}$, PM_{10}, $PM_{2.5}$ absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO_2 – Results of the ESCAPE project. Atmos.

Envir. 2012, 62, 303-317.
17. Putaud, J.-P.; Van Dingenen, R.; Alastuey, A.; Bauer, H.; Birmili, W.; Cyrys, J.; et al., A European aerosol phenomenology – 3: Physical and chemical characteristics of particulate matter from 60 rural, urban, and kerbside sites across Europe. Atmos. Envir. 2010, 44, 1308-1320.
18. Larson, T.; Su, J. G.; Baribeau, A. M.; Buzzelli, M.; Setton, E.; Brauer, M. A spatial model of urban winter woodsmoke concentrations. Environ. Sci. Technol. 2007, 41, 2429–2436.
19. Basagaña, X.; Rivera, M.; Aguilera, I.; Agis, D.; Bouso, L.; Elosua, R.; et al., Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos. Envir. 2012, 54, 634-642.
20. Wang, M.; Beelen, R.; Basagana, X.; Becker, T.; Cesaroni, G; de Hoogh, K.; et al., Evaluation of land use regression models for NO_2 and particulate matter in 20 European study areas: the ESCAPE project. Environ. Sci. Technol. 2012, Submitted.
21. Rosenlund, M.; Forastiere, F.; Stafoggia, M.; Porta, D.; Perucci, M.; Ranzi, A.; et al., Comparison of regression models with land-use and emissions data to predict the spatial distribution of traffic-related air pollution in Rome. J. Expo. Sci. Environ. Epidemiol. 2008, 18, 192–199.

Chapter 5

Appendix

(Online: http://pubs.acs.org/doi/suppl/10.1021/es400156t)

Table S1: All used predictor variables with predefined variable names, units, buffer sizes, directions of effect, and whether data are central or local data, divided in background and traffic variables.

GIS dataset	Predictor variable	Name variable[a]	Unit	Buffer size (radius in m)	Direction of effect	Central / local data
Background variables						
CORINE	Surface area of high density residential land	HDRES_X	m^2	100, 300, 500, 1000, 5000	+	Central
CORINE	Surface area of low density residential land	LDRES_X	m^2	100, 300, 500, 1000, 5000	+	Central
CORINE	Sum of high density and low density residential land	HLDRES_X	m^2	100, 300, 500, 1000, 5000	+	Central
CORINE	Surface area of industry	INDUSTRY_X	m^2	100, 300, 500, 1000, 5000	+	Central
CORINE	Surface area of port	PORT_X	m^2	100, 300, 500, 1000, 5000	+	Central
CORINE	Surface area of urban green[b]	URBGREEN_X	m^2	100, 300, 500, 1000, 5000	-	Central
CORINE	Surface area of semi-natural and forested areas[c]	NATURAL_X	m^2	100, 300, 500, 1000, 5000	-	Central
CORINE	Sum of Urban green and Semi-natural and forested areas	GREEN_X	m^2	100, 300, 500, 1000, 5000	-	Central
Local land use	Similar to CORINE variables	Similar to CORINE variables	m^2	100, 300, 500, 1000, 5000	As CO-RINE	Local
Local land use	Surface area / number of buildings	BUILDINGS_X	m^2 / N(umber)	100, 300, 500, 1000, 5000	+	Local
Local land use	Surface area of water	WATER_X	m^2	100, 300, 500, 1000, 5000	-	Local
Population density	Number of inhabitants	POP_X	N(umber)	100, 300, 500, 1000, 5000	+	Central / Local
Household density	Number of households	HHOLD_X	N(umber)	100, 300, 500, 1000, 5000	+	Local
Altitude	Altitude	SQRALT	m	NA	-	Local
-	Regional estimate[d]	REGIONALESTIMATE	NA	NA	NA	Local
-	Coordinate variables[d]	XCOORD, YCOORD or other combinations	m	NA	NA	Local
Traffic variables						
Central road network	Road length of all roads in a buffer	ROADLENGTH_X	m	25, 50, 100, 300, 500, 1000	+	Central
Central road network	Road length of major roads in a buffer[e]	MAJORROADLENGTH_X	m	25, 50, 100, 300, 500, 1000	+	Central
Central road network	Inverse distance and inverse squared distance to the nearest road	DISTINVNEARC1 DISTINVNEARC2	m^{-1}, m^{-2}	NA	+	Central
Central road network	Inverse distance and inverse squared distance to the nearest major road[e]	DISTINVMAJORC1 DISTINVMAJORC2	m^{-1}, m^{-2}	NA	+	Central
Local road network	Traffic intensity on nearest road	TRAFNEAR	$Veh.day^{-1}$	NA	+	Local
Local road network	Inverse distance and inverse squared distance to the nearest road	DISTINVNEAR1 DISTINVNEAR2	m^{-1}, m^{-2}	NA	+	Local
Local road network	Product of traffic intensity on nearest road (TRAFNEAR) and inverse distance to nearest road (DISTINVNEAR1) and distance squared (DISTINVNEAR2)	INTINVDIST INTINVDIST2	$Veh.day^{-1}m^{-1}$ $Veh.day^{-1}m^{-2}$	NA	+	Local
Local road network	Traffic intensity on nearest major road[f]	TRAFMAJOR	$Veh.day^{-1}$	NA	+	Local
Local road network	Inverse distance and inverse squared distance to the nearest major road[f]	DISTINVMAJOR1 DISTINVMAJOR2	m^{-1}, m^{-2}	NA	+	Local

Local road network	Product of traffic intensity on nearest major road (TRAFMAJOR) and inverse of distance to the nearest major road (DISTINVMAJOR1) and distance squared (DISTINVMAJOR2)[f]	INTMAJORINVDIST INTMAJORINVDIST2	$Veh.day^{-1}m^{-1}$ $Veh.day^{-1}m^{-2}$	NA	+	Local
Local road network	Total traffic load of major roads in a buffer (sum of (traffic intensity * length of all segments))[6]	TRAFMAJORLOAD_X	$Veh.day^{-1}m$	25, 50, 100, 300, 500, 1000	+	Local
Local road network	Total traffic load of all roads in a buffer (sum of (traffic intensity * length of all segments))	TRAFLOAD_X	$Veh.day^{-1}m$	25, 50, 100, 300, 500, 1000	+	Local
Local road network	Heavy-duty traffic intensity on nearest road	HEAVYTRAFNEAR*	$Veh.day^{-1}$	NA	+	Local
Local road network	Product of heavy-duty traffic intensity on nearest road (HEAVYINT) and inverse of distance to the nearest road (INVDIST) and distance squared (INVDIST2)	HEAVYINTINVDIST HEAVYINTINVDIST2	$Veh.day^{-1}m^{-1}$ $Veh.day^{-1}m^{-2}$	NA	+	Local
Local road network	Heavy-duty traffic intensity on nearest major road[6]	HEAVYTRAFMAJOR	$Veh.day^{-1}$	NA	+	Local
Local road network	Total heavy-duty traffic load of major roads in a buffer (sum of (heavy-duty traffic intensity * length of all segments))[f]	HEAVYTRAFMAJORLOAD_X	$Veh.day^{-1}m$	25, 50, 100, 300, 500, 1000	+	Local
Local road network	Total heavy-duty traffic load of all roads in a buffer (sum of (heavy-duty traffic intensity * length of all segments))	HEAVYTRAFLOAD_X	$Veh.day^{-1}m$	25, 50, 100, 300, 500, 1000	+	Local

[a]Variable name: Combining name and buffer size, for example for HDRES_X: HDRES_100, HDRES_300, HDRES_500,HDRES_1000, HDRES_5000
[b]Corine Urban green is the sum of CORINE classes 141 and 142
[c]Corine semi-natural is the sum of CORINE classes 311, 312, 313, 321, 322, 323, 324, 331, 332, 333, 334, 335, 411, 412, 421, 422, 423, 512, 521, 522, 523
[d]Variables were only offered if a model has been developed to test if the model with more explicit variables could be improved with these variables (describing slow trends in background).
[e]Definition of major road for central road network: classes 0, 1, and 2 (+ classes 3 and 4 based on local knowledge and decision)
[f]Definition of major road for local road network: road with traffic intensity > 5,000 mvh/24h
NA is not applicable

Table S2: Description of LUR models for 8 elements for both PM_{10} and $PM_{2.5}$ fractions by study area
Copper (Cu) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N^a	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV (ng/m^3)	N^a	Mean Conc [range] (ng/m^3)
NOS	-4.8 +1.8E+02*DISTINVMAJOR1 +2.5E-04*POP5000 +7.8E-02*INTMAJORINVDIST2	0.91	0.87	5.10	19	16 [0.49-45]	-1.5 +3.8E-02*INTINVDIST2 +9.6E-04*ROADLENGTH300 +1.2E-03*HHOLD500	0.84	0.76	1.64	19	3.3 [-0.45-11]
SST	-0.12 + 3.6E-06*HEAVYTRAFLOAD_500 +1.3E+02*DISTINVMAJORCL1 + 3.9E-03*HEAVYTRAFMAJOR	0.87	0.84	6.05	16	20 [2.6-56]	0.1 +3.0E-07*INDUSTRY5000 +9.4E-06*HEAVYTRAFLOAD_100	0.69	0.61	1.83	18	4.0 [0.35-11]
FIH	0.2 +3.5E-04*POP_1000L +2.6E-06*TRAFMAJORLOAD_50 +1.9E-02*ROADLENGTH_50	0.72	0.61	3.80	20	7.3 [0.96-26]	0.5 +7.3E-04*HHOLD_300L +1.5E-07*TRAFLOAD_100 +5.4E-03*ROADLENGTH_50	0.88	0.84	0.44	20	1.8 [0.47-4.1]
DCO	1.4 +7.2E-07*TRAFLOAD_300	0.92	0.91	3.69	20	15 [2.4-41]	0.4 +8.0E-07*TRAFMAJORLOAD_50 +6.9E-04*MAJORROADLENGTH_500	0.69	0.61	1.71	20	3.8 [0.97-12]
LIK	4.7 +2.4E+01*DISTINVMAJOR1 +5.3E-10*TRAFLOAD500	0.62	0.54	2.59	20	8.1 [3.2-17]	3.5 +1.7E-10*TRAFLOAD500	0.25	0.15	1.33	20	4.3 [1.9-6.9]
UKM	5.5 +1.5E-05*HDRES_300 +2.4E-02*ROADLENGTH_25 +1.8E-02*MAJORROADLENGTH_100 +4.4E-04*MAJORROADLENGTH_100_1000	0.92	0.87	1.91	19	10 [5.0-22]	3.0 +7.5E-03*MAJORROADLENGTH_100 -3.8E-06*NATURAL_1000	0.69	0.60	0.86	19	3.0 [1.7-6.5]
UKO	-0.3 +1.4E-05*TRAFLOAD25 +3.4E-04*ROADLENGTH1000	0.96	0.95	3.50	20	19 [4.1-79]	0.9 +5.3E-08*HLDRES5000 +1.8E-05*TRAFMAJOR +1.9E-02*ROADLENGTH25 +1.2E-02*MAJORROADLENGTH50	0.84	0.79	0.71	19	4.4 [1.4-6.7]
BNL	6.6 +8.0E-07*PORT_5000 +3.9E-04*TRAFNEAR +2.1E-06*TRAFMAJORLOAD_50 +1.6E-07*HEAVYTRAFMAJORLOAD_1000 +8.8E-02*ROADLENGTH25	0.80	0.71	7.38	40	19 [3.4-59]	6.5 +4.8E-08*HDLDRES_5000 +5.0E-07*TRAFMAJORLOAD_50 +1.0E-02*MAJORROADLENGTH50 -6.7E-06*XPLUSY	0.83	0.81	1.15	40	5.0 [1.5-12]
GRU	3.3 +8.1E-07*INDUSTRY_5000 +2.2E-06*TRAFLOAD_100 +1.6E-03*MAJORROADLENGTH_1000	0.94	0.92	3.83	20	22 [5.6-53]	172.8 +2.2E-07*INDUSTRY_5000 +7.2E-05*HEAVYTRAFLOAD_25 +8.7E-04*MAJORROADLENGTH_500 -2.1E-05*XPLUSY	0.92	0.90	0.87	20	6.7 [2.5-11]
GMU	8.4 +1.4E+02*INDUSTRY5000 +3.7E-04*TRAFNEAR +1.5E-01*ROADLENGTH_50	0.75	0.71	11.36	20	29 [11-78]	4.3 +1.6E+01*INDUSTRY5000 +1.1E-06*TRAFLOAD50 +1.8E-03*ROADLENGTH_300	0.81	0.76	1.53	20	7.9 [4.8-14]
AUV	-2.5 +3.9E-07*RES5000 +9.9E-06*TRAFLOAD25	0.89	0.87	1.78	20	11 [2.9-20]	-1.5 +1.6E-07*RES5000 +9.5E-05*RES100	0.73	0.67	0.83	20	3.9 [0.75-6.2]

	+1.2E-03*ROADLENGTH300											
FPA	3.4 +2.6E-05*INDUSTRY_1000 +1.2E-06*TRAFLOAD_300	0.62	0.48	27.12	20	34 [2.0-148]	1.1 +6.3E-06*INDUSTRY_1000 -2.5E-07*NATURAL_5000 +4.5E-04*TRAFMAJOR	0.68	0.51	4.87	20	7.6 [1.6-29]
HUG	7.2 +1.0E-02*INDUSTRY_5000 -2.4E-02*URBGREEN_5000 +1.7E-02*MAJORROADLENGTH_100	0.44	0.25	4.86	20	10 [4.2-26]	4.2 -7.0E-03*URBGREEN_5000 +5.8E-08*TRAFMAJORLOAD_1000	0.49	0.38	1.42	20	4.4 [1.1-7.1]
SWL	-15.9 +4.0E-05*TRAFLOAD25 +3.5E-03*ROADLENGTH300	0.88	0.84	8.94	19	29 [5.5-100]	-1.7 +2.7E-06*TRAFLOAD50 +2.0E-04*ROADLENGTH500	0.87	0.83	1.38	19	5.8 [1.4-15]
ITU	26.0 -6.5E-07*NATURAL5000 +3.9E-03*POP500 +7.0E-06*TRAFMAJORLOAD50	0.90	0.88	11.10	20	54 [9.8-115]	5.0 -1.7E-07*NATURAL5000 +5.0E-04*POP500 +3.0E-08*TRAFLOAD1000	0.88	0.85	2.36	20	13 [2.4-26]
IRO	18.9 +5.1E-05*TRAFMAJORLOAD_25 +9.5E-08*TRAFLOAD_1001 +1.2E-01*MAJORROADLENGTH_50	0.89	0.87	14.15	20	58 [14-133]	7.7 +8.5E-06*TRAFMAJORLOAD_25 +2.2E-08*TRAFLOAD_1001	0.82	0.78	2.91	20	14 [5.9-27]
SPB	-23.1 -6.7E-07*NATURAL_5000 +1.4E-04*HHOLD_5000 +5.7E+02*DISTINVMAJOR1	0.83	0.79	17.25	20	52 [7.5-142]	-7.1 +6.9E-07*PORT_5000 +2.6E-05*HHOLD_5000 +1.9E-08*TRAFMAJORLOAD_500 +4.5E-06*TRAFLOAD_25	0.78	0.71	3.10	20	11 [3.0-25]
SPC	-4.4 +1.3E-06*HDRES_5000 +2.3E-05*INDUSTRY_1000 -4.7E-05*URBGREEN_1000 -9.3E-05*GREEN_300 +2.2E+02*DISTINVMAJOR1 +5.0E-06*TRAFMAJORLOAD_50	0.75	0.67	19.55	40	41 [4.9-142]	1.0 +1.7E-07*HDRES_5000 -1.7E-05*GREEN_300 +1.0E-03*INTMAJORINVDIST1 +6.6E-09*TRAFMAJORLOAD_1000 +6.8E-02*MAJORROADLENGHT_25	0.77	0.67	3.21	40	8.5 [0.71-25]
GRA	17.4 +2.3E-08*TRAFMAJORLOAD_300 +1.8E-05*TRAFLOAD_25 +1.0E-02*ROADLENGTH_100	0.73	0.70	7.74	19	36 [14-73]	9.3 +9.3E-07*HDRES_1000 -1.1E-06*URBGREEN_5000 +4.4E-09*TRAFMAJORLOAD_500 +1.5E-07*TRAFLOAD_100 +9.5E-03*ROADLENGTH_50	0.84	0.76	1.43	19	10 [5.4-16]
GRH	4.6 +1.5E-02*POP_100 +1.1E-01*MAJORROADLENGTH_25	0.55	0.40	4.15	20	9.6 [2.7-24]	2.1 +7.9E+00*DISTINVMAJORC1	0.10	0.05	1.57	20	2.4 [-0.10-6.5]

Iron (Fe) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)
NOS	168.6 +5.0E+00*MAJORROADLENGTH50 +2.9E+00*INTMAJORINVDIST2 +9.1E-02*HHOLD500	0.94	0.92	113.66	19	552 [48-1259]	68.3 +9.4E-01*MAJORROADLENGTH50 +1.2E-02*HEAVYTRAFNEAR -3.1E+00*SQRALT	0.87	0.82	28.40	19	80 [-2.09-205]
SST	263.1 +3.3E-04*TRAFLOAD_50 +2.8E-02*MAJORROADLENGTH_1000	0.75	0.68	277.16	18	817 [57-1995]	-8.5 +1.7E-04*LDRES300 +3.7E-04*INDUSTRY500 +4.8E-06*INDUSTRY500_5000 +8.2E-02*POP_100 +6.4E+02*DISTINVMAJOR1 +2.1E-04*HEAVYTRAFLOAD_100	0.93	0.90	24.49	18	142 [-0.25-295]
FIH	32.6 +1.5E-02*POP_1000L +4.7E-06*TRAFLOAD_300 +1.1E+00*ROADLENGTH_50	0.59	0.47	178.64	20	363 [72-1164]	72.1 -1.6E-06*GREEN_5000L +3.6E-01*HHOLD_100L +8.0E-02*ROADLENGTH_100	0.76	0.63	23.79	20	75 [15-169]
DCO	114.0 -3.0E-04*GREEN_500 +1.4E-05*TRAFLOAD_300	0.94	0.92	67.38	20	362 [105-843]	26.3 +1.9E-06*INDUSTRY_5000 -1.2E-05*GREEN_1000 +1.9E-06*TRAFMAJORLOAD_300 +4.3E-02*MAJORROADLENGTH_100	0.94	0.91	11.71	20	71 [28-158]
LIK	189.4 +9.8E-02*POP300 +8.5E-07*TRAFLOAD100	0.47	0.40	93.28	20	357 [186-609]						75 [34-200]
UKM	174.0 +1.4E-04*HDRES_1000 +1.6E-06*HEAVYTRAFLOAD_1000 +8.8E-01*ROADLENGTH_25 +2.2E-01*MAJORROADLENGTH_100	0.89	0.79	67.48	19	306 [160-608]	39.0 +7.4E-08*TRAFLOAD_1000 +3.0E-02*ROADLENGTH_100	0.69	0.62	16.52	19	67 [40-132]
UKO	191.4 +5.8E-04*POP5000 +4.0E-04*TRAFLOAD25	0.95	0.95	95.44	20	521 [134-1999]	37.4 +8.5E-05*POP5000 +1.3E-02*HEAVYTRAFMAJOR +3.3E-02*INTMAJORINVDIST	0.94	0.92	19.69	20	111 [26-355]
BNL	181.0 +6.0E-04*POP_5000 +7.3E-03*TRAFNEAR +2.2E-06*TRAFMAJORLOAD_500 +2.2E+00*ROADLENGTH25	0.78	0.70	187.70	40	547 [153-1673]	149.0 +1.4E-06*HDLDRES_5000 +1.9E-03*TRAFNEAR +8.7E-06*TRAFMAJORLOAD_50 -1.5E-04*XPLUSY	0.78	0.73	32.44	40	120 [33-278]
GRU	286.9 +2.1E-04*INDUSTRY_1000 +2.4E-02*POP_1000 +8.8E-03*HEAVYTRAFLOAD_25 +2.0E-05*HEAVYTRAFLOAD_25_1000	0.89	0.85	137.19	19	808 [349-1487]	139.1 +3.4E-06*INDUSTRY_5000 +2.3E-03*HEAVYTRAFMAJORLOAD_25 +2.4E-06*TRAFLOAD_300	0.68	0.62	41.27	19	219 [109-310]
GMU	-274.4 +6.1E+02*HLDRES_500 +1.4E-06*TRAFLOAD1000 +5.5E+00*ROADLENGTH_50	0.87	0.82	219.07	20	688 [258-2136]	54.8 +3.0E-07*TRAFLOAD1000 +7.3E-01*ROADLENGTH_50	0.87	0.84	29.16	20	138 [64-315]
AUV	198.0 +4.8E-04*RES300 +2.1E-04*TRAFLOAD25	0.66	0.62	67.12	20	359 [187-542]	44.3 +8.7E-05*RES300 +3.7E+02*DISTINVMAJORC1	0.52	0.46	13.90	20	70 [42-108]
FPA	203.6 +4.5E-04*INDUSTRY_1000	0.63	0.50	477.47	20	747	22.9 +1.5E-06*LDRES_5000	0.83	0.79	44.62	20	129

	+2.3E-05*TRAFLOAD_300					[105-2665]	+3.9E-05*TRAFLOAD_25 +1.9E-02*MAJORROADLENGTH_500					[26-410]
HUG	245.9 +2.4E+00*HDRES_1000 +1.9E-01*MAJORROADLENGTH_300	0.67	0.54	124.38	20	394 [199-964]	50.8 +1.3E-03*ROADLENGTH_1000	0.14	0.04	45.58	20	89 [44-243]
SWL	182.5 +2.1E-06*TRAFLOAD1000 +4.1E-04*TRAFLOAD50 -1.6E-01*DISTWATER	0.90	0.85	186.90	19	756 [141-2190]	61.2 +5.6E-07*TRAFMAJORLOAD1000 +6.2E-05*TRAFLOAD50 -3.9E-02*DISTWATER	0.87	0.83	33.16	19	149 [29-349]
ITU	735.1 -2.2E-05*NATURAL5000 +2.3E-01*POP_300 +9.6E-01*MAJORROADLENGTH_100	0.89	0.86	261.76	20	1343 [378-2599]	150.7 +1.8E-03*RES100 -5.0E-06*NATURAL5000 +1.0E-02*POP_500 +2.1E-05*TRAFLOAD50	0.88	0.83	45.06	20	253 [69-503]
IRO	-62776.1 +1.1E-03*TRAFLOAD_25 +1.3E-02*ROADLENGTH_1000 +2.0E+00*MAJORROADLENGTH_50 +1.3E-02*XPLUSY	0.85	0.82	321.86	20	1297 [465-2639]	146.5 +4.4E-06*HDRES_5000 +1.6E-04*TRAFLOAD_25	0.72	0.67	77.24	20	283 [126-510]
SPB	365.2 +3.6E-03*HDRES_300 +4.5E-01*INTINVDIST1	0.83	0.77	345.75	20	1365 [348-3090]	96.8 +2.2E-04*HDRES_500 +1.5E+03*DISTINVMAJOR1	0.69	0.63	66.96	20	236 [79-488]
SPC	818.4 -1.2E-05*NATURAL_5000 +1.4E-02*HHOLD_1000 +4.0E+03*DISTINVMAJOR1 +9.3E-05*TRAFMAJORLOAD_50	0.72	0.65	394.59	40	1165 [321-3090]	33.7 +3.9E-06*HDRES_5000 +9.0E-05*INDUSTRY_1000 -1.4E-04*URBGREEN_1000 +4.0E-07*TRAFMAJORLOAD_500 +8.7E-01*MAJORROADLENGHT_50	0.74	0.64	64.79	40	196 [44-488]
GRA	630.8 +1.2E-03*INDUSTRY_300 -4.5E-05*URBGREEN_5000 +2.6E-07*TRAFMAJORLOAD_500 +1.2E-05*TRAFLOAD_100 +3.9E-04*TRAFLOAD_25 +3.8E-01*ROADLENGTH_100	0.84	0.75	159.59	19	1113 [671-1892]	190.6 +1.1E-07*TRAFMAJORLOAD_500 +5.1E-01*MAJORROADLENGTH_50	0.28	0.11	80.00	20	258 [150-542]
GRH	379.3 +2.0E-01*ROADLENGTH_100	0.20	0.07	137.30	20	539 [342-953]	106.0 +1.8E-04*HDRES_300	0.37	0.29	23.50	20	115 [73-160]

Potassium (K) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)
NOS	19.3 +7.1E+04*DISTINVMAJORC2 +2.8E-02*HHOLD1000 +3.1E-06*LDRES5000 -3.7E+00*SQRALT	0.93	0.9	31.23	19	223 [74-417]	98.8 +6.0E-03*POP1000 -2.8E+00*SQRALT	0.46	0.32	26.26	19	86 [41-169]
SST	78.2 +9.5E-04*HLDRES300 +1.2E-04*TRAFLOAD_50	0.80	0.76	82.70	18	377 [86-758]	54.4 + 2.2E-04*HLDRES300 + 1.0E-03*TRAFNEAR	0.68	0.55	18.39	18	113 [43-151]
FIH	280.5 -2.7E-04*GREEN_500L +3.4E-05*TRAFMAJORLOAD_50	0.39	0.24	73.02	20	263 [131-528]	125.6 -8.0E-05*GREEN_500L	0.15	0.09	20.30	20	113 [78-151]
DCO	195.0 -4.4E-04*Green_300	0.25	0.15	19.55	20	188 [143-217]	112.9 +1.2E-05*INDUSTRY_L_5000 -2.5E-03*GREEN_L_100	0.57	0.53	9.00	20	115 [91-136]
LIK						399 [314-547]	229.2 +1.8E-03*LDRES100	0.13	0.05	51.27	20	269 [181-418]
UKM	131.5 +4.1E-03*ROADLENGTH_300	0.11	0.01	23.82	19	150 [108-199]						71 [57-121]
UKO	104.7 +7.4E-07*HLDRES5000	0.61	0.56	13.70	20	139 [105-185]	65.6 +2.5E-07*HLDRES5000	0.26	0.14	9.15	20	77 [60-108]
BNL	173.0 +1.2E-04*HDLDRES_300 +4.0E-07*TRAFMAJORLOAD_500	0.51	0.45	27.60	40	223 [159-349]	155.0 +3.5E-07*TRAFMAJORLOAD_300 +1.4E-04*XMINUSY	0.31	0.25	18.85	40	119 [76-167]
GRU	289.8 +5.2E-06*INDUSTRY_5000	0.22	0.14	50.68	20	336 [255-471]						209 [161-317]
GMU	383.2 -2.6E+02*URBGREEN_5000 +1.2E+02*DISTINVNEAR1 +1.1E-05*TRAFLOAD50 +1.6E-01*ROADLENGTH50	0.69	0.63	22.06	20	394 [325-453]	463.8 -1.4E+02*URBGREEN_5000 -8.3E+00*SQRALT +1.3E-01*MAJORROADLENG_50	0.55	0.38	17.79	20	269 [236-325]
AUV	344.0 +2.9E-05*INDUSTRY5000 -2.1E-06*NATURAL5000 +2.9E+00*APT100	0.69	0.60	29.42	19	391 [318-466]	447.5 -2.2E-04*NATURAL500 -7.9E+00*SQRALT +3.3E+00*APT100	0.79	0.69	23.99	19	338 [259-401]
FPA	167.2 -2.2E-06*NATURAL_5000 +1.4E-03*TRAFMAJOR +2.7E-01*ROADLENGTH_25	0.54	0.52	24.03	20	183 [121-250]	94.5 +4.7E-06*HDRES_1000 +4.0E-05*INDUSTRY_500 +4.0E-04*TRAFNEAR	0.49	0.31	10.04	20	104 [78-125]
HUG	518.7 -3.2E-01*URBGREEN_5000	0.34	0.27	50.76	20	442 [307-542]	351.4 +6.5E-01*LDRES_500 -1.8E-01*URBGREEN_5000	0.26	0.15	44.72	20	337 [266-426]
SWL	237.6 -5.8E-05*URBGREEN1000 -6.0E-05*NATURAL1000 +1.6E-05*TRAFLOAD50	0.75	0.67	18.35	19	237 [178-289]	755.5 -1.7E-05*NATURAL5000 -7.2E+00*SQRALT +1.7E-03*HEAVYTRAFLOAD50	0.83	0.78	11.33	19	152 [110-189]
ITU	874.9 +3.4E-03*RES100 -1.6E-04*NATURAL1000 -2.4E+01*SQRALT	0.54	0.48	57.05	20	566 [443-732]	325.6 +1.9E-04*LDRES300 +4.1E-01*ROADLENGHT25	0.21	0.11	47.42	20	384 [299-480]
IRO	-14139.2 +3.1E+03*DISTINVMAJOR1 +3.0E-03*XPLUSY	0.64	0.57	83.94	20	513 [379-877]	-11938.5 +6.8E+02*DISTINVMAJOR1	0.46	0.41	35.37	20	297 [211-398]

							+2.5E-03*XPLUSY					
SPB	742.1 -3.5E-05*URBGREEN_5000 -2.4E+01*SQRALT +1.6E+02*DISTINVNEAR1 +8.7E-05*TRAFLOAD_25	0.85	0.78	41.30	20	482 [275-639]	305.1 -3.8E-05*URBGREEN_300_1000 -1.9E-05*URBGREEN_1000_5000 -3.7E-04*URBGREEN_300 -3.2E-06*NATURAL_5000 +2.5E-06*TRAFLOAD_100 +5.7E-03*MAJORROADLENGHT_500	0.83	0.73	19.53	20	175 [111-254]
SPC	661.8 -2.2E-05*URBGREEN_5000 -2.7E+01*SQRALT +4.6E+02*DISTINVMAJOR1 +1.5E+00*MAJORROADLENGHT_25	0.71	0.62	76.13	40	486 [261-955]	242.3 -8.5E-05*URBGREEN_1000 -9.8E+00*SQRALT +9.7E-02*MAJORROADLENGHT_100	0.31	0.20	57.61	40	184 [95-478]
GRA	327.9 +7.8E-07*TRAFMAJORLOAD_100 +1.9E-01*ROADLENGTH_50	0.33	0.13	41.25	20	380 [292-494]	189.2 +4.8E-07*LDRES_5000 -1.8E-06*NATURAL_5000 +3.0E-08*TRAFLOAD_1000	0.61	0.51	15.54	20	214 [159-258]
GRH	497.8 +1.2E-04*HDRES_1000	0.23	0.09	104.62	20	555 [387-870]	235.1 +1.7E-01*POP_100	0.12	0.03	75.27	20	276 [176-500]

Nickel (Ni) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)
NOS	0.9 +3.7E-07*HDRES1000 +3.1E-07*PORT5000	0.69	0.62	0.32	19	1.2 [0.37-2.2]	0.3 +7.9E-08*HDRES5000	0.56	0.48	0.35	19	0.63 [-0.19-1.5]
SST	1.2 +2.9E-06*INDUSTRY500 -1.2E-08*NATURAL5000 +2.8E+00*DISTINVMAJOR1 +2.0E-07*HEAVYTRAFMAJORLOAD_500	0.88	0.81	0.27	16	1.7 [0.80-2.6]						1.3 [0.45-4.2]
FIH	1.1 +4.6E-04*POP_300L	0.32	0.18	0.72	20	1.6 [0.70-4.0]	0.7 +5.1E-06*POP_5000L	0.18	0.08	0.81	20	1.4 [0.18-3.1]
DCO	0.8 +5.7E-08*PORT_5000 -3.5E-05*GREEN_100	0.59	0.55	0.21	20	0.95 [0.51-1.7]	0.7 +1.0E-04*MAJORROADLENGTH_500	0.18	0.05	0.34	20	1.0 [0.63-2.1]
LIK	0.1 -5.0E-07*URBGREEN1000 +1.6E-03*POP100	0.23	0.16	0.36	20	0.26 [-0.15-1.6]						0.29 [-0.04-1.2]
UKM	0.7 +1.6E+01*DISTINVMAJOR1	0.18	0.06	0.58	19	0.93 [-0.50-2.2]	0.8 -3.4E-08*MMNATURAL_5000 +1.3E+02*DISTINVMAJORC2	0.54	0.50	0.14	19	0.66 [0.38-1.2]
UKO	0.7 +1.3E-06*HHOLD5000 +2.3E-08*TRAFMAJORLOAD100	0.50	0.43	0.20	20	1.00 [0.43-1.6]	0.5 +4.9E-05*POP500	0.19	0.09	0.27	20	0.78 [0.22-1.2]
BNL	4.1 +9.9E-08*PORT_5000 +1.6E-09*TRAFMAJORLOAD_1000 -1.3E-05*XCOORD	0.78	0.73	0.69	40	2.8 [0.96-5.7]	3.7 +8.6E-08*PORT_5000 -1.2E-05*XCOORD	0.76	0.72	0.54	40	2.3 [0.61-4.5]
GRU	0.8 +8.0E-05*ROADLENGTH_1000	0.22	0.11	2.15	20	4.0 [1.7-11]	0.5 +3.5E-05*ROADLENGTH_1000	0.32	0.20	0.75	20	1.9 [0.62-4.6]
GMU	0.7 +9.7E-09*TRAFLOAD_300 +2.7E-03*MAJORROADLENGTH50	0.59	0.49	0.22	19	0.98 [0.52-1.6]	3.5 -1.4E-01*SQRALT +7.1E-10*TRAFLOAD1000 +1.7E-03*ROADLENGTH_50	0.53	0.32	0.19	20	0.51 [0.18-1.1]
AUV	-0.2 +1.0E-02*ROADLENGTH25	0.21	0.11	0.39	20	0.09 [-0.80-0.82]	-0.1 +1.2E-06*RES300	0.18	0.10	0.19	20	0.13 [-0.18-0.73]
FPA	0.7 +1.6E-08*TRAFMAJORLOAD_500	0.63	0.54	0.61	20	1.3 [0.27-3.5]	0.7 +5.0E-05*MAJORROADLENGTH_500	0.19	0.12	0.32	20	0.82 [0.47-1.7]
HUG	0.2 +5.3E-04*INTMAJORINVDIST	0.16	0.03	0.30	20	0.28 [-0.10-1.2]						0.24 [-0.01-0.78]
SWL						1.3 [1.0-2.0]						0.87 [0.22-1.9]
ITU	2.1 +3.1E-05*LDRES100 +1.8E-06*INDUSTRY1000 +7.7E-04*POP_300 +3.7E-03*MAJORROADLENGTH_100	0.84	0.79	0.95	20	5.7 [2.8-10]	1.4 +6.0E-09*TRAFLOAD1000	0.48	0.38	0.65	20	2.6 [1.4-5.0]

IRO	0.0 +6.4E-07*HDRES_1000 +6.6E-06*LDRES_300 +2.6E-03*POP_100	0.90	0.86	0.36	20	2.5 [0.96-5.1]	1.3 +3.0E-02*HHOLD_300	0.28	0.18	0.30	20	1.5 [0.92-2.2]
SPB	6.9 -9.9E-06*URBGREEN_300 -1.2E-07*NATURAL_5000 +4.1E-04*INTINVDIST1 +1.4E+02*DISTINVHARBOUR	0.82	0.74	1.04	20	6.1 [2.4-9.5]	6.3 -1.2E-07*GREEN_5000	0.60	0.55	1.00	20	4.4 [2.1-7.4]
SPC	3.9 -1.8E-05*URBGREEN_300 -9.8E-08*NATURAL_5000 +4.8E-06*HHOLD_5000 +1.5E-03*ROADLENGHT_100	0.81	0.77	1.13	40	4.6 [0.84-9.5]	3.0 -1.3E-05*URBGREEN_300 -7.4E-08*NATURAL_5000 +3.4E-06*HHOLD_100_5000 +9.7E-04*ROADLENGHT_100	0.79	0.76	0.86	40	3.3 [0.42-7.4]
GRA	3.3 +8.2E-07*PORT_5000 +4.4E-05*TRAFNEAR +5.9E-10*TRAFMAJORLOAD_1000	0.59	0.47	1.05	20	4.7 [2.1-7.8]	1.4 +2.0E-08*HDRES_5000 +9.7E-07*PORT_5000	0.85	0.83	0.39	20	2.5 [1.3-4.9]
GRH						5.9 [2.5-11]	10.3 +5.2E-06*PORT_1000 -1.2E-05*URBGREEN_5000	0.24	0.01	2.12	20	4.3 [1.4-8.9]

Sulphur (S) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)
NOS	390.7 -3.8E-05*NATURAL1000 +2.8E-02*HHOLD500 +2.4E-01*ROADLENGTH50	0.65	0.48	42.25	19	413 [289-529]	402.7 -5.5E-05*NATURAL1000	0.16	0.07	92.69	19	369 [217-536]
SST	528.8 +4.0E-02*MAJORROADLENGTH_300	0.30	0.18	45.22	18	580 [451-642]	499.9 +3.2E-01*ROADLENGTH_50	0.34	0.24	28.31	18	525 [456-572]
FIH	683.9 -6.3E-04*URBNATURAL_300L +4.3E-03*POP_1000L	0.46	0.36	59.13	20	692 [528-821]	684.2 -6.4E-04*URBNATURAL_300L	0.31	0.23	50.98	20	649 [520-758]
DCO	723.1 -9.5E-03*GREEN_L_100 +1.8E-03*ROADLENGTH_1000	0.52	0.46	46.57	20	787 [656-879]	664.3 +7.5E-06*INDUSTRY_5000 +5.7E-06*PORT_5000	0.65	0.61	22.96	20	711 [656-757]
LIK						965 [841-1063]						888 [794-954]
UKM	775.0 -2.0E-03*NATURAL_300 +5.6E-03*HEAVYTRAFMAJOR	0.63	0.54	44.62	19	791 [666-924]	653.3 +6.6E-04*TRAFMAJOR +1.3E-01*MAJORROADLENGTH_100	0.39	0.32	40.73	19	687 [599-774]
UKO	840.4 -1.4E-05*NATURAL5000	0.14	0.05	94.78	20	790 [606-954]	734.2 -1.1E-05*NATURAL5000	0.14	0.02	73.29	20	696 [571-825]
BNL	1280.0 +1.3E-05*PORT_5000 +2.8E-03*TRAFNEAR -6.1E-04*YCOORD	0.48	0.39	99.54	40	1070 [812-1341]	1240.0 +1.1E-02*POP_500 -8.5E-04*YCOORD	0.32	0.27	110.51	40	915 [448-1107]
GRU	1133.8 +1.5E-04*HDRES_1000 +6.1E-05*PORT_5000 +7.0E-04*POP_5000	0.59	0.55	86.73	20	1310 [1083-1546]	1067.1 +2.9E-02*POP_500 +7.0E-01*HEAVYINTINVDIST	0.56	0.50	74.27	20	1203 [1049-1419]
GMU	872.5 -6.2E+02*URBGREEN_5000 +2.2E-01*POP100 +3.2E-01*ROADLENGTH50	0.52	0.37	47.04	20	898 [737-974]	1605.2 -3.7E+02*URBGREEN_300 -3.6E+01*SQRALT +6.1E-02*ROADLENGTH100 +3.6E-02*ROADLENGTH_300	0.78	0.70	30.04	20	832 [705-920]
AUV	460.8 +6.7E-06*RES5000 +3.4E-01*INTINVDIST2 +7.9E-01*ROADLENGTH25 +3.6E-01*APT300	0.68	0.55	39.45	20	665 [539-759]	810.6 -1.2E+01*SQRALT +2.5E-03*TRAFMAJOR +6.8E-01*ROADLENGTH25	0.58	0.53	32.81	20	616 [519-673]
FPA	835.8 +6.6E-05*INDUSTRY_1000 -4.0E-04*URBGREEN_500 -9.3E-06*NATURAL_5000 +6.1E-03*TRAFNEAR +3.7E-07*TRAFLOAD_1000	0.86	0.78	69.55	20	905 [564-1129]	804.6 -5.6E-06*NATURAL_5000 +1.0E+02*DISTINVMAJORC2 +2.5E-07*TRAFMAJORLOAD_1000	0.46	0.29	76.47	20	814 [598-997]
HUG						1264 [1098-1341]	1167.6 +1.2E-06*TRAFLOAD_1000	0.64	0.60	33.91	20	1228 [1128-1319]
SWL						903 [829-965]						913 [710-1085]

ITU	1422.8 +3.7E-03*POP_1000 -2.5E+01*SQRALT +1.9E-01*MAJORROADLENGTH_100	0.62	0.57	77.17	20	1134 [962-1367]	924.9 +3.7E-04*RES300 +1.1E-06*TRAFLOAD500	0.52	0.39	75.45	20	1066 [901-1279]
IRO	997.9 +3.4E-05*LDRES_1000 +1.3E-01*POP_100	0.46	0.38	58.41	20	1086 [958-1262]	926.0 +1.2E-03*POP_1000 +2.4E-01*INTMAJORINVDIST 2	0.44	0.33	47.70	20	972 [883-1098]
SPB	912.0 +5.4E-03*HDRES_100 +4.8E-02*HHOLD_300 +3.3E+02*DISTINVNEARC1 +4.3E-06*TRAFLOAD_100	0.88	0.85	54.04	20	1229 [922-1434]	787.2 +1.5E-05*PORT_5000 +3.5E-02*HHOLD_300 +1.7E-07*TRAFLOAD_1000	0.67	0.58	70.11	20	1042 [821-1245]
SPC	1225.4 +2.0E-05*LDRES_5000 -6.5E-06*NATURAL_5000 +2.6E-01*HHOLD_100 -2.5E+01*SQRALT	0.81	0.77	96.80	40	1104 [694-1457]	970.8 +1.8E-05*LDRES_5000 -8.4E-06*NATURAL_5000 +1.9E-01*HHOLD_100 +2.8E+03*DISTINVMAJORC2	0.83	0.81	63.11	40	953 [569-1245]
GRA	1973.8 -1.9E+01*SQRALT	0.52	0.46	76.34	20	1793 [1622-1956]	1790.3 +2.6E-06*LDRES_5000 +1.2E-03*HHOLD_1000 -2.3E+01*SQRALT	0.73	0.67	49.11	20	1657 [1504-1801]
GRH						1705 [1451-2068]						1514 [1239-2008]

Silicon (Si) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)
NOS	123.0 +3.2E+05*DISTINVMAJORC2 +6.9E-01*HHOLD300 +1.2E+01*MAJORROADLENGTH25	0.86	0.82	235.67	19	728 [82-2162]	140.9 +2.8E+03*DISTINVMAJORC1 -6.2E+00*SQRALT +1.1E-04*HDRES500 -2.9E-05*URBGREEN5000	0.84	0.77	35.54	15	92 [-24.08-237]
SST	8.6 +4.3E-03*HLDRES300 +6.6E-04*TRAFLOAD_50	0.82	0.77	400.05	18	1433 [19-3399]	4.6 +5.8E-03*HLDRES100 +7.8E-05*TRAFLOAD_50	0.72	0.65	69.06	18	213 [-25.04-446]
FIH	413.9 +2.6E+00*HHOLD_100L +1.8E-04*TRAFMAJORLOAD_50	0.68	0.57	235.74	20	681 [128-1735]	86.2 -1.1E-03*URBNATURAL_100L +3.0E-01*HHOLD_100L +1.3E-06*TRAFLOAD_300	0.74	0.72	27.53	20	112 [17-209]
DCO	281.7 +1.6E-05*PORT_5000 -1.2E-03*GREEN_300 -1.6E-02*GREEN_100	0.66	0.60	74.22	20	281 [64-474]	30.0 +1.6E-03*MAJORROADLENGTH_1000	0.25	0.17	14.84	20	49 [16-86]
LIK	551.2 +1.1E-02*TRAFNEAR	0.40	0.26	148.44	20	655 [456-1046]	93.8 +1.6E-04*INDUSTRY500 +3.0E-09*TRAFLOAD500	0.24	0.14	33.86	20	117 [53-207]
UKM	225.7 -7.2E-04*NATURAL_500 +1.8E-06*HEAVYTRAFMAJORLOAD_1000	0.72	0.67	60.76	19	264 [117-477]	43.1 +2.5E-07*HEAVYTRAFMAJORLOAD_1000	0.55	0.47	10.91	19	51 [33-79]
UKO	153.3 +3.3E-06*HLDRES5000	0.54	0.49	69.26	20	304 [126-441]	37.1 +5.0E-07*HLDRES5000	0.26	0.17	18.50	20	60 [31-95]
BNL	282.0 +6.6E-04*POP_5000 +3.9E-06*TRAFMAJORLOAD_300	0.40	0.26	203.91	40	475 [151-1531]	146.0 +2.6E-03*TRAFNEAR -1.1E-04*XPLUSY	0.46	0.39	39.09	40	101 [38-285]
GRU	264.5 +1.5E-04*PORT_5000 +1.1E-03*POP_5000 +1.9E+00*ROADLENGTH_25 +7.0E-03*PM10_2008_HT100	0.75	0.70	112.98	20	658 [382-1013]	72.4 +3.5E-04*POP_5000 +1.4E-03*PM10_2008_HT100	0.49	0.40	32.89	20	156 [106-269]
GMU	163.1 +2.7E+02*HLDRES_500 +4.9E-07*TRAFLOAD_1000 +1.0E+00*MAJORROADLENG_50	0.73	0.64	68.75	19	493 [324-769]	16.6 +6.8E+01*HLDRES_300 +1.0E-06*HEAVYTRAFLOAD1000 +3.0E-03*ROADLENGTH_1000 +1.6E-01*MAJORROADLENGT_50	0.74	0.55	12.90	19	116 [73-135]
AUV	470.0 -7.5E-06*NATURAL5000 +1.5E-04*TRAFLOAD25	0.54	0.48	58.23	20	402 [292-606]	60.0 -2.0E-05*NATURAL1000 +2.2E-01*ROADLENGTH25	0.26	0.15	8.25	20	63 [51-89]
FPA	339.0 -1.5E-05*NATURAL_5000 +1.7E-02*TRAFMAJOR	0.74	0.61	152.83	20	467 [79-1030]	49.1 +5.1E-05*INDUSTRY_500 +2.8E-04*POP_1000 +1.4E-01*MAJORROADLENGTH_50	0.59	0.50	15.73	20	66 [14-115]
HUG	396.6 +1.5E+04*DISTINVMAJORC2 +2.1E-02*MAJORROADLENGTH_1000	0.38	0.31	122.18	20	519 [219-817]						97 [76-146]
SWL						494 [395-670]						112 [71-143]
ITU	971.3 -2.2E-05*NATURAL5000	0.67	0.61	283.45	20	1233	79.2 +7.7E-05*INDUSTRY1000	0.65	0.54	35.41	20	184

	+9.7E-01*POP_100 +1.7E+00*MAJORROADLENGTH_50					[604-1970]	+1.6E-02*POP_300 +7.0E-02*ROADLENGHT100					[86-278]
IRO	351.9 +1.7E-03*LDRES_300 +6.0E+01*HHOLD_300 +3.9E-04*TRAFLOAD_25	0.67	0.60	287.89	20	1086 [598-2258]	124.3 +7.6E-01*HHOLD_1000 +3.2E-01*ROADLENGTH_50	0.51	0.43	62.32	20	231 [127-424]
SPB	1526.9 -5.5E+01*SQRALT	0.49	0.41	187.17	20	1244 [648-1644]	-33.9 +5.1E-04*INDUSTRY_500 +1.1E-01*HHOLD_100 +3.1E-04*HHOLD_100_5000	0.48	0.35	37.72	20	181 [106-281]
SPC	1668.6 -6.8E-05*URBGREEN_5000 -5.6E+01*SQRALT +7.5E+00*MAJORROADLENGHT_25	0.62	0.58	198.75	40	1283 [648-2039]	190.0 -3.5E-05*GREEN_1000 +4.8E-01*MAJORROADLENGHT_25	0.18	0.09	39.47	40	185 [106-281]
GRA	1193.8 -5.1E-05*URBGREEN_5000 +9.0E-05*TRAFMAJORLOAD_25 +3.1E-01*ROADLENGTH_100	0.58	0.39	197.03	20	1267 [791-1859]						292 [189-372]
GRH	1035.9 +1.4E-03*HDRES_300	0.26	0.18	231.02	20	1108 [632-1569]	238.5 +3.5E-04*HDRES_300	0.27	0.19	56.06	20	256 [154-362]

Vanadium (V) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)
NOS	0.5 +5.1E+02*DISTINVMAJORC2 +9.8E-04*HHOLD500 +1.5E-07*HDRES5000	0.87	0.83	0.51	19	2.4 [0.65-4.5]	0.8 +1.0E-06*PORT5000	0.51	0.42	0.75	19	1.4 [-0.47-3.1]
SST	0.9 +4.9E-06*HLDRES300 +1.2E-05*INDUSTRY300 +1.1E-02*MAJORROADLENGTH_50	0.76	0.70	0.49	18	2.4 [0.89-3.9]	0.7 + 1.1E-06*HLDRES500	0.19	0.07	0.48	16	1.4 [0.42-2.4]
FIH	4.1 +7.6E-07*PORT_5000L -6.3E-06*GREEN_500L	0.32	0.16	1.92	20	3.9 [1.2-11]	1.5 +4.5E-06*POP_5000L +7.3E-03*HHOLD_100L	0.47	0.30	0.85	20	2.5 [0.82-4.7]
DCO	112.3 +8.9E-08*PORT_5000 +2.0E-05*XMINUSY	0.67	0.59	0.43	20	2.7 [1.8-3.9]	2.1 +8.5E-08*PORT_5000	0.44	0.38	0.36	20	2.4 [1.7-3.3]
LIK	0.5 +1.4E-06*LDRES500 +1.2E-06*INDUSTRY1000	0.39	0.21	0.50	20	1.6 [0.80-2.8]	0.7 +1.3E-05*LDRES100 -5.1E-07*URBGREEN1000 +4.7E-10*TRAFMAJORLOAD_300	0.34	0.27	0.33	20	1.0 [0.57-1.8]
UKM	1.8 +1.8E-04*POP_300 -4.1E-02*SQRALT -1.1E-05*NATURAL_300	0.55	0.36	0.22	19	1.6 [1.1-2.1]	0.8 +1.6E-04*POP_300 +1.5E+02*DISTINVMAJOR2	0.35	0.26	0.26	19	1.1 [0.59-1.6]
UKO	1.9 +1.5E-06*POP5000	0.40	0.29	0.48	20	2.5 [1.1-3.3]	1.3 +1.3E-08*HLDRES5000	0.35	0.24	0.41	20	1.9 [0.90-2.9]
BNL	6.7 +2.3E-07*PORT_5000 -2.0E-05*XCOORD	0.72	0.67	1.25	40	4.5 [1.3-10]	5.6 +2.0E-07*PORT_5000 -1.8E-05*XCOORD	0.68	0.63	1.16	40	3.7 [1.3-9.0]
GRU	37.6 +1.1E-05*LDRES_100 +3.1E-07*PORT_5000 -1.4E-05*XCOORD	0.62	0.52	0.30	20	2.1 [1.1-3.2]	1.6+1.4E-07*PORT_5000+3.6E-06*TRAFMAJOR+3.4E-06*HEAVYTRAFMAJORLOAD_50	0.66	0.48	0.14	20	1.8 [1.5-2.2]
GMU	0.5 -4.3E-01*URBGREEN_100	0.10	0.04	0.17	20	0.42 [0.06-0.71]						0.33 [0.10-0.62]
AUV	0.6 -5.1E-07*NATURAL1000	0.10	0.00	0.33	20	0.55 [0.01-1.2]						0.32 [0.20-0.49]
FPA	1.8 -3.6E-08*NATURAL_5000 +7.2E-03*ROADLENGTH_25	0.64	0.59	0.31	20	1.8 [1.00-2.7]	1.2 +1.8E-07*HDRES_1000	0.39	0.28	0.24	20	1.3 [0.81-2.0]
HUG	1.3 +8.6E-05*MAJORROADLENGTH_500	0.22	0.15	0.20	20	1.4 [0.91-1.8]						0.74 [0.54-1.0]
SWL						0.85 [0.57-1.2]						0.67 [0.22-0.92]
ITU	2.2 +5.8E-07*RES1000 +2.3E-03*MAJORROADLENGTH_100	0.67	0.62	0.64	20	3.7 [2.3-5.6]	1.4 +3.2E-09*TRAFLOAD1000	0.43	0.36	0.37	20	2.0 [1.3-2.7]
IRO	3.6 +6.2E-08*LDRES_5000 -1.5E-01*SQRALT +1.7E+01*DISTINVMAJOR1	0.72	0.68	0.59	20	4.2 [2.3-6.5]	1.9 +4.8E-08*LDRES_5000	0.20	0.09	0.54	20	2.9 [1.7-4.0]
SPB	15.6 -2.9E-07*GREEN_5000 +7.8E-03*ROADLENGHT_50	0.67	0.54	2.87	20	12 [6.7-22]	1.5 +8.4E-07*PORT_5000 +9.0E-04*ROADLENGHT_300	0.61	0.48	2.60	20	9.5 [2.6-17]
SPC	8.6 +6.9E-07*PORT_5000	0.82	0.79	2.62	40	8.0	2.9 +3.4E-06*INDUSTRY_1000	0.83	0.78	1.93	40	6.4

	+1.0E-05*HHOLD_5000 -6.8E-01*SQRALT					[-1.12-22]	-3.5E-05*URBGREEN_300 +1.3E-05*HHOLD_5000 -3.8E-01*SQRALT +3.8E-04*ROADLENGHT_300					[0.86-17]
GRA	4.8 +2.2E-07*INDUSTRY_5000 +2.2E-06*PORT_5000	0.75	0.72	1.33	20	7.6 [4.5-13]	4.5 +3.5E-08*HDRES_5000 +2.2E-06*PORT_5000 -9.8E-07*URBGREEN_1000 -3.6E-08*NATURAL_5000	0.93	0.92	0.61	20	6.2 [3.9-12]
GRH	11.6 +1.9E-05*PORT_1000 -1.3E-05*URBGREEN_1000	0.52	0.38	4.13	20	12 [6.7-26]	9.2 +1.7E-05*PORT_1000 -1.0E-05*URBGREEN_1000	0.54	0.41	3.52	20	9.7 [5.4-22]

Zinc (Zn) models

	PM_{10}						$PM_{2.5}$					
Study Area	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)	LUR model	R^2	R^2 LOOCV	RMSE LOOCV(ng/m^3)	N[a]	Mean Conc [range] (ng/m^3)
NOS	17.5 +1.7E-01*MAJORROADLENGTH50 -3.5E-06*NATURAL1000	0.78	0.75	5.27	19	22 [7.7-43]	10.7+4.2E-04*MAJORROADLENGTH 1000 +8.8E+01*DISTINVMAJORC1 +2.7E-03*HHOLD500 -2.8E-01*SQRALT	0.80	0.70	2.40	19	13 [3.8-23]
SST	9.1 +1.2E-06*INDUSTRY5000 +1.8E-03*INTMAJORINVDIST +6.3E-06*TRAFLOAD_50	0.89	0.87	3.62	18	24 [7.5-50]	10.1 +6.1E-07*TRAFLOAD_100	0.35	0.24	2.26	18	12 [7.1-16]
FIH	27.8 -3.2E-05*GREEN_500L	0.20	0.10	7.43	20	23 [13-51]	17.9 -3.0E-06*GREEN_1000L	0.21	0.16	2.21	20	16 [12-19]
DCO	15.3 -4.7E-05*GREEN_300 +1.2E-07*TRAFLOAD_500	0.77	0.72	2.88	20	20 [12-33]	10.6 +2.9E-03*HHOLD_500	0.18	0.11	1.88	20	13 [9.4-18]
LIK	30.3 +6.9E-04*LDRES100 +2.4E-09*TRAFLOAD500	0.25	0.14	21.77	20	58 [28-119]						49 [20-113]
UKM	10.0 +1.6E-03*ROADLENGTH_300 +1.3E-02*MAJORROADLENGTH_100	0.68	0.58	3.38	19	18 [13-30]	10.5-7.9E-06*NATURAL_1000+3.4E-02*ROADLENGTH_25+6.2E-03*MAJORROADLENGTH_100+5.4E-04*MAJORROADLENGTH_100_500	0.73	0.63	1.82	19	12 [7.3-18]
UKO	14.6 +1.1E-07*HLDRES5000 +5.3E-01*HEAVYINTINVDIST2 +1.3E-06*HEAVYTRAFLOAD500	0.80	0.77	3.69	20	23 [12-45]	10.0 +2.1E-05*HHOLD5000 +4.9E-05*TRAFMAJOR +2.7E+02*DISTINVMAJOR2	0.70	0.63	2.19	20	15 [8.1-22]
BNL	112.0 +7.1E-07*TRAFLOAD_300 +2.7E-04*XMINUSY	0.65	0.57	16.95	40	43 [19-139]	85.6 +2.0E-07*TRAFMAJORLOAD_300 +2.0E-04*XMINUSY	0.66	0.58	11.19	40	28 [13-87]
GRU	33.7 +9.4E-05*POP_5000 +1.4E-05*TRAFMAJORLOAD_25 +1.6E-05*NOX_2008_HT100	0.61	0.50	10.99	19	60 [37-88]	23.7 +5.5E-06*LDRES_1000 +9.5E-06*INDUSTRY_1000 +2.9E-06*PORT_5000 -1.5E-05*NATURAL_1000 +1.4E-04*HEAVYTRAFLOAD_50	0.74	0.58	5.18	19	41 [26-94]
GMU	16.0 +5.0E-02*ROADLENGTH50 +3.4E-03*ROADLENGTH_500 +1.1E-01*ROADLENGTHDTV5000_50	0.74	0.65	7.37	20	34 [11-70]	13.6 +1.4E+01*INDUSTRY1000 -2.8E+01*URBGREEN_5000 +5.1E-04*ROADLENGTH50_500 +3.0E-02*MAJORROADLENGTH_50 +1.9E-02*ROADLENGTH50	0.69	0.53	2.79	20	20 [13-30]
AUV	9.1 +8.5E-07*RES5000 -1.6E-05*URBGREEN5000 +3.7E-04*TRAFNEAR +4.9E-02*BLDG300	0.86	0.79	2.83	20	32 [17-41]	20.5 +1.9E-05*RES300 -3.4E-07*NATURAL5000 +9.4E-08*TRAFLOAD500	0.57	0.44	3.29	20	22 [11-30]
FPA	15.5 +5.8E-04*TRAFNEAR +9.6E-04*MAJORROADLENGTH_300_1000 +5.7E-03*MAJORROADLENGTH_300	0.80	0.73	11.07	20	40 [18-101]	15.3 +1.3E-07*LDRES_5000 -6.5E-07*URBGREEN_5000 -2.3E-07*NATURAL_5000 +3.1E-04*TRAFNEAR +4.3E-04*MAJORROADLENGTH_1000	0.85	0.76	3.66	20	23 [12-41]

HUG						40 [27-80]						32 [24-73]
SWL	45.6 -1.5E+00*SQRALT +5.9E-08*TRAFLOAD1000 +7.5E-06*TRAFLOAD50 +2.2E-02*MAJORROADLENGTH50	0.93	0.86	4.86	19	36 [20-65]	7.0 +5.1E-06*LDRES1000 +1.8E-05*HDRES500 +5.5E-04*HEAVYTRAFMAJORLOAD50	0.81	0.75	3.14	19	23 [14-37]
ITU	39.8 +5.1E-04*RES100 -8.2E-07*NATURAL5000 +1.2E-03*POP_500 +5.9E-06*TRAFLOAD50	0.93	0.91	5.93	20	62 [27-104]	24.5 +2.4E-05*LDRES300 -3.5E-07*NATURAL5000 +1.2E-06*TRAFLOAD50 +3.9E-08*TRAFLOAD1000	0.85	0.78	3.59	20	35 [19-50]
IRO	26.9 +4.9E-07*HDRES_5000 +3.0E-05*TRAFLOAD_25	0.85	0.81	8.41	20	45 [23-82]	-462.2 +9.1E+01*DISTINVMAJOR1 +1.9E-08*TRAFLOAD_1000 +9.7E-05*XPLUSY	0.78	0.75	2.47	20	23 [14-33]
SPB	47.6 +5.2E-06*INDUSTRY_5000 -3.8E+00*SQRALT +6.8E-08*TRAFMAJORLOAD_1000	0.78	0.71	21.46	20	105 [43-171]	-9.7 +4.6E-06*INDUSTRY_5000 +3.4E-08*TRAFLOAD_1000 +1.0E+03*DISTINVHARBOUR	0.77	0.73	11.13	20	58 [28-101]
SPC	19.6 +8.0E-06*PORT_5000 +2.5E-05*POP_500_5000 +1.7E-03*POP_500 +6.6E-06*TRAFMAJORLOAD_50	0.86	0.82	19.68	40	73 [12-171]	13.0 +2.2E-06*INDUSTRY_5000 +3.0E-05*POP_5000 -1.3E+00*SQRALT +8.1E-08*TRAFMAJORLOAD_500	0.85	0.81	11.02	40	39 [5.8-101]
GRA	89.8 -5.1E-06*URBGREEN_5000 +7.7E-03*HHOLD_100 -2.4E+00*SQRALT +5.0E-06*TRAFMAJORLOAD_25 +8.8E-04*ROADLENGTH_500	0.86	0.73	10.75	20	68 [35-106]	12.0 +1.3E-04*INDUSTRY_300 +7.7E-06*PORT_5000 -3.6E-06*URBGREEN_5000 -7.9E-07*NATURAL_5000 +5.7E-04*ROADLENGTH_1000	0.93	0.90	4.60	20	42 [23-77]
GRH	14.3 -6.0E-05*URBGREEN_300 +3.3E+01*DISTINVMAJORC1 +7.9E-03*ROADLENGTH_100	0.64	0.60	3.09	20	21 [15-33]	8.8 +5.0E-03*ROADLENGTH_100	0.59	0.54	1.51	20	13 [8.7-17]

[a]Number of sites.

Table S3. Detection limits (DL, ng/m^3) and % samples > DL in $PM_{2.5}$ and PM_{10}

Element	DL[a]		PM_{10}	$PM_{2.5}$
	Min	Max		
Cu	0.5	1.5	99	95
Fe	1.3	6.1	100	100
K	0.9	20.9	100	100
Ni	0.1	0.3	90	88
S	0.0	17.0	100	100
Si	4.2	16.1	100	100
V	0.1	0.3	96	95
Zn	0.6	2.0	100	100

[a]range across study areas

Table S4. Coefficient of variation (%) obtained from duplicate samples by study area

Study Area	N	Cu	Fe	K	Ni	S	Si	V	Zn
NOS	19	4.2	4.8	7.0	34.7	5.9	7.5	13.0	3.6
SST	22	141.8	6.4	6.8	304.4	3.4	6.8	16.2	70.0
FIH	24	13.6	19.7	15.6	31.4	5.2	27.7	14.9	12.0
DCO	19	12.8	6.3	4.9	96.5	4.4	15.6	7.0	8.4
LIK	20	6.6	4.1	4.0	104.1	2.9	5.6	20.8	3.1
UKM	15	11.8	10.7	8.4	100.4	5.4	7.8	20.1	7.3
UKO	14	6.3	6.7	5.7	132.2	5.2	6.4	18.8	7.2
BNL	19	9.5	8.2	7.4	13.1	6.0	13.1	8.8	11.6
GRU	23	8.6	4.5	26.8	22.1	11.8	4.7	19.8	6.0
GMU	13	5.8	3.2	7.0	80.1	3.0	4.0	58.0	9.8
AUV	23	15.9	16.4	4.8	147.9	3.2	16.2	63.0	9.5
FPA	22	5.3	8.4	8.4	39.4	8.4	4.3	14.4	33.8
HUG	23	6.8	5.6	9.4	102.8	6.7	5.5	18.8	9.0
SWL	14	12.6	14.6	7.3	23.1	7.7	12.8	24.6	6.0
ITU	24	9.0	10.5	8.4	18.4	6.4	12.6	17.6	8.6
IRO	25	5.3	5.2	4.5	18.3	4.4	4.4	13.4	5.3
SPB/SPC	25	7.4	6.7	9.1	13.9	4.3	7.7	7.9	7.3
GRA	22	5.7	4.5	5.8	14.7	3.4	4.4	8.0	5.8
GRH	21	18.9	6.2	7.9	19.4	2.3	5.3	4.2	13.4
Median		8.6	6.4	7.3	34.7	5.2	6.8	16.2	8.4

Table S5: Tables showing correlations (R^2) between prediction of LUR models for 8 elements (both PM_{10} and $PM_{2.5}$ fractions) and NO_2, NO_x, $PM_{2.5}$ABS, $PM_{2.5}$, PM_{10} and PM_{coarse} at 20 sites not used in the modelling.Copper (Cu)

Compositon	Study Area	$PM_{2.5}$ NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse	PM_{10} NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse
Copper (Cu)	NOS	0.18	0.25	0.29	0.18	0.15	0.02	0.72	0.67	0.91	0.33	0.60	0.22
	SST	0.78	0.80	0.82	0.41	0.59	0.61	0.64	0.64	0.57	0.30	0.41	0.43
	FIH	0.60	0.67	0.47	0.53	0.49	0.41	0.47	0.56	0.58	0.33	0.53	0.45
	DCO	0.79	0.81	0.77	0.53	0.71	0.44	0.92	0.84	0.71	0.43	0.82	0.74
	LIK	0.15	0.07	0.23	0.45	0.52	0.02	0.36	0.29	0.33	0.32	0.38	0.04
	UKM	0.60	0.75	0.72	0.23	0.02	0.15	0.62	0.67	0.51	0.32	0.12	0.32
	UKO	0.74	0.60	0.11	0.57	0.23	0.19	0.95	0.94	0.03	0.86	0.34	0.52
	BNL	0.79	0.76	0.92	0.80	0.81	0.41	0.75	0.74	0.59	0.40	0.63	0.72
	GRU	0.67	0.66	0.81	0.53	0.44	0.79	0.75	0.72	0.80	0.45	0.42	0.79
	GMU	0.80	0.71	0.75	0.01	0.00	0.80	0.54	0.47	0.37	0.02	0.04	0.67
	AUV	0.01	0.03	0.01	0.20	0.09	0.01	0.73	0.66	0.80	0.55	0.45	0.65
	FPA	0.30	0.23	0.55	0.77	0.66	0.59	0.74	0.64	0.85	0.66	0.64	0.82
	HUG	0.33	0.34	0.48	0.17	0.07	0.21	0.38	0.35	0.33	0.01	0.22	0.55
	SWL	0.78	0.51	0.54	0.23	0.82	0.66	0.60	0.60	0.76	0.33	0.51	0.61
	ITU	0.20	0.06	0.30	0.07	0.61	0.77	0.40	0.25	0.50	0.31	0.88	0.87
	IRO	0.28	0.29	0.66	0.51	0.39	0.53	0.48	0.48	0.82	0.76	0.65	0.74
	SPB	0.68	0.70	0.66	0.64	0.38	0.33	0.62	0.68	0.69	0.62	0.20	0.20
	SPC	0.57	0.52	0.68	0.47	0.16	0.47	0.42	0.45	0.59	0.46	0.05	0.36
	GRA	0.11	0.23	0.35	0.26	0.50	0.21	0.26	0.43	0.72	0.55	0.36	0.47
	GRH	0.55	0.57	0.01	0.15	0.58	0.58	0.35	0.28	0.55	0.66	0.67	0.43
	Median	0.60	0.57	0.55	0.47	0.44	0.41	0.60	0.60	0.59	0.40	0.42	0.52
Iron (Fe)	NOS	0.79	0.72	0.70	0.36	0.80	0.38	0.68	0.66	0.80	0.28	0.74	0.29
	SST	0.63	0.74	0.56	0.08	0.51	0.52	0.79	0.74	0.86	0.45	0.79	0.80
	FIH	0.39	0.53	0.10	0.28	0.65	0.67	0.53	0.56	0.34	0.28	0.44	0.39
	DCO	0.92	0.86	0.79	0.46	0.85	0.69	0.93	0.85	0.70	0.45	0.82	0.72
	LIK							0.30	0.19	0.54	0.45	0.66	0.18
	UKM	0.21	0.16	0.09	0.14	0.41	0.35	0.39	0.36	0.30	0.34	0.28	0.52
	UKO	0.87	0.93	0.16	0.88	0.56	0.66	0.93	0.97	0.03	0.76	0.31	0.44
	BNL	0.85	0.82	0.74	0.51	0.65	0.68	0.86	0.81	0.73	0.46	0.85	0.70
	GRU	0.82	0.79	0.73	0.30	0.28	0.48	0.69	0.68	0.95	0.71	0.72	0.65
	GMU	0.71	0.66	0.65	0.02	0.00	0.86	0.80	0.75	0.72	0.07	0.01	0.82
	AUV	0.38	0.68	0.35	0.16	0.49	0.35	0.63	0.36	0.68	0.15	0.17	0.67
	FPA	0.64	0.55	0.72	0.66	0.80	0.87	0.76	0.66	0.86	0.65	0.64	0.82
	HUG	0.70	0.54	0.38	0.04	0.10	0.10	0.51	0.52	0.74	0.04	0.10	0.44
	SWL	0.77	0.67	0.52	0.62	0.54	0.35	0.85	0.60	0.53	0.43	0.80	0.61
	ITU	0.45	0.27	0.59	0.24	0.89	0.89	0.46	0.26	0.59	0.22	0.95	0.92
	IRO	0.34	0.37	0.80	0.50	0.44	0.64	0.54	0.48	0.84	0.74	0.71	0.86
	SPB	0.65	0.69	0.76	0.62	0.20	0.20	0.82	0.83	0.75	0.36	0.44	0.35
	SPC	0.35	0.34	0.41	0.26	0.10	0.45	0.60	0.60	0.74	0.59	0.19	0.63
	GRA	0.43	0.58	0.60	0.50	0.30	0.53	0.24	0.44	0.70	0.51	0.53	0.51
	GRH	0.25	0.13	0.63	0.85	0.63	0.42	0.14	0.07	0.69	0.44	0.21	0.34
	Median	0.64	0.67	0.59	0.33	0.53	0.52	0.63	0.60	0.72	0.45	0.59	0.63

Compositon	Study Area	$PM_{2.5}$ NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse	PM_{10} NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse
Potassium	NOS	0.41	0.35	0.42	0.63	0.24	0.36	0.20	0.22	0.17	0.56	0.62	0.49
(K)	SST	0.48	0.43	0.51	0.16	0.80	0.77	0.69	0.69	0.73	0.19	1.00	1.00
	FIH	0.09	0.11	0.10	0.67	0.06	0.06	0.29	0.36	0.74	0.99	0.26	0.18
	DCO	0.04	0.05	0.10	0.47	0.13	0.00	0.02	0.04	0.02	0.09	0.03	0.00
	LIK												
	UKM							0.21	0.29	0.40	0.30	0.13	0.02
	UKO	0.66	0.48	0.10	0.52	0.17	0.15	0.66	0.48	0.10	0.52	0.17	0.15
	BNL	0.14	0.15	0.24	0.31	0.09	0.01	0.57	0.49	0.46	0.23	0.64	0.47
	GRU							0.40	0.38	0.63	0.65	0.48	0.75
	GMU	0.01	0.02	0.01	0.25	0.30	0.00	0.20	0.25	0.31	0.42	0.58	0.08
	AUV	0.45	0.50	0.52	0.75	0.44	0.33	0.37	0.26	0.39	0.68	0.40	0.39
	FPA	0.36	0.30	0.51	0.38	0.39	0.55	0.64	0.48	0.63	0.66	0.61	0.76
	HUG	0.49	0.55	0.14	0.31	0.17	0.06	0.66	0.53	0.25	0.55	0.27	0.08
	SWL	0.16	0.30	0.13	0.72	0.00	0.00	0.85	0.70	0.64	0.44	0.82	0.66
	ITU	0.20	0.27	0.12	0.11	0.08	0.05	0.18	0.09	0.44	0.14	0.50	0.57
	IRO	0.33	0.32	0.55	0.49	0.48	0.49	0.34	0.39	0.75	0.86	0.80	0.71
	SPB	0.44	0.39	0.36	0.38	0.45	0.45	0.41	0.43	0.26	0.03	0.56	0.48
	SPC	0.11	0.10	0.10	0.14	0.38	0.28	0.09	0.11	0.13	0.15	0.48	0.39
	GRA	0.10	0.03	0.04	0.47	0.05	0.01	0.45	0.49	0.45	0.18	0.24	0.89
	GRH	0.08	0.03	0.98	0.52	0.25	0.19	0.19	0.11	0.65	0.58	0.40	0.26
	Median	0.20	0.30	0.24	0.38	0.25	0.24	0.37	0.36	0.40	0.42	0.56	0.48
Nickel	NOS	0.43	0.29	0.40	0.41	0.11	0.08	0.43	0.29	0.40	0.41	0.11	0.08
(Ni)	SST							0.55	0.50	0.56	0.45	0.25	0.26
	FIH	0.33	0.20	0.01	0.00	0.15	0.16	0.17	0.19	0.02	0.00	0.53	0.63
	DCO	0.68	0.61	0.75	0.70	0.78	0.51	0.66	0.56	0.47	0.70	0.75	0.82
	LIK							0.00	0.03	0.02	0.33	0.35	0.01
	UKM	0.89	0.92	0.69	0.13	0.00	0.13	0.80	0.89	0.57	0.11	0.02	0.08
	UKO	0.65	0.46	0.08	0.69	0.19	0.26	0.83	0.67	0.00	0.59	0.10	0.15
	BNL	0.43	0.34	0.26	0.15	0.25	0.53	0.54	0.45	0.35	0.20	0.34	0.66
	GRU	0.12	0.11	0.31	0.34	0.57	0.16	0.12	0.11	0.31	0.34	0.57	0.16
	GMU	0.33	0.28	0.32	0.16	0.17	0.45	0.90	0.89	0.88	0.04	0.03	0.84
	AUV	0.00	0.01	0.05	0.00	0.00	0.01	0.03	0.00	0.01	0.03	0.16	0.07
	FPA	0.52	0.40	0.60	0.74	0.85	0.89	0.74	0.58	0.69	0.42	0.43	0.64
	HUG							0.07	0.07	0.25	0.01	0.38	0.43
	SWL												
	ITU	0.09	0.06	0.07	0.02	0.01	0.04	0.57	0.39	0.66	0.31	0.93	0.77
	IRO	0.01	0.03	0.00	0.00	0.01	0.00	0.21	0.29	0.17	0.01	0.00	0.02
	SPB	0.50	0.31	0.38	0.12	0.94	0.95	0.75	0.57	0.58	0.25	0.90	0.84
	SPC	0.57	0.36	0.38	0.34	0.57	0.46	0.57	0.36	0.38	0.33	0.57	0.45
	GRA	0.01	0.03	0.01	0.05	0.28	0.00	0.18	0.37	0.27	0.23	0.45	0.09
	GRH	0.16	0.15	0.03	0.01	0.00	0.00						
	Median	0.43	0.29	0.31	0.15	0.19	0.16	0.55	0.39	0.38	0.31	0.38	0.43

Compositon	Study Area	$PM_{2.5}$ NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse	PM_{10} NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse
Sulphur (S)	NOS	0.36	0.26	0.21	0.38	0.10	0.08	0.54	0.46	0.40	0.79	0.28	0.23
	SST	0.83	0.82	0.75	0.35	0.69	0.71	0.75	0.59	0.75	0.54	0.48	0.48
	FIH	0.01	0.02	0.03	0.31	0.04	0.05	0.30	0.33	0.03	0.12	0.57	0.63
	DCO	0.44	0.36	0.45	0.46	0.52	0.68	0.46	0.44	0.53	0.97	0.65	0.44
	LIK												
	UKM	0.20	0.19	0.16	0.02	0.00	0.01	0.00	0.02	0.01	0.00	0.11	0.01
	UKO	0.20	0.09	0.00	0.27	0.00	0.04	0.20	0.09	0.00	0.27	0.00	0.04
	BNL	0.46	0.46	0.40	0.38	0.20	0.25	0.54	0.50	0.37	0.27	0.25	0.62
	GRU	0.37	0.37	0.44	0.48	0.50	0.40	0.43	0.42	0.58	0.62	0.61	0.66
	GMU	0.02	0.03	0.03	0.55	0.35	0.00	0.00	0.01	0.02	0.46	0.45	0.03
	AUV	0.07	0.01	0.06	0.31	0.05	0.05	0.30	0.23	0.49	0.42	0.56	0.37
	FPA	0.42	0.22	0.35	0.30	0.27	0.48	0.54	0.36	0.66	0.56	0.56	0.74
	HUG	0.77	0.67	0.61	0.01	0.02	0.22						
	SWL												
	ITU	0.29	0.21	0.26	0.16	0.25	0.31	0.40	0.25	0.54	0.31	0.76	0.72
	IRO	0.33	0.34	0.80	0.61	0.54	0.69	0.05	0.11	0.00	0.01	0.04	0.03
	SPB	0.35	0.21	0.35	0.23	0.49	0.55	0.27	0.21	0.40	0.10	0.19	0.21
	SPC	0.49	0.32	0.30	0.25	0.46	0.50	0.34	0.18	0.17	0.21	0.74	0.37
	GRA	0.08	0.09	0.04	0.34	0.18	0.02	0.02	0.06	0.04	0.19	0.23	0.00
	GRH												
	Median	0.35	0.22	0.30	0.31	0.27	0.28	0.34	0.25	0.40	0.31	0.56	0.37
Silicon (Si)	NOS	0.04	0.06	0.08	0.14	0.24	0.26	0.43	0.43	0.41	0.38	0.85	0.43
	SST	0.48	0.53	0.49	0.13	0.80	0.79	0.70	0.71	0.74	0.19	1.00	1.00
	FIH	0.36	0.35	0.03	0.11	0.59	0.62	0.44	0.56	0.26	0.23	1.00	0.98
	DCO	0.50	0.47	0.49	0.72	0.57	0.50	0.63	0.58	0.57	0.89	0.80	0.67
	LIK	0.01	0.01	0.00	0.11	0.04	0.04	0.12	0.17	0.22	0.00	0.00	0.21
	UKM	0.14	0.09	0.04	0.09	0.34	0.44	0.11	0.07	0.06	0.10	0.71	0.52
	UKO	0.66	0.48	0.10	0.52	0.17	0.15	0.66	0.48	0.10	0.52	0.17	0.15
	BNL	0.47	0.55	0.40	0.30	0.27	0.54	0.82	0.74	0.65	0.37	0.82	0.66
	GRU	0.37	0.35	0.39	0.35	0.30	0.63	0.44	0.42	0.60	0.51	0.51	0.67
	GMU	0.60	0.54	0.60	0.36	0.17	0.54	0.80	0.74	0.76	0.18	0.10	0.69
	AUV	0.36	0.24	0.26	0.09	0.25	0.69	0.36	0.24	0.26	0.09	0.25	0.69
	FPA	0.69	0.68	0.78	0.45	0.48	0.70	0.52	0.39	0.68	0.81	0.71	0.68
	HUG							0.44	0.43	0.67	0.02	0.25	0.63
	SWL												
	ITU	0.23	0.08	0.39	0.06	0.74	0.81	0.39	0.21	0.44	0.12	0.77	0.73
	IRO	0.22	0.19	0.21	0.16	0.23	0.42	0.03	0.08	0.13	0.10	0.04	0.05
	SPB	0.00	0.00	0.00	0.01	0.00	0.00	0.42	0.26	0.27	0.06	0.97	0.97
	SPC	0.06	0.09	0.10	0.06	0.00	0.28	0.05	0.10	0.09	0.06	0.19	0.39
	GRA							0.19	0.48	0.54	0.13	0.35	0.32
	GRH	0.25	0.13	0.63	0.85	0.63	0.42	0.25	0.13	0.63	0.85	0.63	0.42
	Median	0.36	0.24	0.26	0.14	0.25	0.54	0.43	0.42	0.49	0.18	0.57	0.67

Compositon	Study Area	$PM_{2.5}$ NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse	PM_{10} NO_2	NO_X	$PM_{2.5}$ABS	$PM_{2.5}$	PM_{10}	Coarse
Vanadium	NOS	0.33	0.18	0.28	0.34	0.07	0.07	0.49	0.44	0.44	0.86	0.54	0.44
(V)	SST	0.25	0.16	0.30	0.15	0.44	0.40	0.56	0.61	0.65	0.10	0.77	0.78
	FIH	0.25	0.26	0.01	0.00	0.61	0.71	0.11	0.20	0.04	0.07	0.12	0.11
	DCO	0.60	0.48	0.29	0.37	0.56	0.83	0.59	0.50	0.42	0.41	0.59	0.87
	LIK	0.00	0.00	0.30	0.00	0.02	0.46	0.03	0.03	0.03	0.04	0.02	0.11
	UKM	0.80	0.91	0.72	0.12	0.01	0.09	0.05	0.02	0.14	0.01	0.17	0.09
	UKO	0.66	0.48	0.10	0.52	0.17	0.15	0.69	0.49	0.03	0.54	0.12	0.13
	BNL	0.44	0.35	0.25	0.13	0.26	0.58	0.44	0.35	0.25	0.13	0.26	0.58
	GRU	0.76	0.77	0.68	0.25	0.19	0.29	0.00	0.00	0.06	0.13	0.07	0.01
	GMU							0.00	0.01	0.00	0.01	0.00	0.00
	AUV							0.00	0.01	0.02	0.01	0.02	0.02
	FPA	0.22	0.15	0.13	0.01	0.01	0.07	0.40	0.27	0.24	0.19	0.20	0.44
	HUG							0.26	0.22	0.43	0.08	0.06	0.22
	SWL												
	ITU	0.09	0.06	0.07	0.02	0.01	0.04	0.78	0.76	0.70	0.57	0.51	0.31
	IRO	0.20	0.13	0.06	0.10	0.03	0.06	0.43	0.42	0.43	0.51	0.40	0.46
	SPB	0.27	0.19	0.15	0.01	0.49	0.44	0.52	0.34	0.39	0.11	0.92	0.92
	SPC	0.42	0.25	0.28	0.30	0.62	0.31	0.36	0.19	0.20	0.28	0.69	0.30
	GRA	0.01	0.02	0.01	0.07	0.28	0.00	0.01	0.03	0.01	0.02	0.21	0.00
	GRH	0.03	0.04	0.02	0.18	0.12	0.05	0.02	0.03	0.04	0.21	0.14	0.07
	Median	0.27	0.19	0.25	0.13	0.23	0.29	0.36	0.22	0.20	0.13	0.21	0.30
Zinc	NOS	0.49	0.50	0.39	0.89	0.49	0.52	0.81	0.74	0.67	0.32	0.73	0.32
(Zn)	SST	0.82	0.87	0.75	0.25	0.74	0.75	0.75	0.81	0.77	0.22	0.70	0.72
	FIH	0.02	0.06	0.03	0.28	0.00	0.00	0.09	0.11	0.10	0.67	0.06	0.06
	DCO	0.24	0.20	0.19	0.53	0.30	0.19	0.77	0.68	0.57	0.56	0.71	0.75
	LIK							0.17	0.08	0.54	0.21	0.41	0.45
	UKM	0.51	0.52	0.60	0.29	0.30	0.25	0.45	0.59	0.62	0.33	0.05	0.05
	UKO	0.86	0.72	0.11	0.82	0.33	0.39	0.66	0.66	0.51	0.83	0.76	0.68
	BNL	0.02	0.03	0.08	0.18	0.00	0.01	0.13	0.15	0.22	0.29	0.07	0.01
	GRU	0.68	0.67	0.66	0.28	0.22	0.28	0.71	0.70	0.57	0.33	0.28	0.62
	GMU	0.13	0.18	0.18	0.54	0.42	0.04	0.75	0.76	0.77	0.22	0.30	0.74
	AUV	0.13	0.05	0.00	0.05	0.13	0.22	0.51	0.70	0.64	0.46	0.87	0.51
	FPA	0.58	0.41	0.67	0.69	0.81	0.84	0.67	0.58	0.73	0.72	0.86	0.90
	HUG												
	SWL	0.07	0.01	0.13	0.00	0.16	0.45	0.76	0.51	0.52	0.30	0.82	0.76
	ITU	0.66	0.66	0.53	0.39	0.39	0.25	0.68	0.52	0.76	0.39	0.82	0.73
	IRO	0.34	0.34	0.66	0.53	0.42	0.57	0.35	0.39	0.87	0.68	0.59	0.72
	SPB	0.01	0.02	0.02	0.01	0.00	0.18	0.10	0.07	0.11	0.01	0.40	0.47
	SPC	0.46	0.30	0.29	0.25	0.39	0.38	0.60	0.46	0.48	0.47	0.39	0.52
	GRA	0.04	0.08	0.03	0.17	0.36	0.01	0.11	0.31	0.31	0.22	0.47	0.02
	GRH	0.06	0.04	0.65	0.35	0.85	0.98	0.68	0.57	0.11	0.36	0.70	0.85
	Median	0.46	0.30	0.29	0.29	0.36	0.28	0.66	0.57	0.57	0.33	0.65	0.62

Table S6. Average model R^2 and LOOCV R^2 for each study area by fraction

Study area	PM_{10}			$PM_{2.5}$		
	Number of models	Average of model R^2	Average of LOOCV R^2	Number of models	Average of model R^2	Average of LOOCV R^2
NOS	8	0.83	0.77	8	0.63	0.54
SST	8	0.76	0.70	7	0.56	0.47
FIH	8	0.46	0.34	8	0.46	0.38
DCO	8	0.66	0.61	8	0.49	0.42
LIK	6	0.39	0.29	4	0.24	0.15
UKM	8	0.58	0.49	7	0.56	0.49
UKO	8	0.61	0.56	8	0.46	0.38
BNL	8	0.64	0.56	8	0.59	0.54
GRU	8	0.61	0.54	7	0.62	0.52
GMU	8	0.62	0.54	7	0.71	0.58
AUV	8	0.58	0.50	7	0.52	0.43
FPA	8	0.68	0.59	8	0.56	0.45
HUG	6	0.37	0.26	4	0.38	0.29
SWL	4	0.87	0.81	4	0.84	0.80
ITU	8	0.76	0.72	8	0.61	0.53
IRO	8	0.75	0.70	8	0.53	0.46
SPB	8	0.77	0.70	8	0.68	0.59
SPC	8	0.76	0.71	8	0.66	0.60
GRA	8	0.65	0.54	7	0.74	0.67
GRH	6	0.40	0.29	7	0.32	0.22
Grand Total	150	0.64	0.56	141	0.56	0.48

Chapter 6

Long-Term Exposure to Elemental Constituents of Particulate Matter and Cardiovascular Mortality in 19 European Cohorts: Results from the ESCAPE and TRANSPHORM Projects

Meng Wang, Rob Beelen, Massimo Stafoggia, Evangelia Samoli, Zorana Andersen, Mark Nieuwenhuijsen, Paul Fischer, Barbara Hoffmann, Oie Raacshou-Nielsen, Paolo Vineis, Wei Xun, Konstantina Dimakopoulou, Mette Sørensen, Menno Keuken, Bas Bueno-de-Mesquita, Timothy Key, Michiel Bots, Claes-Göran Östensson, Göran Pershagen, Annette Peters, Enrica Migliore, Ulf De Faire, Timo Lanki, Nancy Pedersen, Chiara Badaloni, Joachim Heinrich, Hans Concin, Jaakko Kukkonen, Laura Fratiglioni, Gabriele Nagel, Klea Katsouysanni, Gerard Hoek, Bert Brunekreef

On behalf of the ESCAPE and the TRANSPHORM consortium

(Submitted)

Abstract

Background: Associations between long-term exposure to ambient particulate matter (PM) and cardiovascular (CVD) mortality have been widely recognized. However, health effects of long-term exposure to constituents of PM on total CVD mortality have not been explored.
Aims: The aim of this study is to examine the association of PM composition with cardiovascular mortality.
Methods: We used data from 19 European ongoing cohorts within the framework of the ESCAPE and TRANSPHORM projects. Residential annual average exposure to elemental constituents (Copper, Iron, Potassium, Nickel, Sulfur, Silicon, Vanadium and Zinc) within $PM_{2.5}$ and PM_{10} was estimated using Land Use Regression models. Cohort-specific analyses were conducted using Cox proportional hazards models with a standardized protocol. Random-effects meta-analysis was used to calculate pooled effect estimates.
Results: The total population consisted of 322,291 participants, with 9,545 CVD deaths. We found no associations between any of the elemental constituents in $PM_{2.5}$ or PM_{10} and CVD mortality. Most of the hazard ratios were close to unity, with the exception of PM2.5 Si, and S in $PM_{2.5}$ and PM_{10}.
Conclusion: In a joint analysis of 19 European cohorts, we found no significant association between long-term exposure to 8 elemental constituents of particles and total cardiovascular mortality.

Introduction

Recent studies of health effects of particulate matter (PM) show accumulating evidence of adverse effects on cardiovascular (CVD) mortality [1, 2]. However, effect estimates of long-term exposure to $PM_{2.5}$ and PM_{10} (particles <2.5 μm and 10 μm in aerodynamic diameter) varied among different studies and geographical locations, showing elevated risks in some cities and areas in Europe and the United States [3-6] but no or little association in others [7, 8]. The recent study within the ESCAPE (European Study of Cohorts for Air Pollution Effects) project by Beelen et al. (2013)[9] reported no associations between PM mass concentrations and CVD mortality in 19 ongoing cohorts in Europe while the magnitude of the association of individual cohort differed among study areas, suggesting that more specific examinations are needed taking individual particle components into account rather than including size as only indicator.

Ambient $PM_{2.5}$ and PM_{10} represent a heterogeneous mixture of constituents from diverse sources e.g. fossil fuel combustion, biomass burning and human activity [10]. However, little is known about which PM constituents are associated with higher risks. Several studies showed evidence of acute effects of PM components on CVD mortality, but results varied among studies [11-14]. The reasons for these differences are not clear but may include methodological differences as well as true variations related to PM composition and/or population susceptibility. . Long-term effects of specific elemental components on CVD mortality have only been examined by a single study which suggested that the constituents responsible for the mortality risks were from combustion of fossil fuel, biomass burning and crustal origin [15].

The aim of this study is to investigate effects of long term exposure to PM constituents on all CVD mortality using standardized methods within the ESCAPE (European Study of Cohorts for Air Pollution Effects) and TRANSPHORM (Transport related Air Pollution and Health impacts – Integrated Methodologies for Assessing Particulate Matter) projects using data from 19 ongoing cohort studies.

Methods

The association between PM constituents and cardiovascular mortality was analyzed in each cohort separately using a common statistical protocol for exposure assessment, outcome definition, confounder models and statistical analysis. Cohort-specific results were pooled together and evaluated at the coordinating institute (IRAS, Utrecht University). The individual effect estimates were combined by random-effects meta-analysis [16]. Pooling of the cohort data was not possible due to data transfer and privacy issues. Random-effects meta-analysis had the advantage of including both inter-area and inter-cohort differences which were not entirely addressed by the available data on confounders.

Table 1: Description of the included cohort studies: total study population (N total)[a], number of deaths per outcome (N outcome)[a], mean age at baseline (SD), baseline period, total follow-up time (average) and short description of the study area. Order of studies is based on North to South gradient.

Cohort	Study area description	N total	N CVD[b]	Age (years)	Baseline period	Follow-up time (person years)
FINRISK, Finland	Greater Helsinki Area and Turku city and its rural surroundings	10,224	225	47.9 (13.2)	1992; 1997; 2002; 2007	108,434 (10.6)
HUBRO, Norway	City of Oslo	18,234	332	48.3 (15.2)	2000-2001	175,076 (9.6)
SNAC-K, Sweden	City of Stockholm	2401	140	70.3 (8.1)	2001-2004	15,568 (6.5)
SALT/Twin gene, Sweden	Stockholm County	5473	206	58.0 (9.9)	1998-2002	47,767 (8.7)
60-y/IMPROVE, Sweden	Stockholm County	3612	81	60.4 (0.1)	1997-1999	40,612 (11.2)
SDPP, Sweden	Stockholm County	7408	55	47.1 (5.0)	1992-1998	102,831 (13.9)
DCH, Denmark	City of Copenhagen and surrounding areas	35,458	678	56.7 (4.4)	1993-1997	469,571 (13.2)
EPIC-MORGEN, Netherlands	Cities of Amsterdam, Maastricht and Doetinchem and surrounding rural areas	16,446	185	43.9 (10.9)	1993-1997	217,722 (13.2)
EPIC-PROSPECT, Netherlands	City of Utrecht and surrounding rural areas	15,670	290	57.7 (6.0)	1993-1997	202,809 (12.9)
SALIA, Germany	Areas in the cities of Dortmund, Duisburg, Essen, Gelsenkirchen and Herne situated in the Ruhr Area and adjacent towns Borken and Dülmen	4352	206	54.5 (0.6)	1985-1987; 1990-1994	81,093 (18.6)
EPIC-Oxford, UK	Urban and rural areas within 400 km around London-Oxford area	38,941	661	45.8 (13.7)	1993-2001	491,542 (12.6)
KORA, Germany	City of Augsburg and two adjacent rural counties	8399	270	49.5 (13.8)	1994-1995; 1999-2001	88,592 (10.5)
VHM&PP, Austria	State of Vorarlberg, excluding high mountain areas (> 600m) and areas within 300m of state border	117,824	5,858	41.9 (14.9)	1985-2005	2,039,328 (17.3)
SAPALDIA, Switzerland	Cities of Geneva, Lugano and Basel	1250	14	41.1 (11.8)	1991	20,294 (16.2)
E3N, France	Cities of Paris, Grenoble, Lyon and Marseille and surrounding rural areas	10,915	89	53.0 (6.7)	1993-1996	147,021 (13.5)
EPIC-Turin, Italy	City of Turin	7261	46	50.4 (7.5)	1993-1998	97,549 (13.4)
SIDRIA-Turin, Italy	City of Turin	5054	29	44.2 (6.2)	1999	55,667 (11.0)
SIDRIA-Rome, Italy	City of Rome	9177	64	44.3 (6.0)	1999	102,856 (11.2)
EPIC-Athens, Greece	Greater Athens Area	4192	116	49.4 (11.7)	1994-1999	46,852 (11.2)

[a] Number of observations without missing value in any confounder variable of model 3 (main model)
[b] CVD = all cardiovascular mortality

Study populations

Our analysis included 19 cohorts from 12 countries where PM measurements were available: Finland (FINRISK), Norway (HUBRO), Sweden (SNAC-K, SALT, Sixty, SDPP), Denmark (DCH), the Netherlands (EPIC-MORGEN, EPIC-PROSPECT), Germany (SALIA, KORA), the United Kingdom (EPIC-Oxford), Austria (VHMPP), Switzerland (SAPALDIA), France (E3N), Italy (EPIC-Turin, SIDRIA-Turin, EPIC-Rome), and Greece (EPIC-Athens) (Table 1). Detailed information of the study design and the characteristics of each cohort have been described in online supplement 1. The study areas of most cohorts consisted of a large city and its surrounding area. Some of the cohorts included large regions of the country such as EPIC-MORGEN in the Netherlands, EPIC-Oxford covering much of the UK, the VHM&PP cohort in Vorarlberg, Austria, and SAPALDIA in three cities of Switzerland. All included cohort studies were approved by the institutional medical Ethics Committees and conducted in accordance with the Declaration of Helsinki. All subjects provided written informed consent.

Definition of cardiovascular mortality outcome

In all cohorts, follow-up was based upon linkage to mortality registries. All CVD outcome mortality was defined on the basis of the underlying cause of death recorded on death certificates as ICD-9: 400-440; ICD-10: I10-I70.

Exposure assessment

Within ESCAPE, we a priori selected 8 elements (Copper (Cu), iron (Fe), potassium (K), nickel (Ni), sulfur (S), silicon (Si), vanadium (V) and zinc (Zn)). These elements were included because they reflected a variety of particle sources, had evidences for toxicity and had a high percentage (>75%) of well-detected samples. Elemental concentrations at the baseline residential addresses of study participants were estimated by Land Use Regression models (LUR) following a standardized procedure described previously [17]. Briefly, three two-week measurements of PM were conducted during different seasons between October 2008 and May 2011 in each cohort study area (1 year per study area). Annual average concentrations of PM constituents were obtained by adjusting temporal variation measured at a continuous background sampling site in the entire period [18]. PM filters with aerodynamic diameter <2.5μm ($PM_{2.5}$) and <10μm (PM_{10}) were weighed before and after each measurement centrally at IRAS, Utrecht University and were then sent to the Cooper Environmental Services (Portland, OR, USA) to detect metal components. All filters were analyzed for elemental compositions using X-Ray Fluorescence (XRF). Details of the measurement, analysis and results have been published elsewhere [17]. Area-specific LUR models were developed for each element to explain annual concentrations using traffic and land use predictor variables in a GIS database. Models explained modest to large fractions of variation of PM constituents (see table S2-17 in online supplement 2). However, 16 of 224 models of the 12 countries (14 study areas) could not be built as no single predictor variable met the predefined inclusion criteria [17] (see table S1 in online supplement 2). To avoid extreme values of PM constituents at cohort addresses, we truncated the values of the predictor variables to the upper or lower limit of

Table 2: Population characteristics at baseline of the included cohort studies. Order of studies is North to South. Detailed description of each cohort can be found in online supplement 3.

Cohort	% women	% never smokers	Cigarettes / day[a]	Years of smoking[a]	BMI[a] (kg/m^2)	Fruit intake[b]	Alcohol intake[c]	% married / living with partner	% low educational level	% (self)employed
FINRISK, Finland	53.8%	45.4%	3.8 (7.8)	8.6 (12.2)	26.4 (4.6)	66.3%	0.9 (1.3)	70.1%	31.0%	69.2%
HUBRO, Norway	56.1%	46.0%	6.7 (8.4)	11.5 (14.4)	25.7 (4.1)	39.9%	50.6%	49.8%	17.7%	NA
SNAC-K, Sweden	60.0%	44.3%	7.1 (9.5)	9.8 (15.2)	26.0 (4.1)	NA	21.8%	54.2%	21.2%	28.6%
SALT/Twin gene, Sweden	55.7%	38.7%	8.5 (9.7)	16.7 (17.3)	28.6 (4.1)	NA	NA	68.0%	21.5%	NA
60-y/IMPROVE, Sweden	52.5%	41.0%	8.0 (9.1)	15.2 (16.4)	26.8 (4.2)	64.2%	8.9 (9.7)	71.6%	27.5%	51.4%
SDPP, Sweden	61.7%	37.4%	8.5 (8.8)	12.3 (12.4)	25.6 (4.0)	92.4%	1.3 (1.9)	83.6%	25.5%	91.8%
DCH, Denmark	54.1%	36.3%	6.3 (10.4)	18.7 (17.1)	26.0 (4.1)	183.2 (151.2)	21.7 (22.8)	69.2%	29.6%	80.1%
EPIC-MORGEN, Netherlands	54.4%	35.0%	10.4 (11.1)	14.3 (13.7)	25.2 (4.0)	171.9 (129.2)	12.7 (18.0)	67.7%	11.9%	NA
EPIC-PROSPECT, Netherlands	100%	45.0%	5.7 (7.4)	15.2 (16.5)	25.5 (4.1)	231.6 (139.2)	9.0 (12.4)	76.9%	22.2%	NA
SALIA, Germany	100%	74.5%	2.6 (6.6)	4.4 (10.5)	NA	NA	NA	NA	28.8%	NA
EPIC-Oxford, UK	77.5%	63.3%	5.0 (8.3)	6.7 (11.2)	24.0 (3.9)	259.9 (204.5)	9.1 (11.7)	70.8%	36.5%	72.5%
KORA, Germany	50.8%	43.7%	9.2 (13.3)	12.0 (14.2)	27.2 (4.6)	59.5%	16.3 (22.3)	75.7%	12.6%	58.3%
VHM&PP, Austria	56.1%	69.9%	NA	NA	24.8 (4.3)	NA	NA	68.4%	NA	69.3%
SAPALDIA, Switzerland	52.0%	42.4%	11.5 (14.5)	10.7 (12.4)	23.7 (4.0)	NA	NA	54.3%	15.0%	84.4%
E3N, France	100%	49.3%	NA	NA	22.8 (3.2)	242.0 (164.7)	12.0 (15.1)	NA	5.0%	NA
EPIC-Turin, Italy	47.7%	42.6%	7.2 (8·2)	17.6 (16.3)	25.3 (3.8)	318.2 (182.2)	18.1 (20.3)	85.6%	43.6%	NA
SIDRIA-Turin, Italy	51.8%	37.5%	9.3 (10·2)	11.3 (10.6)	NA	NA	NA	95.4%	17.5%	72.2%
SIDRIA-Rome, Italy	52.8%	34.6%	10.1 (10·5)	11.7 (10.4)	NA	NA	NA	100%	44.9%	NA
EPIC-Athens, Greece	55.0%	39.5%	1.7 (15·0)	10.8 (13.1)	27.5 (4.5)	402.6 (258.2)	9.2 (14.5)	78.0%	23.6%	66.9%

[a] Mean (SD)

[b] Mean (SD) (g/day) or percentage with daily fruit consumption. For SDPP it is percentage daily/weekly fruit consumption.

[c] Mean (SD) (g/day) or percentage with daily alcohol consumption. For FINRISK it is number of glasses of alcoholic drink during last week. For SDPP it number of glasses of alcoholic drink per day. For HUBRO it is percentage with weekly alcohol consumption.

NA is not available

those at the measurement sites.This procedure was previously applied for nitrogen oxides and mass particulate matter[19, 20] and has been shown to result in more realistic exposure estimates for these pollutants[21].

Statistical analyses

Cohort specific analyses

Cox proportional hazards models were used for the cohort specific analyses, with age as the underlying time scale. Censoring occurred at the time of death for non-CVD causes, out migration, loss to follow-up for other reasons, or at end of follow-up, whichever came first. Air pollution exposure was analyzed as a linear variable. Potential confounders were available from questionnaires at baseline. A priori we specified three confounder models with increasing level of adjustment. Confounder models were decided based upon previous cohort studies of air pollution and mortality and availability of data in a majority of the cohorts. Model 1 included age (time axis), gender, and calendar time (year(s) of enrollment) only. Model 2 adjusted for additional individual level variables: smoking status (never/former/present), smoking intensity, smoking duration, environmental tobacco smoke, fruit intake, vegetables intake, alcohol consumption (linear and squared term), body mass index (BMI) (linear and squared term), educational level (low, medium, high), occupational class (white/blue collar classification), employment status, and marital status. Model 3 further adjusted for area-level socio-economic status (SES) variables (mostly mean income of neighborhood or municipality). Model 3 was selected as the main confounder model. Only subjects with complete information for Model 3 variables were included in the analyses (see table S1-19 in online supplement 3).

As constituents from similar sources may be highly correlated with each other, we conducted two-pollutant models for all elements, adjusting for all the standard pollutants (NO_2, NOx, $PM_{2.5}$, PM_{10}, $PM_{2.5}$ absorbance and coarse particles) and elements separately. We restricted our two-pollutant models to the cohorts with Pearson correlations between LUR based estimates of the two pollutants lower than 0.7 to avoid multicollinearity [22].

In sensitivity analyses, we added to Model 3 prevalent hypertension and physical activity, and as further classical cardiovascular risk factors: prevalent diabetes and cholesterol level. Extended confounder models were used in sensitivity analyses because some of the air pollution effect might be mediated by these factors.

Meta-analysis

Meta-analyses of cohort-specific estimates were conducted using the Dersimonian-Laird method with random effects. Hazard ratios (HR) and 95% confidence intervals (CIs) were calculated for fixed increments. The increments for each constituent were selected to cover the range in concentrations over most study areas. Heterogeneity between cohorts was quantified by the I^2 statistic and tested by the X^2 test from Cochran's Q statistic[23].

Effect modification was tested by pooling estimates from different regions of ESCAPE (north: Sweden, Norway, Finland, Denmark; west and middle: the United

Kingdom, the Netherlands, Germany, France, Austria, Switzerland; south: Italy and Greece).

In addition, we estimated pooled effects restricted to cohorts with leave-one-out cross validation (LOOCV) R^2 lower and higher than 0.5 for each element.

All cohort-specific analyses and meta-analyses were done in STATA, version 10-12 (StataCorp, College Station, TX, USA). We defined statistical significance as p value <0.05 in the Cox and meta-regression models.

Results

Characteristics of the study population

The entire study population consisted of 322,291 participants contributing 4,551,184 person year at risk (average time of follow-up 12.4 years), with 9,545 CVD deaths during follow-up (Table 1). Cohorts were recruited mostly in the 1990s. Cohorts differed in the number of participants, the mean baseline age, and availability of data on confounders (Table 2 and table S1-19 in online supplement 3). Age gender, smoking status, and area-level SES were available for all cohorts. Smoking intensity and duration were available as continuous variables for all cohorts, except VHM&PP and E3N. VHM&PP had data on occupation and employment status, but not on education.

Air pollution exposure

Concentration distributions of estimated particle constituents varied substantially between and within study areas (see figure S1-2 in online supplement 4). Cu, Fe and Zn (in $PM_{2.5}$ and PM_{10}) showed highest overall concentrations in southern Europe and high spatial contrasts in all study areas. Highest concentrations and contrasts of crustal elements (K and Si) were observed for cohorts in north and south Europe. Exposure contrasts of V and especially S were much larger between than within study areas. Correlations between estimated elemental concentrations and total mass concentrations in the same size fractions ($PM_{2.5}$ or PM_{10}) were modest on average ($0.5<r<0.7$) for traffic tracers (e.g. Cu, Fe and Zn) and for elements which made up a relatively large fraction of PM (e.g. K, S, Si) and were lower ($0.2<r<0.5$) for V and Ni (see table S1 in online supplement 5). Correlation ranges were wide for all the elements across study areas.

Main results

We found no consistent associations between any of the elemental constituents of $PM_{2.5}$ or PM_{10} and CVD mortality in the main model based on pooled analyses of 19 cohorts (model 3 in Table 3), even though some cohort-specific associations were found for some elements (Figure 1). Most of the pooled HRs were close to unity for PM constituents, with exception of S and $PM_{2.5}$ Si. HR for $PM_{2.5}$ Si was 1.17 (95% CI: 0.93-1.47) per 100ng/m^3 and for S in $PM_{2.5}$ and PM_{10} the HRs were 1.08 (95% CI: 0.95-1.22) and 1.09 (95% CI: 0.90-1.32) per 200ng/m^3 respectively. The crude adjustment models (model 1 with adjustment only for calendar year and gender in Table 3) showed relatively high HRs and wide confidence intervals. The HRs were reduced to unity with shrinking confidence interval in each stage after

adjustment for individual level confounders (model 2 in Table 3) and area-level socio-economic status variables (model 3 in Table 3). Nevertheless, positive effects of some elements were observed in a few cohorts on CVD mortality individually such as the Dutch EPIC- PROSPECT cohort ($PM_{2.5}$ Si, Zn and PM_{10} S), and the German SALIA ($PM_{2.5}$ Si and PM_{10} Si, K, Zn) and KORA ($PM_{2.5}$ Fe, Si, Zn and PM_{10} Cu, Fe, Si, S) cohorts.

Heterogeneity between the results of individual cohorts varied substantially across constituents (I^2: 0% ~ 58%, significant for Cu, Fe, S, Si and Zn), and was generally larger for the elemental constituents than for PM mass (Table 3). HRs of constituents-CVD meta-analyses showed similar results for both using random effects (default method) and fixed effects (Figure 1).

Table 3 Association[a] between CVD mortality and exposure to PM constituents: Results from random-effects meta-analyses (HRs and 95% CIs), p-value of model 3 and I^2 (p_{het}) of test for heterogeneity of effect estimates between cohorts

N[b]	Exposure		Model 1[c]	Model 2[c]	Model 3[c]	Pm[d]	P_{het}	I^2
19	Cu	$PM_{2.5}$	1.03(0.88-1.21)	0.96(0.83-1.11)	0.90(0.77-1.07)	0.26	0.03	42.77
19		PM_{10}	1.00(0.89-1.13)	0.95(0.84-1.08)	0.93(0.82-1.06)	0.29	0.01	48.83
19	Fe	$PM_{2.5}$	1.07(0.94-1.23)	1.01(0.89-1.15)	0.99(0.87-1.11)	0.82	0.09	31.93
19		PM_{10}	1.03(0.91-1.16)	0.98(0.86-1.12)	0.96(0.84-1.09)	0.55	0.00	53.49
18	K	$PM_{2.5}$	0.98(0.94-1.02)	0.98(0.94-1.02)	0.98(0.94-1.02)	0.37	0.62	0.00
18		PM_{10}	1.02(0.94-1.11)	1.01(0.94-1.07)	1.00(0.93-1.08)	0.92	0.27	15.03
14	Ni	$PM_{2.5}$	1.05(0.86-1.28)	0.98(0.78-1.24)	0.97(0.78-1.21)	0.80	0.02	48.63
17		PM_{10}	1.12(0.93-1.36)	1.02(0.87-1.19)	1.01(0.88-1.16)	0.90	0.35	9.11
18	S	$PM_{2.5}$	1.27(1.01-1.61)	1.09(0.97-1.24)	1.08(0.95-1.22)	0.25	0.68	0.00
18		PM_{10}	1.21(0.97-1.50)	1.11(0.92-1.33)	1.09(0.90-1.32)	0.40	0.06	36.46
16	Si	$PM_{2.5}$	1.28(0.98-1.66)	1.21(0.94-1.54)	1.17(0.93-1.47)	0.21	0.01	52.04
18		PM_{10}	1.09(0.93-1.28)	1.03(0.91-1.17)	1.01(0.90-1.13)	0.85	0.08	33.87
15	V	$PM_{2.5}$	1.14(0.86-1.52)	1.00(0.79-1.27)	1.00(0.80-1.24)	0.99	0.19	24.10
18		PM_{10}	1.04(0.83-1.30)	1.01(0.86-1.18)	1.00(0.85-1.17)	0.99	0.32	11.10
19	Zn	$PM_{2.5}$	1.08(0.90-1.29)	1.06(0.90-1.25)	1.04(0.88-1.24)	0.63	0.00	58.59
19		PM_{10}	1.06(0.91-1.24)	1.03(0.90-1.19)	1.00(0.86-1.16)	0.99	0.00	57.47
19		$PM_{2.5}$[e]	1.18(1.00-1.38)	1.04(0.93-1.17)	0.99(0.91-1.08)	0.80	0.80	0.00
19		PM_{10}[e]	1.10(0.95-1.26)	1.04(0.92-1.16)	1.02(0.91-1.14)	0.69	0.69	20.20

[a]Effects are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn

[b]N: number of cohorts in the meta-analysis

[c]Model 1: adjusted for gender and calendar time; Model 2: as in Model 1 also adjusting for smoking status, smoking intensity, smoking duration, environmental tobacco smoke, fruit intake, vegetables intake, alcohol consumption, body mass index, education level, occupational class, employment status, marital status; and Model 3: as in model 2 also adjusting for area-level socio-economic status

[d]Pm: p-value of HR from meta regression in model 3; [e]Association between CVD mortality and $PM_{2.5}$, PM_{10} mass concentrations reported by Beelen et al. (Submitted)[9].

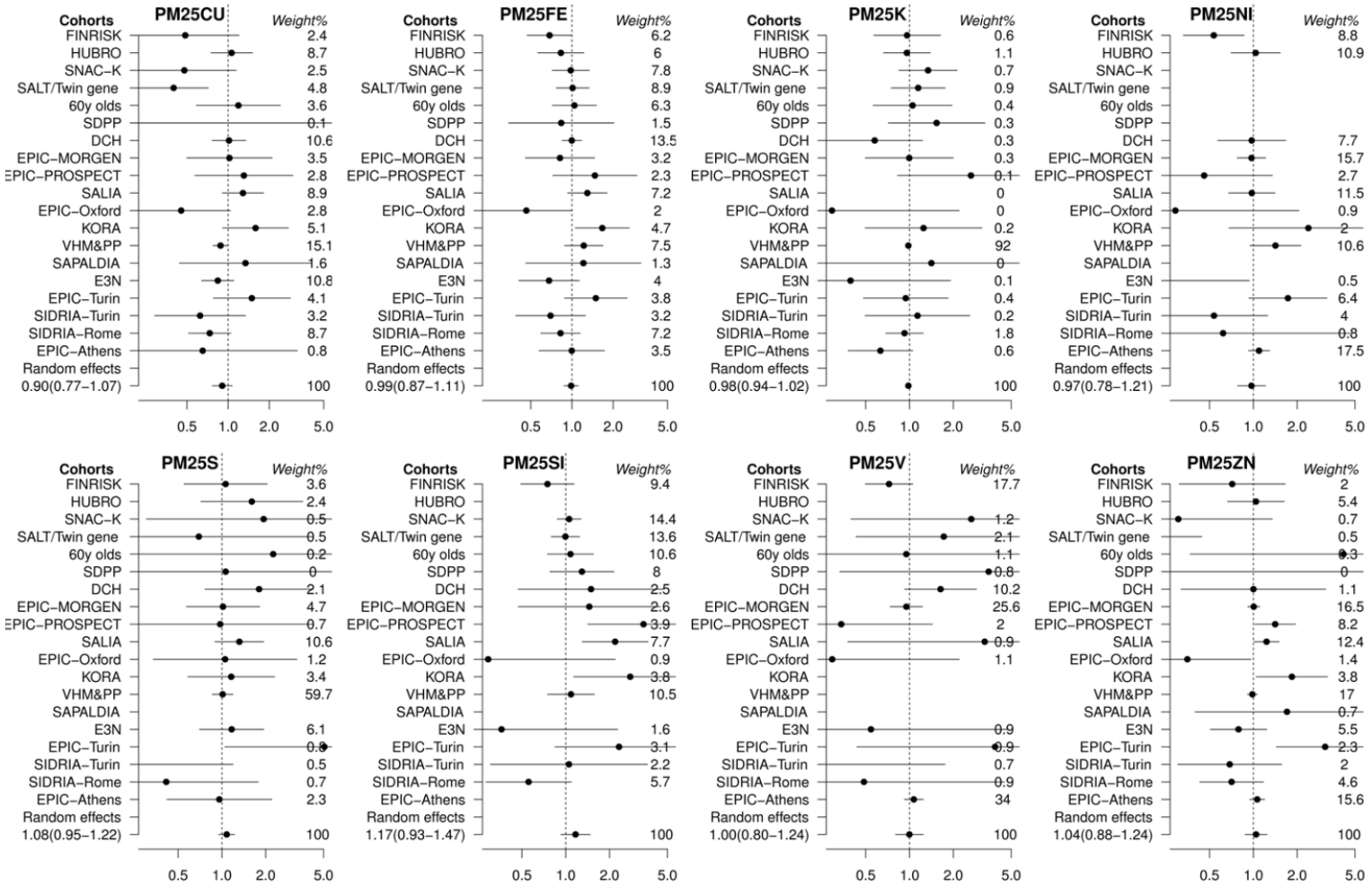


Fig.1 Adjusted association between CVD mortality and exposure to $PM_{2.5}$ constituents: Results from cohort-specific analyses and from random effect-effects meta-analyses[a]

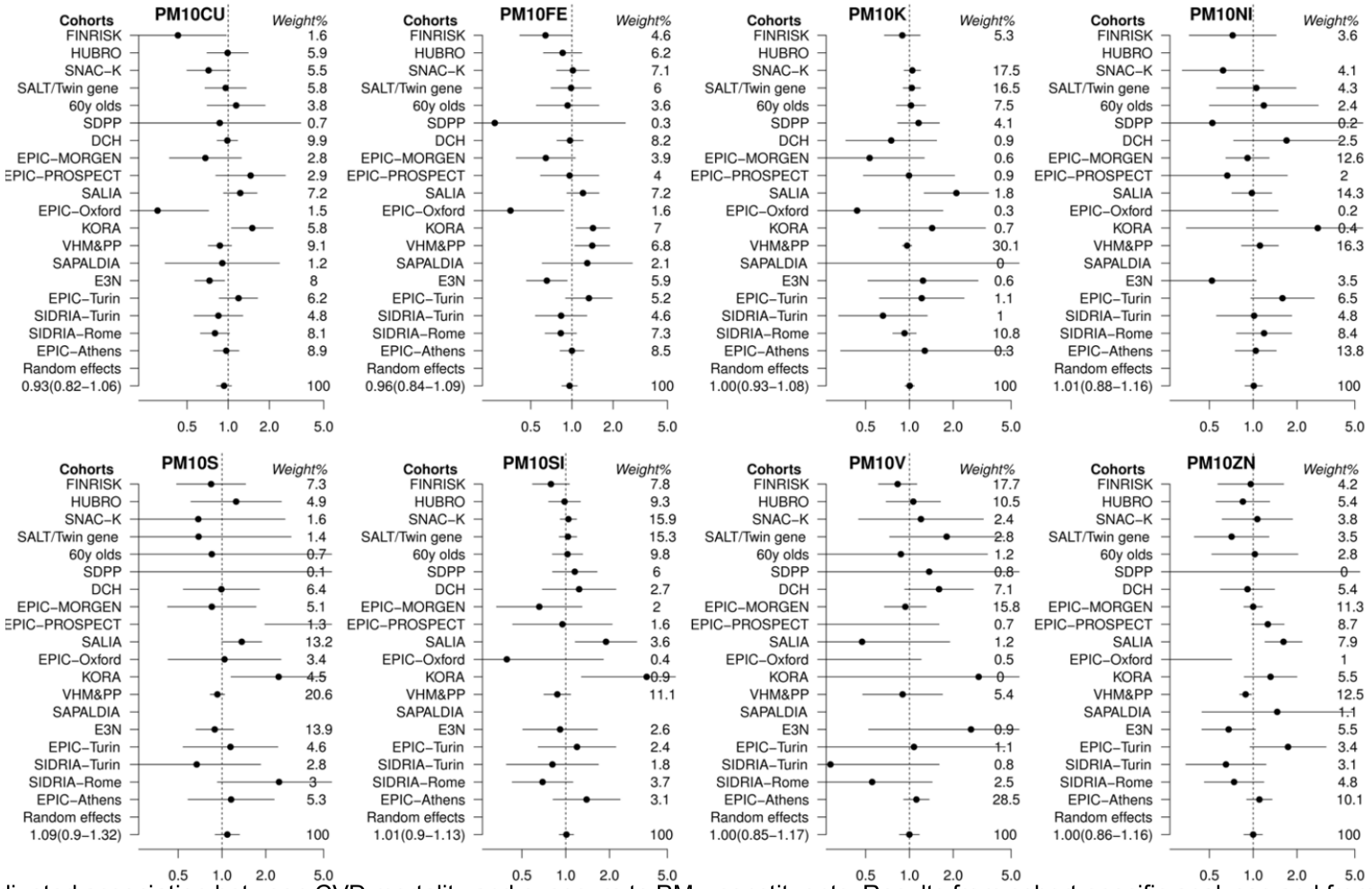


Fig.1 Adjusted association between CVD mortality and exposure to PM_{10} constituents: Results from cohort-specific analyses and from random effect-effects meta-analyses[a] *(Continued)*.

[a]HRs are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn.

Sensitivity analyses

Additional adjustments for hypertension and physical activity, diabetes and cholesterol, and noise did not change the main results of the pooled analyses (table S1-4 in online supplement 6). Using two-pollutant models adjusting for NO NO_2, NOx or PM metrics respectively, while restricting analyses to cohorts with correlations of the two pollutants lower than 0.7 did not change the HRs compared to the single pollutant models in the same cohorts. With a few exceptions, the HR of S was significantly elevated after adjusting for Cu in the same fraction of PM ($PM_{2.5}$ 1.18, 95%CI: 1.02-1.36; PM_{10} 1.35, 95%CI: 1.06-1.72) per 200 ng/m^3 (Figure S1-2 in online supplement 6), but in view of the many analyses performed, this could be a chance finding. Restricting the meta-analyses to cohorts for which the exposure model had LOOCV R^2's of at least 0.5 resulted in removal of many cohorts for estimates of K, Ni, S and V in $PM_{2.5}$. No heterogeneity was observed between stratified cohorts with LOOCV R^2 higher and lower than 0.5. No association was found after this restriction (Table S5 in online supplement 6). Restricting analyses to cohorts with port and industrial sources for trace markers of PM_{10} V and Ni, we found slightly elevated but not significant associations with CVD mortality (V 1.04, 95%CI: 0.90-1.21; Ni 1.07, 95%CI: 0.80-1.42 per 3ng/m^3). Stratified analyses showed slight differences of HRs among regions but generally no association was identified for any of the elemental constituents on CVD mortality (Table S6 in online supplement 6). Stratified analyses showed elevated effect of $PM_{2.5}$ S on CVD mortality for cohorts with average ages older than 50 years (1.35 95%CI: 1.03-1.77 per 200ng/m^3) with respect to that with ages younger than 50 (1.01 95%CI: 0.88-1.17).

Discussion

This study found no significant association between a comprehensive set of elemental constituents of PM and overall cardiovascular mortality based on 19 European cohorts. Most of the hazard ratios were close to unity, with exception for $PM_{2.5}$ Si and S in $PM_{2.5}$ and PM_{10}.

The strengths of this study include: 1) we recruited a large population of 19 study cohorts from 12 countries with a relatively long follow-up history and with detailed personal information; 2) we incorporated a comprehensive set of potential confounders including smoking status and intensity, area-level SES and other important variables, which are at least as intense as those in previous studies e.g. the large U.S. ACS study [24]; 3) we implemented standardized land use regression models for many of the measured PM elemental constituents for fine scale predictions [17, 25]; few previous European and American studies assessed PM constituents by modeling but assigned individual exposures from the nearest monitoring stations.

Our study is an extension study of Beelen et al. (2013)[9] who reported no significant associations between any of the pollutants NO_2, NOx, $PM_{2.5}$, PM_{10} and $PM_{2.5}$ absorbance and a number of CVD mortality outcomes based on the ESCAPE study cohorts. Both studies recruited the same number of cohorts across Europe

for PM and applied similar standardized protocols in exposure assessment, confounder selections and statistical analyses. The median correlations between elemental constituents and PM mass concentrations were from low to modest across study areas, suggesting that mortality effects from elemental constituents could be different from the particle mass mortality effects.

However, we did not identify any significant associations between CVD mortality and PM elemental constituents across the 19 cohorts. Most of the hazard ratios were close to unity for PM constituents and were in line with the findings from $PM_{2.5}$ and PM_{10} [9]. This is different from the relatively high correlations between $PM_{2.5}$ and constituents (0.67[K]<r<0.84[Zn]) in the California Teachers Study (CTS) of which the HRs of all the constituents exhibited similar elevated risks of mortality outcomes to the effect estimates of $PM_{2.5}$ [15].

To the best of our knowledge, associations between long-term exposure to elemental constituents and total CVD mortality have not been studied. Only one study reported long-term effects of several PM constituents on ischemic heart disease (IHD) mortality in the CTS study [15]. In contrast to the CTS study [15], we pay particular attention to within-area exposures of constituents to individual cohort members as substantial spatial variability has been identified within each study areas for most of the elements (Tsai et al. in preparation). Exposure error may occur if within-city variability was not taken into account [26]. In addition, the CTS cohort has relatively larger exposure contrasts (between counties) to the elements (Fe, K, Si and Zn) and older population on average which may increase susceptibility to air pollution exposure compared to the cohorts in our study.

Although elemental effects on cardiovascular disease have been examined in a few short-term studies, results varied substantially across areas. In countrywide studies, effects estimates were significantly increased only for V and Ni in the United States [27, 28]. In our study, V and Ni were mainly explained by a combination of industry, port, residential density and traffic variables, in agreement with previous European source apportionment studies which suggested fuel (and its derivates e.g. shipping emission) /residual oil combustion and industrial emissions [29]. We restricted analyses to the cohorts with port and industry variables for PM_{10} V and Ni respectively (DCH, EPIC-MORGEN, SNAC-K, SALT, Sixty, SDPP, EPIC-Athens for both V and Ni and SALIA for V, SIDRIA/EPIC-Turin for Ni). We observed stronger associations though not statistically significant. This is in agreement with the finding of a concurrent study which showed negative associations between Ni or V and lung function in cohorts with elemental concentrations influenced by distance to port in the LUR models [30]. The variations in particular of Cu, Fe and Zn were explained by traffic related variables (e.g. traffic intensity and road length) in many cohort areas. Cu, Fe and Zn are recognized as trace markers of traffic from a combination of brake/tire-wear, vehicle exhausts and road dusts [29]. Some large scale, short-term studies in the U.S showed no association between CVD outcomes and Cu, Fe and Zn [27, 28] in line with our study results whereas four other studies found associations with Cu, Fe or Zn in smaller regions [11, 12, 31, 32]. We found stronger associations between CVD mortality and $PM_{2.5}$ Si, S in $PM_{2.5}$ and PM_{10} than other constituents though not statistically significant. Mineral and road dust are the primary sources of Si [29]. The cardio-toxicity effects of $PM_{2.5}$ Si have been

presented in some studies [11-13] but not in others [27, 28, 31, 32]. S is part of vehicle exhausts but is mostly determined by secondary aerosol formation. In this study, the spatial variation of S was mostly explained by traffic and residential density variables[17]) and thus S represents traffic emissions and possibly residential oil combustion. We found significant associations for $PM_{2.5}$ S in the cohorts older than 50 year on average with respect to that in the subset of younger ages. This finding was consistent with the results of a previous study that elder people were more vulnerable to $PM_{2.5}$ than younger people [33].

$PM_{2.5}$ and PM_{10} are often associated with both the constituent concentrations and the health outcome, making it difficult to disentangle element contributions from those of PM mass[34]. Mostofsky et al. (2012)[34] evaluated several solution strategies, including modeling with constituent proportions, PM species as modifier or residual analyses which yielded fairly similar effect estimates across models. Following the strategy proposed by Mostofsky et al. (2012)[34], we conducted two-pollutant model by adjusting individual pollutants (including PM metrics and nitrogen oxides). We further restricted the analyses to the cohorts with correlations between the two pollutants lower than 0.7 as regression model may become highly unstable when incorporating two pollutants that are highly correlated[22]. The null effects of constituents remained after adjusting for NO_2/NOx and PM metrics, except for S in $PM_{2.5}$ and PM_{10} adjusted for Cu in the same size fractions. The increased effects of S in $PM_{2.5}$ and PM_{10} could be explained by chance finding given the large number of two-pollutant models for one component. Consistently, the Cu effects decreased after the adjustment of S in the same size fraction (Figure S1-2 in online supplement 6).

We observed higher heterogeneity in pooled HRs for some constituents than for $PM_{2.5}$ and PM_{10} in our previous study[9]. This is partially due to the differences in particle compositions across study areas (correlations between PM mass and constituents varied substantially between study areas)[17]

A limitation in our study is that the LUR models used for exposure assessment were based on air pollution measurements in the period 2008-2011 while cohort studies included in ESCAPE started in the past (1985-2007 with most studies starting in the mid-90s). Three recent studies in the Netherlands[35], Great Britain[36], Rome [37], and Vancouver[38] have shown that for periods up to 10 years spatial air pollution contrasts of NO_2 often remained the same. This indicates that the LUR models based on current NO_2 data are able to predict historical exposure well. This finding may be applicable to traffic-related constituents such as Cu, Fe and Zn whereas is still unclear for the other constituents. In addition, the LUR models explained fairly large concentration contrasts of constituents in most of the study areas but had poor predictions for several $PM_{2.5}$ constituents (e.g. K, Ni, S and V) in several study areas. This is due to lack of proper variables for specific sources, for instance biomass burning for K and regional transport for S. We restricted the analyses to the cohorts with reliable estimates and the effects of these constituents on CVD mortality did not change from the main results.

In summary, we did not find significant associations between any of the elemental constituents of particles and overall cardiovascular mortality in the 19 European cohort studies. Elevated risks were found for CVD mortality and $PM_{2.5}$ Si,

S in $PM_{2.5}$ and PM_{10} respectively, although not significant.

References

1. Brook, R. D.; Rajagopalan, S.; Pope, C. A., 3rd; Brook, J. R.; Bhatnagar, A.; Diez-Roux, A. V.; et al., Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **2010,** *121*, (21), 2331-78.
2. Hoek, G.; Krishnan, R. M.; Beelen, R.; Peters, A.; Ostro, B.; Brunekreef, B.; et al., Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environ Health* **2013,** *12*, (1), 43.
3. Laden, F.; Schwartz, J.; Speizer, F. E.; Dockery, D. W., Reduction in fine particulate air pollution and mortality - Extended follow-up of the Harvard six cities study. *American Journal of Respiratory and Critical Care Medicine* **2006,** *173*, (6), 667-672.
4. Lipsett, M. J.; Ostro, B. D.; Reynolds, P.; Goldberg, D.; Hertz, A.; Jerrett, M.; et al., Long-Term Exposure to Air Pollution and Cardiorespiratory Disease in the California Teachers Study Cohort. *American Journal of Respiratory and Critical Care Medicine* **2011,** *184*, (7), 828-835.
5. Lepeule, J.; Laden, F.; Dockery, D.; Schwartz, J., Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. *Environ Health Perspect* **2012,** *120*, (7), 965-70.
6. Cesaroni, G.; Badaloni, C.; Gariazzo, C.; Stafoggia, M.; Sozzi, R.; Davoli, M.; et al., Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ Health Perspect* **2013,** *121*, (3), 324-31.
7. Puett, R. C.; Hart, J. E.; Suh, H.; Mittleman, M.; Laden, F., Particulate matter exposures, mortality, and cardiovascular disease in the health professionals follow-up study. *Environ Health Perspect* **2011,** *119*, (8), 1130-5.
8. Beelen, R.; Hoek, G.; van den Brandt, P. A.; Goldbohm, R. A.; Fischer, P.; Schouten, L. J.; et al., Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environ Health Perspect* **2008,** *116*, (2), 196-202.
9. Beelen, R.; Stafoggia, M.; Raaschou-Nielsen, O.; Andersen, Z. J.; Xun, W.; Katsouyanni, K.; et al., Long-term exposure to air pollution and cardiovascular mortality: an analysis of 22 European cohorts within the ESCAPE project. *Submitted* **2013**.
10. Kelly, F. J.; Fussell, J. C., Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment* **2012,** *60*, 504-526.
11. Zhou, J.; Ito, K.; Lall, R.; Lippmann, M.; Thurston, G., Time-series analysis of mortality effects of fine particulate matter components in Detroit and Seattle. *Environ Health Perspect* **2011,** *119*, (4), 461-6.
12. Valdes, A.; Zanobetti, A.; Halonen, J. I.; Cifuentes, L.; Morata, D.; Schwartz, J., Elemental concentrations of ambient particles and cause specific mortality in Santiago, Chile: a time series study. *Environ Health* **2012,** *11*, 82.
13. Ito, K.; Mathes, R.; Ross, Z.; Nadas, A.; Thurston, G.; Matte, T., Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. *Environ Health Perspect* **2011,** *119*, (4), 467-73.
14. Cao, J.; Xu, H.; Xu, Q.; Chen, B.; Kan, H., Fine particulate matter constituents and cardiopulmonary mortality in a heavily polluted Chinese city. *Environ Health Perspect* **2012,** *120*, (3), 373-8.
15. Ostro, B.; Lipsett, M.; Reynolds, P.; Goldberg, D.; Hertz, A.; Garcia, C.; et al., Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study. *Environ Health Perspect* **2010,** *118*, (3), 363-9.
16. DerSimonian, R.; Laird, N., Meta-analysis in clinical trials. *Control Clin Trials* **1986,** *7*, (3), 177-88.
17. de Hoogh, K.; Wang, M.; Adam, M.; Badaloni, C.; Beelen, R.; Birk, M.; et al., Development of land use regression models for particle composition in twenty study areas in europe. *Environ Sci Technol* **2013,** *47*, (11), 5778-86.
18. Eeftens, M.; Tsai, M. Y.; Ampe, C.; Anwander, B.; Beelen, R.; Bellander, T.; et al., Spatial variation of PM2.5, PM_{10}, $PM_{2.5}$ absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO_2 - Results of the ESCAPE project.

Atmospheric Environment **2012,** *62*, 303-317.
19. Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; et al., Development of NO_2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. *Atmospheric Environment* **2013,** *72*, 10-23.
20. Eeftens, M.; Beelen, R.; de Hoogh, K.; Bellander, T.; Cesaroni, G.; Cirach, M.; et al., Development of Land Use Regression Models for $PM_{2.5}$, $PM_{2.5}$ Absorbance, PM_{10} and PMcoarse in 20 European Study Areas; Results of the ESCAPE Project. *Environmental Science & Technology* **2012,** *46*, (20), 11195-11205.
21. Wang, M.; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B., Systematic Evaluation of Land Use Regression Models for NO_2. *Environmental Science & Technology* **2012,** *46*, (8), 4481-4489.
22. Dominici, F.; Peng, R. D.; Barr, C. D.; Bell, M. L., Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. *Epidemiology* **2010,** *21*, (2), 187-94.
23. Higgins, J. P.; Thompson, S. G., Quantifying heterogeneity in a meta-analysis. *Stat Med* **2002,** *21*, (11), 1539-58.
24. Pope, C. A., 3rd; Burnett, R. T.; Thun, M. J.; Calle, E. E.; Krewski, D.; Ito, K.; et al., Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* **2002,** *287*, (9), 1132-41.
25. Wang, M.; Beelen, R.; Basagana, X.; Becker, T.; Cesaroni, G.; de Hoogh, K.; et al., Evaluation of Land Use Regression Models for NO_2 and Particulate Matter in 20 European Study Areas: The ESCAPE Project. *Environ Sci Technol* **2013,** *47*, (9), 4357-64.
26. Jerrett, M.; Burnett, R. T.; Ma, R.; Pope, C. A., 3rd; Krewski, D.; Newbold, K. B.; et al., Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* **2005,** *16*, (6), 727-36.
27. Lippmann, M.; Ito, K.; Hwang, J. S.; Maciejczyk, P.; Chen, L. C., Cardiovascular effects of nickel in ambient air. *Environ Health Perspect* **2006,** *114*, (11), 1662-9.
28. Bell, M. L.; Ebisu, K.; Peng, R. D.; Samet, J. M.; Dominici, F., Hospital admissions and chemical composition of fine particle air pollution. *Am J Respir Crit Care Med* **2009,** *179*, (12), 1115-20.
29. Viana, M.; Kuhlbusch, T. A. J.; Querol, X.; Alastuey, A.; Harrison, R. M.; Hopke, P. K.; et al., Source apportionment of particulate matter in Europe: A review of methods and results. *Journal of Aerosol Science* **2008,** *39*, (10), 827-849.
30. Eeftens, M.; Hoek, G.; Gruzieva, O.; Mölter, A.; Agius, R.; Brunekreef, B.; et al., Elemental composition of particulate matter and the association with lung function in 5 European birth cohorts – results of the ESCAPE & TRANSPHORM projects. *Submitted* **2013**.
31. Ostro, B.; Feng, W. Y.; Broadwin, R.; Green, S.; Lipsett, M., The effects of components of fine particulate air pollution on mortality in california: results from CALFINE. *Environ Health Perspect* **2007,** *115*, (1), 13-9.
32. Sarnat, J. A.; Marmur, A.; Klein, M.; Kim, E.; Russell, A. G.; Sarnat, S. E.; et al., Fine particle sources and cardiorespiratory morbidity: an application of chemical mass balance and factor analytical source-apportionment methods. *Environ Health Perspect* **2008,** *116*, (4), 459-66.
33. Miller, K. A.; Siscovick, D. S.; Sheppard, L.; Shepherd, K.; Sullivan, J. H.; Anderson, G. L.; et al., Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med* **2007,** *356*, (5), 447-58.
34. Mostofsky, E.; Schwartz, J.; Coull, B. A.; Koutrakis, P.; Wellenius, G. A.; Suh, H. H.; et al., Modeling the association between particle constituents of air pollution and health outcomes. *Am J Epidemiol* **2012,** *176*, (4), 317-26.
35. Eeftens, M.; Beelen, R.; Fischer, P.; Brunekreef, B.; Meliefste, K.; Hoek, G., Stability of measured and modelled spatial contrasts in NO(2) over time. *Occup Environ Med* **2011,** *68*, (10), 765-70.
36. Gulliver, J.; de Hoogh, K.; Hansell, A.; Vienneau, D., Development and Back-Extrapolation of NO_2 Land Use Regression Models for Historic Exposure Assessment in Great Britain. *Environ Sci Technol* **2013,** *47*, (14), 7804-11.
37. Cesaroni, G.; Porta, D.; Badaloni, C.; Stafoggia, M.; Eeftens, M.; Meliefste, K.; et al., Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. *Environmental Health* **2012,** *11*.
38. Wang, R. R.; Henderson, S. B.; Sbihi, H.; Allen, R. W.; Brauer, M., Temporal stability of land use regression models for traffic-related air pollution. *Atmospheric Environment* **2013,** *64*, 312-319.

Chapter 6

Appendix

Online supplement 1	Description of each cohort and study area
Online supplement 2	Exposure assessment procedures and LUR model results for all study areas
Online supplement 3	Study population characteristics at baseline for each cohort (variables included in main model 3)
Online supplement 4	Description of exposure to particle components concentrations at participant addresses in each cohort
Online supplement 5	Correlations between elemental constituent in $PM_{2.5}$ and PM_{10} at participant addresses in each cohort
Online supplement 6	Sensitivity analysis

Online supplement 1: Description of each cohort and study area

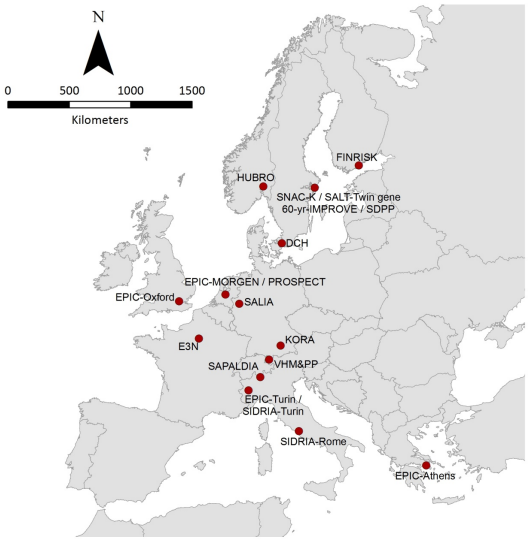


Figure S1: Cohort locations in which both PM and nitrogen oxides were measured

The National FINRISK Study (FINRISK), Finland

FINRISK surveys have been conducted every five years since 1972 to monitor the risk factor trends of chronic diseases, including cardiovascular diseases, diabetes, cancer, asthma, and allergy. For each survey, a stratified random sample has been selected from the 25-64 (74 since 1997) year old inhabitants in different regions of Finland. The ESCAPE study used FINRISK data from four surveys (1992, 1997, 2002, and 2007) and two study regions (the cities of Helsinki and Vantaa, and Turku city with its nearby municipalities). The FINRISK study protocol has been described elsewhere[1].

The surveys included a self-administered questionnaire (the questions focus mainly on socioeconomic factors, medical history, health behaviour, and psychosocial factors) and a clinical examination including measurements of height, weight and blood pressure and blood sampling. The participants have been annually followed up through 31 December, 2008 (up to 16 years) for fatal and nonfatal coronary and stroke events, and total mortality. The National Hospital Discharge Register and the National Causes of Death Register were used to identify these events. These registers cover every hospitalization in Finland and every death of permanent residents in Finland, yielding in practice 100% coverage of the follow-up events[2, 3]. In addition, we used the drug reimbursement records from the Social Insurance Institution of Finland to identify subjects who had developed diabetes or hypertension during the follow-up period.

The population-based Oslo Health Study (HUBRO), Norway

HUBRO was designed to identify health needs and the priorities of the health sector within Oslo, to monitor the developments and trends of diseases and their associated risks, to estimate the prevalence and later the incidence of chronic diseases, to investigate the social and geographical differences in health and associated risk factors and to initiate research to further investigate the aetiology of major health problems[4]. HUBRO was carried out in the city of Oslo from May 2000 to September 2001. All men and women born in the following years: 1924, 1925, 1940, 1941, 1954, 1955, 1960, 1969, and 1970, who resided in Oslo on December 31, 1999, were invited to participate. 58 178 subjects were invited and 22 699 individuals (39%) participated in the study. The questionnaires covered the following topics: health and chronic diseases, family history of disease, risk factors and lifestyles, social network, education, occupation, use of health services, and use of medicine. A physical exam was performed to obtain data on blood pressure, pulse recording, and collection of venous non-fasting blood samples. HUBRO was linked to the Norwegian Cause of Death Registry including deaths up to December 31, 2010, and was also linked to the Cancer Registry of Norway including cancers up to December 31, 2009.

SNAC-K, **The Swedish National study of Aging and Care in Kungsholmen (SNAC-K), Sweden**

SNAC-K is an ongoing longitudinal study aiming to investigate the ageing process and identify possible preventive strategies to improve health and care in elderly adults[5]. The study population

consists of randomly sampled individuals >=60 years old and in a central area of Stockholm (Kungsholmen) between March 2001 and June 2004. The sample was stratified for age and year of assessment giving sub-cohorts with 60, 66, 72, 78, 81, 84, 87, 90, 93, 96, and 99+ year olds. Information was collected through social interviews, assessment of physical functioning, clinical examination (incl. geriatric, neurological and physical assessments) as well as cognitive assessment. At baseline, information regarding events prior to the study period was gathered. The follow-up interval is six years for the younger age cohorts, and three years for the older age cohorts (81+). During the follow-up intervals, medical events of all subjects are registered through linkage with primary care registry and hospital discharge registry (available for all subjects in Sweden). In case of death, hospital and cause of death registries provide the clinical information, and informant interviews are carried out. The same protocol as for the baseline data collection is used during the follow-up, though only concerning the follow-up period. Website of study: http://www.aldrecentrum.se/snack- /index.htm. Any outcomes based on the Swedish nationwide health registries (such as the myocardial infarction and stroke registries, the cause-of-death register and the national patient register) have been used.

Stockholm Screening Across the Lifespan Twin study (SALT) & Twin GENE (subcohort), Sweden

The SALT study was set-up to screen all twins born in Sweden before 1958 for the most common complex diseases with a focus on cardiovascular diseases[6, 7]. Twin Gene is a sub-study involving establishing a biobank with DNA and serum from SALT participants. SALT is a telephone interview, which took place between 1998-2002. For the purposes of this study, only twins living in Stockholm County are included in the analyses. Information concerning birth order and weight, zygosity, contact with twin partner and family constellation, diseases, use of medication, occupation, education, life style habits, gender- and age-specific (hormone replacement therapy) and memory problems (age > 65) was collected. In Twin Gene, twins born before 1958 were contacted 2004-2008, a total number of ~2500 participants was available. Health and medication data were collected from questionnaires. Blood sampling material was mailed to study subjects, who contacted a local health care centre for blood sampling and a health check-up. Height, weight, circumference of waist and hip, and blood pressure was measured and blood was collected. Any outcomes based on the Swedish nationwide health registries (such as the myocardial infarction and stroke registries, the cause-of-death register and the national patient register) have been used.

Stockholm 60 year olds & IMPROVE, Sweden

The 60 year olds cohort is a study aiming to identify biological and socio-economic risk factors and predictors for cardiovascular diseases[8]. Recruitment took place between August 1997 and March 1999. A random sample of every third man and woman living in Stockholm County, who was born between 1 July 1937 and 30 June 1938, was invited to the 60 year olds study. In total ~4100 subjects were included. Height, weight, BMI, Waist/Hip ratio and resting ECD, blood pressure and fasting blood samples were taken during a physical examination, while a comprehensive questionnaire was completed, including information on socioeconomic, medical and life-style factors. The study was supplemented 2003 by the IMPROVE project (an ongoing multi cohort study into genetics and CVDs). In Stockholm, IMPROVE is a sub-cohort consisting of ~500 participants from the 60 year olds cohort with inclusion criteria of having at least three risk factors for the metabolic syndrome. For IMPROVE, three follow-ups were conducted, blood and urine were collected, socio-economic data, quantitative B-mode ultra sound examination of carotid arteries and replicate B-mode ultrasound was performed, and vascular events were recorded. Any outcomes based on the Swedish nationwide health registries (such as the myocardial infarction and stroke registries, the cause-of-death register and the national patient register) have been used.

Stockholm SDPP, - Stockholm diabetes preventive programme (SDPP), Sweden

The Stockholm diabetes prevention programme, a population-based prospective study, aimed at investigating the aetiology of type 2 diabetes and developing prevention strategies for type 2 diabetes[9]. An initial survey included all men and women in the targeted age group in Stockholm County; for men in four municipalities (Värmdö, Upplands Bro, Tyresö and Sigtuna), and for women these four plus a fifth municipality (Upplands Väsby). All were screened by a questionnaire regarding presence of own diabetes and diabetes in relatives. Subjects with family history of diabetes (FHD) and randomly selected subjects without FHD, all without previously diagnosed diabetes, were invited to a health examination. This baseline study, 1992-1994 for men and 1996-1998 for women, comprised 7949 subjects, aged 35-56 years, and about 50% had FHD. In the follow-up study 8-10 years later, 2383 men (2002-2004) and 3329 women (2004-2006) participated. At the health examinations, both at baseline and follow-up, an extensive questionnaire (information

on lifestyle factors, such as physical activity, dietary habits, tobacco use, alcohol consumption, health status, socioeconomic status and psychosocial conditions) was completed. Diabetes heredity was confirmed and measurements of weight, height, hip and waist circumference as well as blood pressure were performed. In addition, an oral glucose tolerance test (OGTT) was made, and blood was sampled at fasting state and 2 hour after glucose intake. Outcomes based on the Swedish nationwide health registries (such as the myocardial infarction and stroke registries, the cause-of-death register, and the national patient register) have been used.

Danish Diet Cancer and Health study (DCH), Denmark

The primary aim of the DCH study is to investigate diet and lifestyle in relation to incidence of cancer and other chronic diseases[10]. The study combines the collection of questionnaire data with storing of biological specimen in order to investigate genetic susceptibility and gene-environment interactions with regard to diet, dietary compounds, and the risk of cancer, and indeogenous markers of nutritional, metabolic, and hormonal characteristics of study participants. Historical residential history of the study participants is available, which facilitate studies of air pollution and noise. The study enrolled participants in two areas, Copenhagen and Aarhus, Denmark. 160 725 individuals aged 50-64 years were invited to participate between December 1993 and May 1997. All participants were Danish-born, living in the Copenhagen or Aarhus areas and without medical history of cancer diagnosis registered in the Danish Cancer Registry at the time of invitation. Out of the 160 725 people invited, which were a random sample of all eligible individuals in the specified areas, 57 053 were enrolled. Due to the geographical limitations of the land use regression, only the almost 40 000 participants from the Copenhagen area were included in the ESCAPE analyses. On enrolment, each participant completed self-administered questionnaires (in Danish) that included questions on dietary habits, health status, family history of cancer, social factors, reproductive factors, smoking, environmental smoking, and lifestyle habits. Anthropometric measurements including blood pressure and blood samples were also obtained. The DCH cohort is followed up regularly by use of complete nationwide registers hence the loss to follow-up is virtually nil. Data on cancer incidence from the Danish Cancer Registry and data on cause-specific mortality from the Danish Mortality Registry were used.

Study on the influence of Air pollution on Lung function, Inflammation and Aging (SALIA), Germany

The SALIA study was initiated in 1985 as part of Environmental Health surveys to monitor health effects of outdoor air pollution in the heavily polluted Ruhr Area[11, 12]. It was an element of the Clean Air Plan initiated by the Government of North-Rhine Westphalia in Germany. The geographic regions covered were parts of Dortmund (1985, 1990), Duisburg (1990), Essen (1990), Gelsenkirchen (1986,1990) and Herne (1986). They were chosen to represent a range of polluted areas with high traffic load and steel and coal industries. Two non-industrial small towns, Dülmen (1985) and Borken (1985, 1986, 1987, 1990, 1993, 1994) were chosen as reference areas. The Research Institute for Environmental Medicine in Düsseldorf (then Medical Institute of Environmental Hygiene) coordinated the studies. The baseline investigations of SALIA were cross-sectional surveys. They were conducted on 4757 women in the local health departments in March and April between 1985 and 1994. Sampling included all women of German nationality aged 54 to 55 residing in the selected areas. Women were selected because men in these areas mainly worked in the mining industry with very high occupational exposure probably masking the effects of air pollution. Postal questionnaires were sent out and included information about airway diseases and covariates. The filled in questionnaires were checked at the day of investigation. Overall questionnaire response was 70%. Specific measurements (lung function, determination of immunological markers, and xenobiotics) were added in subgroups. All investigations were done according to standardized operating procedures.

Height and weight was measured at the day of investigation. These measurements are not available for more than 10% of all women. Therefore BMI was not included in the ESCAPE analysis, after having demonstrated that BMI did influence the results only marginally.

Follow-ups were set up to investigate the effects of outdoor air pollution and changes in pollution on respiratory health and mortality. In 1990, women investigated in 1984/1986 had a first follow-up investigation including a questionnaire and a lung function testing. A mortality follow-up of all women having participated in the baseline investigation was conducted in 2003 and in 2008 by the Institute of Epidemiology Helmholtz Munich. All surviving women were asked to participate in a questionnaire follow-up in 2006 and invited to eventually participate in a follow-up investigation. All women with lung-function available at baseline were invited to a more detailed follow-up investigation, which started in 2007.

The mortality analyses of ESCAPE use questionnaire data from the baseline investigation and the data from the mortality follow-ups. All these data were available to be included in the ESCAPE analysis.

All women with geocoded addresses at baseline were included in the analysis (4,663). Two continuous covariates were used as year of recruitment, early (1985, 1986, and 1987) and late (1990, 1993, and 1994) years. Coding was year of recruitment – 1900, recruitment before 1990 was coded as 90 in the late variable, recruitment after 1990 was coded as 87 in the early variable. No dietary covariates were available, environmental tobacco smoke was a combined variable from home and work place, occupational exposure was extreme temperature and dust. Area SES was defined as income-rate per five-digit postcode-area.

The Cooperative Health Research in the Region of Augsburg (KORA), Germany

KORA is a cohort study based on four cross-sectional surveys of a random sample of inhabitants of the Augsburg region[13]. Main objectives of the baseline study were to investigate cardiovascular and other chronic diseases regarding: 1) to assess health indicators (morbidity, mortality) and health care (utilization, costs), 2) to quantify the prevalence of risk factors, and 3) to study the impact of lifestyle, metabolic and genetic factors. The follow-up studies aimed to assess also time-trends in risk factors and health over a period of seven to ten years. Two cross-sectional population-representative surveys were conducted in 1994-1995 (survey S3) and 1999-2001 (survey S4) in the city of Augsburg and two adjacent rural counties to include all inhabitants of the Augsburg region with German nationality aged 25 to 74 (n=400 000). Follow-up examinations of survey S3 and S4 participants were carried out seven to ten years later. Baseline examination included standardized interviews, physical examination, and blood sampling. All investigations were done according to standardized operating procedures.

Follow-up investigations were conducted in 2004-2005 for survey S3 and in 2006-2008 for survey S4. 2974 and 3080 of survey S3 and S4 participants attended the follow-up examinations including standardized computer-assisted interview, self-administered questionnaire, physical examination, and blood sampling. Survival was ascertained for S3 participants in 2008 through Population Registry search and is available from the time of recruitment until December 31 2007. Survival of S4 participants was ascertained through a combination of returned questionnaires and subsequent Population Registry search and is available from recruitment until December 31, 2008. Causes of death are abstracted for all deaths from the death certificates. For the ESCAPE analyses a study/baseline indicator was included instead of calendar time.

The Vorarlberg Health Monitoring and Prevention Program (VHM&PP), Austria

The VHM&PP study is a prevention program routinely performed by the Agency of Social and Preventive Medicine and covers all adults of the whole province[14, 15]. It has been ongoing since 1985 and data are presently available until 2005. Recruitment and follow-up has been ongoing that means during the whole period new persons were recruited and already recruited persons came for follow-up visits. The total adult population of the state Vorarlberg is covered, with voluntary enrolment. Data are available from 1985 to 2005 at present on 185 330 persons, corresponding to about 65% participation. Their age at recruitment rangew from 18-97 years (mean=42). The screening examination takes place in the practice of the local physicians; a self-administered questionnaire is also applied. The same protocol was applied at baseline and follow-up examinations. A total of 132 242 geocodes were assigned exposures. 30 718 geocodes (18.85%) were omitted if: 1) they were entirely outside of the Vorarlberg state, 2) within 300m of the state boundary (lack of GIS data in neighboring countries), and 3) if their elevation was > 600m.

Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA), Switzerland

SAPALDIA is a multi-center study performed in eight geographic areas representing the range of environmental, meteorological, and socio-demographic conditions in Switzerland[16]. A random population sample across eight geographic areas (Aarau, Basel, Davos, Geneva, Lugano, Montana, Payerne, and Wald) was obtained in 1991, with follow-ups in 2002 and 2010. The main aim of the study was to assess the effect of air pollution (outdoor and indoor) on respiratory and cardiovascular health, with a special focus on how the respiratory and cardio-vascular systems interact in this regard, and on the role of lifestyle and genetic background. In 1991, 9651 subjects, aged 18 to 60 years, were recruited via detailed interviews and more than 90% provided valid spirometry results. The follow-up in 2002 obtained health information and anthropometric data from physical re-examination with spirometry and blood sampling, blood pressure measurement, and heart rate variability in a subsample (<50yrs). The most recent follow-up (SAPALDIA 3) was in 2010. In the third assessment, study subjects were also asked in detail about chronic diseases having been

diagnosed and treated since the second survey. Questionnaire domains are the following: respiratory health and disease, cardiovascular health and disease, chronic disease and relevant risk factors, women's health, allergies, medications, drug use, exposure to air pollutants, sleep apnea, and health care resources used. SAPALDIA did not obtain information on physical activity, alcohol intake, and nutrition at baseline in 1991. Within ESCAPE, only the areas of Basel, Geneva, and Lugano were included, with PM measurements in Lugano only.

Italian Studies on Respiratory Disorders in Childhood and Environment (SIDRIA)

The SIDRIA study has been an extension of the ISAAC initiative in Italy (International Study on Asthma and Allergies in Childhood), a worldwide survey to analyse variations in prevalence of symptoms asthma, rhinitis, and atopic eczema[17]. A cross-sectional survey was carried out between October 1994 and March 1995 in eight centres in northern and central Italy using standardised questionnaires (response rate=94%). Parents of first and second graders from a representative sample of primary schools, and adolescents in the third year of a representative sample of junior high schools answered a self-administered questionnaire on the child's health status, as well as their personal respiratory health status and various risk factors, including education, occupation, housing conditions, smoking habits, and traffic intensity in their area of residence. The data used within ESCAPE are from the subset of parents recruited in two metropolitan areas: Rome and Turin, in the context of a project co-funded by the Ministry of Health (Programma Strategico Ambiente e Salute, Ricerca Finalizzata ex-art.12, 2006). A record linkage has been performed with the Municipal Registry Office Databases to collect the residential history of parents who were living in Rome and Turin with their children at the time of the survey. In the city of Turin the project was performed through a collaboration between SIDRIA and the regional Unit of Epidemiology (ASL TO3), in the context of the Turin Longitudinal Study, a census-based cohort study following up health outcomes of people censused in Turin since 1971. It was possible to identify ~16 000 adults.

European Prospective Investigation into Cancer and Nutrition (EPIC)

The European Prospective Investigation into Cancer and Nutrition (EPIC), which covers a large cohort of half a million men and women from 23 European centers in 10 Western European countries, was designed to study the relationship between diet and the risk of chronic diseases, particularly cancer[18]. Five of these centers were included in ESCAPE. The selection of ESCAPE participants was done centrally at Imperial College, UK using the central EPIC database.

EPIC- Monitoring Project on Risk Factors and chronic diseases in the Netherlands (MORGEN), The Netherlands

The MORGEN cohort consists of a general population sample of 10 260 men and 12 394 women aged 20–59 years from three Dutch towns (Amsterdam, Doetinchem and Maastricht)[19]. From 1993 to 1997 each year a new random sample, consisting of 6000 subjects, was examined. A total of 50 766 persons were invited to participate in the MORGEN cohort. Those who replied received two questionnaires by mail (a general questionnaire on socio- demographic factors, lifestyle and health indicators, and an FFQ and were invited to visit the local Public Health Service for a medical examination). The EPIC-MORGEN cohort and the EPIC-PROSPECT cohort have been joined to form the EPIC-NL cohort. All members of the EPIC-NL cohort are followed for changes in vital status and the occurrence of diseases by linkage with several registries, including the Municipality registry for vital status, the Dutch National Cancer registry for occurrence of cancer, the Central Bureau of Statistics registry for causes of death, and a National Hospital Discharge Diagnosis registry for occurrence of cardiovascular diseases or type 2 diabetes. Changes in some exposure status are assessed by questionnaires during follow-up. Part of the MORGEN cohort (Doetinchem participants) is re-invited every five years for a physical examination in addition to questionnaire information. The MORGEN cohort of EPIC-NL is linked to the Dutch Cancer Registry because participants are residing in several geographical areas covered by different regional integral cancer centres.

EPIC-Prospect, the Netherlands

A total of ~17 500 healthy women, living in Utrecht and surroundings, were enrolled[20]. Women were recruited from breast cancer screening participants, age 50-70 years at enrolment. The purpose of the EPIC-PROSPECT study is to assess the relation between nutrition and cancer and other chronic diseases. Baseline information was collected between 1993-1997 on the basis of two self-administered questionnaires and a medical examination. The general questionnaire contains questions on demographic characteristics, presence of chronic diseases of interest, and risk factors for chronic diseases of interest, i.e. blood pressure, serum cholesterol, reproductive history of women, family history, smoking habits, drinking of alcohol, and physical activity. Dietary intake was assessed using detailed food frequency questionnaires. A medical examination was performed

including measurement of blood pressure, anthropometric measurements and taking of blood. All EPIC-PROSPECT participants are followed-up by questionnaire at 3-5 year intervals. The questionnaire collects information on changes in lifestyle habits as well as on health status. All incident and prevalent cancer cases were identified through linkage to the regional cancer registry, IKMN (Integraal Kankercentrum Midden Nederland), then from the National Cancer Registry from 2008 onwards. Vital status and cause-specific mortality information is obtained through linkage to the municipality registries and Central Buro of Statistics.

European Prospective Investigation into Cancer and Nutrition (EPIC) –Oxford, UK

The Oxford cohort was recruited from the nationwide general population in urban and rural areas throughout the United Kingdom, although a large percentage comes from Southern parts of England and big cities such as London[21]. The cohort contains 65 429 men and women over 20 years of age recruited through medical general practices or by post between 1993 and 1999, with an emphasis on vegetarians. The questionnaires gathered information on diet (FFQ and 24hr recall), social and demographic factors, lifestyle, anthropometrics, medical history of diseases, and prevalent cancers; approximately 20 000 gave a blood sample. Participants who consented were followed-up from recuritment by "flagging" on the NHS Central Registers (NHSCRs) in England and Wales (via the Office for National Statistics), Scotland (via the General Registry Office) and Northern Ireland (via the Northern Ireland Cancer Registry) via automatic notifications. The date of each event and coding of the cancer site or type and the causes of death were recorded according to the 10th revision of the International Classification of Diseases (ICD-10). For incident cancers, tumour morphology is also coded, according to WHO ICD-O. EPIC-Oxford website: http://www.epic-oxford.org. The study population was restricted to ~45 650 participants living within 400Km threshold of ESCAPE monitoring sites.

EPIC -Turin

Recruitment took place from 1993 and involved blood donors and other healthy volunteers, accruing 10 604 participants by 1998[22]. Co-operation with the local cancer registry and the local health authority allows for access to hospital discharge information and all newly diagnosed cancer cases. Follow-up started in 1998, including collaboration with the local cancer registry, the demographic computerized archives of the Torino area and the discharge report database for hospital patients.

EPIC-Greece

Recruitment of volunteers in EPIC-Greece started in 1994, and was completed in 1999[18]. In total, 16 619 women and 11 953 men were recruited from Greece nationwide. Data collection on medical and reproductive history, socio-demographic and lifestyle factors and habitual diet was performed via interview and a baseline examination that recorded measurement of anthropometric data and collection of blood samples. The follow-up of study participants was initiated in January 1997 and focused on the update of information on lifestyle factors and the health status. Due to the lack of a national cancer registry and the country-wide nature of EPIC-Greece, information is being collected through self-administered questionnaires or telephone interviews. Reported diagnoses of interest were further ascertained through consultation of medical files in hospitals and clinics all over Greece, or through the collection of death certificates from the regional death registries, in case of death. Participants that contribute to the ESCAPE analyses are residents from the Prefecture of Attica (which comprises mainly the Greater Athens Area, and hence called EPIC-Athens in the manuscript). Based on GIS availability, we included only the members if the EPIC cohort who were residents of 16 municipalities, specifically Athens, Agios Ioannis Rentis, Amaroussion, Egaleo, Galatsi, Halandri, Ilioupolis, Kalithea, Moschato, Nea Ionia, Nea Smyrni, Nikaia, Peristeri, Pireaus, Tavros, and Zografou.

Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale (E3N), France

E3N is a large ongoing prospective cohort consisting of 98 995 French women born between 1925 and 1950, subscribing to the health insurance plan for public education system employees, and who voluntarily enrolled in 1990-1991[23]. The main objective of the study was to investigate the risk factors for breast cancer among women in particular hormonal factors and diet. This study began in 1990 when a baseline questionnaire (Q1) was sent to the 103 089 out of the 494 458 women subscribed to the health insurance plan for public education system employees women aged 40–65 years who agreed to participate. Follow up questionnaires were sent in January 1992 (Q2) and then approximately every two years thereafter. The most recent update questionnaires was sent in June 2008 (Q9) and another one in 2010. The base population covers the whole country

of France and participation was based on voluntary agreement. To date, participants have been followed for 18 years (from 1991 to 2008) with complete data available from 2005. All the questionnaires are self-administered and are sent by mail to participants in French language and returned to the study centre at IGR, Paris. Biological material was collected in 1996 on 25 000 women out of the 68 000 (who lived in communes with at least 1000 participants) invited to participate in the setting up of the biological bank. While the E3N study includes a large population in all France, exposure assessment for the ESCAPE project was available only for 4 cities: Paris, Lyon, Grenoble and Marseille. PM measurements were only done in Paris. E3N is the French component of EPIC.

References

1. Vartiainen, E.; Laatikainen, T.; Peltonen, M.; Juolevi, A.; Mannisto, S.; Sundvall, J.; et al., Thirty-five-year trends in cardiovascular risk factors in Finland. *Int J Epidemiol* 2010, *39*, (2), 504-18.
2. Pajunen, P.; Koukkunen, H.; Ketonen, M.; Jerkkola, T.; Immonen-Raiha, P.; Karja-Koskenkari, P.; et al., The validity of the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil* 2005, *12*, (2), 132-7.
3. Tolonen, H.; Salomaa, V.; Torppa, J.; Sivenius, J.; Immonen-Raiha, P.; Lehtonen, A., The validation of the Finnish Hospital Discharge Register and Causes of Death Register data on stroke diagnoses. *Eur J Cardiovasc Prev Rehabil* 2007, *14*, (3), 380-5.
4. Sogaard, A. J.; Selmer, R.; Bjertness, E.; Thelle, D., The Oslo Health Study: The impact of self-selection in a large, population-based survey. *Int J Equity Health* 2004, *3*, (1), 3.
5. Lagergren, M.; Fratiglioni, L.; Hallberg, I. R.; Berglund, J.; Elmstahl, S.; Hagberg, B.; et al., A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). *Aging Clin Exp Res* 2004, *16*, (2), 158-68.
6. Lichtenstein, P.; Sullivan, P. F.; Cnattingius, S.; Gatz, M.; Johansson, S.; Carlstrom, E.; et al., The Swedish Twin Registry in the third millennium: an update. *Twin Res Hum Genet* 2006, *9*, (6), 875-82.
7. Magnusson, P. K.; Almqvist, C.; Rahman, I.; Ganna, A.; Viktorin, A.; Walum, H.; et al., The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Res Hum Genet* 2013, *16*, (1), 317-29.
8. Wandell, P. E.; Wajngot, A.; de Faire, U.; Hellenius, M. L., Increased prevalence of diabetes among immigrants from non-European countries in 60-year-old men and women in Sweden. *Diabetes Metab* 2007, *33*, (1), 30-6.
9. Eriksson, A. K.; Ekbom, A.; Granath, F.; Hilding, A.; Efendic, S.; Ostenson, C. G., Psychological distress and risk of pre-diabetes and Type 2 diabetes in a prospective study of Swedish middle-aged men and women. *Diabet Med* 2008, *25*, (7), 834-42.
10. Tjonneland, A.; Olsen, A.; Boll, K.; Stripp, C.; Christensen, J.; Engholm, G.; et al., Study design, exposure variables, and socioeconomic determinants of participation in Diet, Cancer and Health: a population-based prospective cohort study of 57,053 men and women in Denmark. *Scand J Public Health* 2007, *35*, (4), 432-41.
11. Heinrich, J.; Thiering, E.; Rzehak, P.; Kramer, U.; Hochadel, M.; Rauchfuss, K. M.; et al., Long-term exposure to NO_2 and PM_{10} and all-cause and cause-specific mortality in a prospective cohort of women. *Occup Environ Med* 2013, *70*, (3), 179-86.
12. Gehring, U.; Heinrich, J.; Kramer, U.; Grote, V.; Hochadel, M.; Sugiri, D.; et al., Long-term exposure to ambient air pollution and cardiopulmonary mortality in women. *Epidemiology* 2006, *17*, (5), 545-51.
13. Holle, R.; Happich, M.; Lowel, H.; Wichmann, H. E., KORA--a research platform for population based health research. *Gesundheitswesen* 2005, *67 Suppl 1*, S19-25.
14. Ulmer, H.; Kelleher, C. C.; Fitz-Simon, N.; Diem, G.; Concin, H., Secular trends in cardiovascular risk factors: an age-period cohort analysis of 698,954 health examinations in 181,350 Austrian men and women. *J Intern Med* 2007, *261*, (6), 566-76.
15. Ulmer, H.; Bachmann, J.; Huber, K.; Concin, H.; Bischof, H. P., [Follow-up studies of routine health screening in Vorarlberg 1986-1994]. *Wien Klin Wochenschr* 1997, *109*, (5), 160-4.
16. Downs, S. H.; Schindler, C.; Liu, L. J.; Keidel, D.; Bayer-Oglesby, L.; Brutsche, M. H.; et al., Reduced exposure to PM_{10} and attenuated age-related decline in lung function. *N Engl J Med* 2007, *357*, (23), 2338-47.
17. Cesaroni, G.; Badaloni, C.; Porta, D.; Forastiere, F.; Perucci, C. A., Comparison between various indices of exposure to traffic-related air pollution and their impact on respiratory health in adults. *Occup Environ Med* 2008, *65*, (10), 683-90.
18. Riboli, E.; Hunt, K. J.; Slimani, N.; Ferrari, P.; Norat, T.; Fahey, M.; et al., European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* 2002, *5*, (6B), 1113-24.
19. Beulens, J. W.; Monninkhof, E. M.; Verschuren, W. M.; van der Schouw, Y. T.; Smit, J.; Ocke, M. C.; et al., Cohort profile: the EPIC-NL study. *Int J Epidemiol* 2010, *39*, (5), 1170-8.
20. Boker, L. K.; van Noord, P. A.; van der Schouw, Y. T.; Koot, N. V.; Bueno de Mesquita, H. B.; Riboli, E.; et al., Prospect-EPIC Utrecht: study design and characteristics of the cohort population. European Prospective Investigation into Cancer and Nutrition. *Eur J Epidemiol* 2001, *17*, (11), 1047-53.
21. Davey, G. K.; Spencer, E. A.; Appleby, P. N.; Allen, N. E.; Knox, K. H.; Key, T. J., EPIC-Oxford: lifestyle characteristics and nutrient intakes in a cohort of 33 883 meat-eaters and 31 546 non meat-eaters in the UK. *Public Health Nutr* 2003, *6*, (3), 259-69.
22. Guarrera, S.; Ricceri, F.; Polidoro, S.; Sacerdote, C.; Allione, A.; Rosa, F.; et al., Association between total number of deaths, diabetes mellitus, incident cancers, and haplotypes in chromosomal region 8q24 in a prospective study. *Am J Epidemiol* 2012, *175*, (6), 479-87.
23. Clavel-Chapelon, F.; Jadand, C.; Goulard, H.; Guibout-Peigne, C., [E3N, a cohort study on cancer risk factors in MGEN women. Description of protocol, main characteristics and population]. *Bull Cancer* 1996, *83*, (12), 1008-13.

Online supplement 2: Exposure assessment procedures and LUR model results for all study areas

Exposure assessment

Air pollution concentrations at the baseline residential addresses of study participants were estimated by Land Use Regression (LUR) models following a standardized procedure that has been described elsewhere [1-3]. In brief, air pollution monitoring campaigns were performed between October 2008 and May 2011 in all study areas. Three two-week measurements of particles with aerodynamic diameter <2.5μm ($PM_{2.5}$) and <10μm (PM_{10}) were performed at 20 sites in 19 study areas. The three measurements were then averaged, adjusting for temporal trends using data from a background monitoring site with continuous data [4, 5]. PM filters were weighted before and after each measurement centrally at IRAS, Utrecht University and were then sent to Cooper Environmental Services (Portland, OR, USA) to detect metal components. All filters were analyzed for elemental composition using X-Ray Fluorescence (XRF)[3]. Predictor variables on nearby traffic intensity, population/household density and land use were derived from Geographic Information Systems (GIS), and were evaluated to explain spatial variation of annual average concentrations using regression modeling. LUR model results for all study areas are shown below. The LUR models were used to estimate ambient air pollution concentration at the participants' addresses. If values of predictor variables for the cohort addresses were outside the range of values for the monitoring sites, values were truncated to the minimum and maximum values at the monitoring sites. Truncation was performed to prevent unrealistic predictions (e.g. related to too small distance to roads in GIS) and because we did not want to extrapolate the derived model beyond the range for which it was developed. Truncation has been shown to improve predictions at independent sites[6]. In total, 13 of 208 constituents' models were not available for cohorts in corresponding areas (Table S1). The lack of any model was probably related to small within-area variability (e.g. S), poor precision of the measurements with low concentrations (Ni and V), lack of availability of predictor variables representing the major source of a component, and complexity of the study area.

Table S1 Study areas of cohorts where LUR models were not available

Constitutes	$PM_{2.5}$	PM_{10}
Cu	NA	NA
Fe	NA	NA
K	SALIA, Germany	HUBRO, Norway
Ni	SNAC-K, Sweden SALT/Twin gene, Sweden 60-y/IMPROVE, Sweden SDPP, Sweden	SAPALDIA, Switzerland HUBRO, Norway
S	SAPALDIA, Switzerland	SAPALDIA, Switzerland
Si	SAPALDIA, Switzerland EPIC-Athens, Greece HUBRO, Norway	SAPALDIA, Switzerland
V	SAPALDIA, Switzerland KORA, Germany VHM&PP, Austria HUBRO, Norway	SAPALDIA, Switzerland
Zn	NA	NA

Table S2 PM_{10} and $PM_{2.5}$ Cu model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	0.61	X		X						
	$PM_{2.5}$	0.84	X		X						
HUBRO	PM_{10}	0.87	X		X						
	$PM_{2.5}$	0.76	X	X	X						
SNAC-K	PM_{10}	0.84	X	X							
	$PM_{2.5}$	0.61		X			X				
SALT	PM_{10}	0.84	X	X							
	$PM_{2.5}$	0.61		X			X				
60-yr	PM_{10}	0.84	X	X							
	$PM_{2.5}$	0.61		X			X				
SDPP	PM_{10}	0.84	X	X							
	$PM_{2.5}$	0.61		X			X				
DCH	PM_{10}	0.91		X							
	$PM_{2.5}$	0.61	X	X							
EPIC-MORGEN,	PM_{10}	0.71	X	X					X		
	$PM_{2.5}$	0.81	X			X					X
EPIC-PROSPECT,	PM_{10}	0.71	X	X					X		
	$PM_{2.5}$	0.81	X			X					X
SALIA	PM_{10}	0.92	X	X			X				
	$PM_{2.5}$	0.9	X	X			X				X
EPIC- Oxford	PM_{10}	0.95	X	X							
	$PM_{2.5}$	0.79	X			X					
KORA	PM_{10}	0.71	X				X				
	$PM_{2.5}$	0.76	X	X			X				
VHM&PP	PM_{10}	0.95	X				X	X			
	$PM_{2.5}$	0.38		X				X			
SAPALDIA,	PM_{10}	0.84	X	X							
	$PM_{2.5}$	0.83	X	X							
E3N	PM_{10}	0.48		X			X				
	$PM_{2.5}$	0.51	X				X	X			
EPIC- Turin	PM_{10}	0.88	X		X			X			
	$PM_{2.5}$	0.85		X	X			X			
SIDRIA-Turin	PM_{10}	0.88	X		X			X			
	$PM_{2.5}$	0.85		X	X			X			
SIDRIA-Rome	PM_{10}	0.87	X	X							
	$PM_{2.5}$	0.78	X	X							
EPIC- Athens	PM_{10}	0.7	X	X							
	$PM_{2.5}$	0.76	X	X		X		X			

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants
[d]RES includes low and high density residential
[e]Industry indicates industry area
[f]Green includes natural green and urban green variables
[g]Port indicates port area
[h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

Table S3 PM_{10} and $PM_{2.5}$ Fe model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	0.47	X	X	X						
	$PM_{2.5}$	0.63	X		X			X			
HUBRO	PM_{10}	0.92	X		X						
	$PM_{2.5}$	0.82	X							X	
SNAC-K	PM_{10}	0.68	X	X							
	$PM_{2.5}$	0.9	X		X	X	X				
SALT	PM_{10}	0.68	X	X							
	$PM_{2.5}$	0.9	X		X	X	X				
60-yr	PM_{10}	0.68	X	X							
	$PM_{2.5}$	0.9	X		X	X	X				
SDPP	PM_{10}	0.68	X	X							
	$PM_{2.5}$	0.9	X		X	X	X				
DCH	PM_{10}	0.92		X				X			
	$PM_{2.5}$	0.91	X	X			X	X			
EPIC-MORGEN,	PM_{10}	0.7	X	X	X						
	$PM_{2.5}$	0.73	X			X					X
EPIC-PROSPECT,	PM_{10}	0.7	X	X	X						
	$PM_{2.5}$	0.73	X			X					X
SALIA	PM_{10}	0.85	X	X	X		X				
	$PM_{2.5}$	0.62	X	X			X				
EPIC- Oxford	PM_{10}	0.95	X		X						
	$PM_{2.5}$	0.92	X		X						
KORA	PM_{10}	0.82	X	X		X					
	$PM_{2.5}$	0.62	X	X			X				
VHM&PP	PM_{10}	0.62	X			X					
	$PM_{2.5}$	0.46	X			X					
SAPALDIA,	PM_{10}	0.85	X	X							
	$PM_{2.5}$	0.83	X	X							
E3N	PM_{10}	0.5		X			X				
	$PM_{2.5}$	0.79	X	X		X					
EPIC- Turin	PM_{10}	0.86	X		X			X			
	$PM_{2.5}$	0.83	X		X	X		X			
SIDRIA- Turin	PM_{10}	0.86	X		X			X			
	$PM_{2.5}$	0.83	X		X	X		X			
SIDRIA-Rome	PM_{10}	0.82	X	X							X
	$PM_{2.5}$	0.67	X			X					
EPIC- Athens	PM_{10}	0.75	X	X			X	X			
	$PM_{2.5}$	0.11	X	X							

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants
[d]RES includes low and high density residential
[e]Industry indicates industry area
[f]Green includes natural green and urban green variables
[g]Port indicates port area
[h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

Table S4 PM_{10} and $PM_{2.5}$ K model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	0.24	X					X			
	$PM_{2.5}$	0.09						X			
HUBRO	PM_{10}	0.9	X		X	X				X	
	$PM_{2.5}$	0.32			X					X	
SNAC-K	PM_{10}	0.76	X			X					
	$PM_{2.5}$	0.55	X			X					
SALT	PM_{10}	0.76	X			X					
	$PM_{2.5}$	0.55	X			X					
60-yr	PM_{10}	0.76	X			X					
	$PM_{2.5}$	0.55	X			X					
SDPP	PM_{10}	0.76	X			X					
	$PM_{2.5}$	0.55	X			X					
DCH	PM_{10}	0.15						X			
	$PM_{2.5}$	0.53					X	X			
EPIC-MORGEN,	PM_{10}	0.45		X		X					
	$PM_{2.5}$	0.25		X							X
EPIC-PROSPECT,	PM_{10}	0.45		X		X					
	$PM_{2.5}$	0.25		X							X
SALIA	PM_{10}	0.14					X				
	$PM_{2.5}$	NA									
EPIC- Oxford	PM_{10}	0.56				X					
	$PM_{2.5}$	0.14				X					
KORA	PM_{10}	0.63	X					X			
	$PM_{2.5}$	0.38	X					X		X	
VHM&PP	PM_{10}	0.6			X		X	X			
	$PM_{2.5}$	0.69			X			X		X	
SAPALDIA,	PM_{10}	0.67	X					X			
	$PM_{2.5}$	0.78	X					X		X	
E3N	PM_{10}	0.52	X					X			
	$PM_{2.5}$	0.31	X			X	X				
EPIC- Turin	PM_{10}	0.48				X		X		X	
	$PM_{2.5}$	0.11	X			X					
SIDRIA- Turin	PM_{10}	0.48				X		X		X	
	$PM_{2.5}$	0.11	X			X					
SIDRIA- Rome	PM_{10}	0.57	X								X
	$PM_{2.5}$	0.41	X								X
EPIC- Athens	PM_{10}	0.13	X								
	$PM_{2.5}$	0.51		X		X		X			

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants
[d]RES includes low and high density residential
[e]Industry indicates industry area
[f]Green includes natural green and urban green variables
[g]Port indicates port area
[h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

Table S5 PM_{10} and $PM_{2.5}$ Ni model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	0.18			X						
	$PM_{2.5}$	0.08			X						
HUBRO	PM_{10}	0.62				X			X		
	$PM_{2.5}$	0.48				X					
SNAC-K	PM_{10}	0.81	X	X			X	X			
	$PM_{2.5}$	NA									
SALT	PM_{10}	0.81	X	X			X	X			
	$PM_{2.5}$	NA									
60-yr	PM_{10}	0.81	X	X			X	X			
	$PM_{2.5}$	NA									
SDPP	PM_{10}	0.81	X	X			X	X			
	$PM_{2.5}$	NA									
DCH	PM_{10}	0.55						X	X		
	$PM_{2.5}$	0.05		X							
EPIC-MORGEN,	PM_{10}	0.73		X					X		X
	$PM_{2.5}$	0.72							X		X
EPIC-PROSPECT,	PM_{10}	0.73		X					X		X
	$PM_{2.5}$	0.72							X		X
SALIA	PM_{10}	0.11		X							
	$PM_{2.5}$	0.2		X							
EPIC- Oxford	PM_{10}	0.43	X		X						
	$PM_{2.5}$	0.09			X						
KORA	PM_{10}	0.49	X	X							
	$PM_{2.5}$	0.32	X	X						X	
VHM&PP	PM_{10}	0.11	X								
	$PM_{2.5}$	0.1				X					
SAPALDIA,	PM_{10}	NA									
	$PM_{2.5}$	NA									
E3N	PM_{10}	0.54		X							
	$PM_{2.5}$	0.12		X							
EPIC- Turin	PM_{10}	0.79	X		X	X	X				
	$PM_{2.5}$	0.38		X							
SIDRIA-Turin	PM_{10}	0.79	X		X	X	X				
	$PM_{2.5}$	0.38		X							
SIDRIA-Rome	PM_{10}	0.86			X	X					
	$PM_{2.5}$	0.18			X						
EPIC- Athens	PM_{10}	0.47	X	X					X		
	$PM_{2.5}$	0.83			X				X		

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants
[d]RES includes low and high density residential
[e]Industry indicates industry area
[f]Green includes natural green and urban green variables
[g]Port indicates port area
[h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

Table S6 PM_{10} and $PM_{2.5}$ S model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	NA									
	$PM_{2.5}$	0.23						X			
HUBRO	PM_{10}	0.48	X			X		X			
	$PM_{2.5}$	0.07						X			
SNAC-K	PM_{10}	0.18		X							
	$PM_{2.5}$	0.24	X								
SALT	PM_{10}	0.18		X							
	$PM_{2.5}$	0.24	X								
60-yr	PM_{10}	0.18		X							
	$PM_{2.5}$	0.24	X								
SDPP	PM_{10}	0.18		X							
	$PM_{2.5}$	0.24	X								
DCH	PM_{10}	0.46		X				X			
	$PM_{2.5}$	0.61					X		X		
EPIC-MORGEN,	PM_{10}	0.39	X						X		X
	$PM_{2.5}$	0.27			X						X
EPIC-PROSPECT,	PM_{10}	0.39	X						X		X
	$PM_{2.5}$	0.27			X						X
SALIA	PM_{10}	0.55			X	X			X		
	$PM_{2.5}$	0.5	X		X						
EPIC- Oxford	PM_{10}	0.05						X			
	$PM_{2.5}$	0.02						X			
KORA	PM_{10}	0.37	X		X			X			
	$PM_{2.5}$	0.7	X	X				X		X	
VHM&PP	PM_{10}	0.55	X		X	X					
	$PM_{2.5}$	0.53	X							X	
SAPALDIA,	PM_{10}	NA									
	$PM_{2.5}$	NA									
E3N	PM_{10}	0.78	X	X			X	X			
	$PM_{2.5}$	0.29	X	X				X			
EPIC- Turin	PM_{10}	0.57	X		X					X	
	$PM_{2.5}$	0.39		X		X					
SIDRIA- Turin	PM_{10}	0.57	X		X					X	
	$PM_{2.5}$	0.39		X		X					
SIDRIA-Rome	PM_{10}	0.38			X	X					
	$PM_{2.5}$	0.33	X		X						
EPIC- Athens	PM_{10}	0.46								X	
	$PM_{2.5}$	0.67			X	X				X	

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants
[d]RES includes low and high density residential
[e]Industry indicates industry area
[f]Green includes natural green and urban green variables
[g]Port indicates port area
[h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

Table S7 PM_{10} and $PM_{2.5}$ Si model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	0.57	X		X						
	$PM_{2.5}$	0.72		X	X			X			
HUBRO	PM_{10}	0.82	X		X						
	$PM_{2.5}$	0.77	X			X		X		X	
SNAC-K	PM_{10}	0.77	X			X					
	$PM_{2.5}$	0.65	X			X					
SALT	PM_{10}	0.77	X			X					
	$PM_{2.5}$	0.65	X			X					
60-yr	PM_{10}	0.77	X			X					
	$PM_{2.5}$	0.65	X			X					
SDPP	PM_{10}	0.77	X			X					
	$PM_{2.5}$	0.65	X			X					
DCH	PM_{10}	0.6						X	X		
	$PM_{2.5}$	0.17		X							
EPIC-MORGEN,	PM_{10}	0.26		X	X						
	$PM_{2.5}$	0.39	X								X
EPIC-PROSPECT,	PM_{10}	0.26		X	X						
	$PM_{2.5}$	0.39	X								X
SALIA	PM_{10}	0.7	X		X				X		
	$PM_{2.5}$	0.4			X						
EPIC-Oxford	PM_{10}	0.49				X					
	$PM_{2.5}$	0.17				X					
KORA	PM_{10}	0.64	X	X		X					
	$PM_{2.5}$	0.55	X	X		X					
VHM&PP	PM_{10}	0.48	X	X							
	$PM_{2.5}$	0.15	X					X			
SAPALDIA,	PM_{10}	NA									
	$PM_{2.5}$	NA									
E3N	PM_{10}	0.61	X					X			
	$PM_{2.5}$	0.5	X		X		X				
EPIC- Turin	PM_{10}	0.61	X		X			X			
	$PM_{2.5}$	0.54	X		X		X				
SIDRIA-Turin	PM_{10}	0.61	X		X			X			
	$PM_{2.5}$	0.54	X		X		X				
SIDRIA-Rome	PM_{10}	0.6	X		X	X					
	$PM_{2.5}$	0.43	X		X						
EPIC-Athens	PM_{10}	0.39	X					X			
	$PM_{2.5}$	NA									

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants
[d]RES includes low and high density residential
[e]Industry indicates industry area
[f]Green includes natural green and urban green variables
[g]Port indicates port area
[h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

Table S8 PM_{10} and $PM_{2.5}$ V model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	0.16						X	X		
	$PM_{2.5}$	0.3			X						
HUBRO	PM_{10}	0.83	X		X	X					
	$PM_{2.5}$	0.42							X		
SNAC-K	PM_{10}	0.7	X			X	X				
	$PM_{2.5}$	0.07				X					
SALT	PM_{10}	0.7	X			X	X				
	$PM_{2.5}$	0.07				X					
60-yr	PM_{10}	0.7	X			X	X				
	$PM_{2.5}$	0.07				X					
SDPP	PM_{10}	0.7	X			X	X				
	$PM_{2.5}$	0.07				X					
DCH	PM_{10}	0.59							X		X
	$PM_{2.5}$	0.38							X		
EPIC-MORGEN,	PM_{10}	0.67							X		X
	$PM_{2.5}$	0.63							X		X
EPIC-PROSPECT,	PM_{10}	0.67							X		X
	$PM_{2.5}$	0.63							X		X
SALIA	PM_{10}	0.52				X			X		X
	$PM_{2.5}$	0.48	X						X		
EPIC-Oxford	PM_{10}	0.29			X						
	$PM_{2.5}$	0.24				X					
KORA	PM_{10}	0.04						X			
	$PM_{2.5}$	NA									
VHM&PP	PM_{10}	0.00						X			
	$PM_{2.5}$	NA									
SAPALDIA,	PM_{10}	NA									
	$PM_{2.5}$	NA									
E3N	PM_{10}	0.59	X					X			
	$PM_{2.5}$	0.28				X					
EPIC- Turin	PM_{10}	0.62	X			X					
	$PM_{2.5}$	0.36		X							
SIDRIA-Turin	PM_{10}	0.62	X			X					
	$PM_{2.5}$	0.36		X							
SIDRIA-Rome	PM_{10}	0.68	X			X				X	
	$PM_{2.5}$	0.09				X					
EPIC-Athens	PM_{10}	0.72					X		X		
	$PM_{2.5}$	0.92				X		X	X		

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants
[d]RES includes low and high density residential
[e]Industry indicates industry area
[f]Green includes natural green and urban green variables
[g]Port indicates port area
[h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

Table S9 PM_{10} and $PM_{2.5}$ Zn model details

Cohort	Type	LOOCV	[a]Traffic (≤100m)	[b]Traffic (>100m)	[c]POP	[d]RES	[e]Industry	[f]Green	[g]Port	[h]Alt	[i]XY
FINRISK	PM_{10}	NA									
	$PM_{2.5}$	NA									
HUBRO	PM_{10}	0.75	X					X			
	$PM_{2.5}$	0.7	X	X	X					X	
SNAC-K	PM_{10}	0.87	X				X				
	$PM_{2.5}$	0.24	X								
SALT	PM_{10}	0.87	X				X				
	$PM_{2.5}$	0.24	X								
60-yr	PM_{10}	0.87	X				X				
	$PM_{2.5}$	0.24	X								
SDPP	PM_{10}	0.87	X				X				
	$PM_{2.5}$	0.24	X								
DCH	PM_{10}	0.72	X					X			
	$PM_{2.5}$	0.11			X						
EPIC-MORGEN,	PM_{10}	0.57		X							X
	$PM_{2.5}$	0.58		X							X
EPIC-PROSPECT,	PM_{10}	0.57		X							X
	$PM_{2.5}$	0.58		X							X
SALIA	PM_{10}	0.5	X		X						
	$PM_{2.5}$	0.58	X			X	X	X	X		
EPIC-Oxford	PM_{10}	0.77	X	X		X					
	$PM_{2.5}$	0.63	X		X						
KORA	PM_{10}	0.65	X	X							
	$PM_{2.5}$	0.53	X	X			X	X			
VHM&PP	PM_{10}	0.79	X		X	X		X			
	$PM_{2.5}$	0.44		X		X		X			
SAPALDIA,	PM_{10}	0.86	X	X						X	
	$PM_{2.5}$	0.75	X			X					
E3N	PM_{10}	0.73	X	X							
	$PM_{2.5}$	0.76	X	X		X		X			
EPIC- Turin	PM_{10}	0.91	X		X	X		X			
	$PM_{2.5}$	0.78	X	X		X		X			
SIDRIA-Turin	PM_{10}	0.91	X		X	X		X			
	$PM_{2.5}$	0.78	X	X		X		X			
SIDRIA-Rome	PM_{10}	0.81	X			X					
	$PM_{2.5}$	0.75	X	X							X
EPIC-Athens	PM_{10}	0.73	X	X	X			X		X	
	$PM_{2.5}$	0.9		X			X	X	X		

[a]Traffic (≤100m) includes all traffic intensity and road length variables within 100m
[b]Traffic (>100m) includes all traffic intensity and road length variables beyond 100m
[c]POP indicates the number of inhabitants ; [d]RES includes low and high density residential
[e]Industry indicates industry area; [f]Green includes natural green and urban green variables
[g]Port indicates port area; [h]Alt: altitude of measured sites
[i]XY: coordinate variables which indicates the trends of concentrations

References

1. Eeftens, M.; Beelen, R.; de Hoogh, K.; Bellander, T.; Cesaroni, G.; Cirach, M.; Declercq, C.; 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*, (20), 11195-205.
2. Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; et al., Development of NO_2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. *Atmospheric Environment* **2013,** *72*, 10-23.
3. de Hoogh, K.; Wang, M.; Adam, M.; Badaloni, C.; Beelen, R.; Birk, M.; et al., Development of land use regression models for particle composition in twenty study areas in europe. *Environ Sci Technol* **2013,** *47*, (11), 5778-86.
4. Cyrys, J.; Eeftens, M.; Heinrich, J.; Ampe, C.; Armengaud, A.; Beelen, R.; et al., Variation of NO_2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. *Atmospheric Environment* **2012,** *62*, 374-390.
5. Eeftens, M.; Tsai, M. Y.; Ampe, C.; Anwander, B.; Beelen, R.; Bellander, T.; et al., Spatial variation of $PM_{2.5}$, PM_{10}, $PM_{2.5}$ absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO_2 - Results of the ESCAPE project. *Atmospheric Environment* **2012,** *62*, 303-317.
6. Wang, M.; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B., Systematic Evaluation of Land Use Regression Models for NO_2. *Environmental Science & Technology* **2012,** *46*, (8), 4481-4489.

Online supplement 3: Study population characteristics at baseline for each cohort (variables included in main model 3)

Table S1: Study population characteristics at baseline for HUBRO with complete confounder information in main model 3 (N = 18,234)

Variable	Mean (SD)
Age at baseline	48.3 (15.2)
Number of cigarette equivalents/day (lifetime average)	6.7 (8.4)
Years of regular smoking	11.5 (14.4)
BMI (kg/m^2)	25.7 (4.1)
Unemployment rate (municipality level) (%)	1.8 (0.7)
	N (%)
Gender	
- Women	10,236 (56.1%)
- Men	7998 (43.9%)
Calendar year	
- 2000	7928 (43.5%)
- 2001	10,306 (56.5%)
Smoking status	
- Current	4752 (26.1%)
- Former	5094 (27.9%)
- Never	8388 (46.0%)
Alcohol consumption	
- Weekly	9228 (50.6%)
- Occasionally	7358 (40.4%)
- Never/not past year	1648 (9.0%)
Intake of fruit	
- Daily	7284 (39.9%)
- Weekly	8881 (48.7%)
- Rarely	2069 (11.3%)
Intake of vegetables	
- Daily	2646 (14.5%)
- Weekly	12,503 (68.6%)
- Rarely	3085 (16.9%)
Marital status	
- Single	5645 (31.0%)
- Married/living with partner	9089 (49.8%)
- Divorced/separated	2474 (13.6%)
- Widowed	1026 (5.6%)
Educational level	
- Low	3234 (17.7%)
- Medium	6597 (36.2%)
- High	8403 (46.1%)

Table S2: Study population characteristics at baseline for DCH with complete confounder information in main model 3 (N = 35,458)

Variable	Mean (SD)
Age at baseline	56.7 (4.4)
Number of cigarette equivalents/day (current)	6.3 (10.4)
Years of regular smoking	18.7 (17.1)
Intake of fruit (g/day)	183.2 (151.2)
Intake of vegetables (g/day)	175.9 (99.2)
Alcohol consumption (g/day)	21.7 (22.8)
BMI (kg/m^2)	26.0 (4.1)
Average income (municipality) (100,000 Dkr)	1.9 (0.4)
	N (%)
Gender	
- Women	19,171 (54.1%)
- Men	16,287 (45.9%)
Calendar year	
- 1993	86 (0.2%)
- 1994	3712 (10.5%)
- 1995	11,034 (31.1%)
- 1996	14,726 (41.5%)
- 1997	5900 (16.6%)
Smoking status	
- Current	12,737 (35.9%)
- Former	9851 (27.8%)
- Never	12,870 (36.3%)
Marital status	
- Single	2317 (6.5%)
- Married/living with partner	24,544 (69.2%)
- Divorced/separated	6539 (18.4%)
- Widowed	2058 (5.8%)
Educational level	
- Low	10,490 (29.6%)
- Medium	16,844 (47.5%)
- High	8124 (22.9%)
Environmental tobacco smoke at work and/or home	
- No	12,654 (35.7%)
- Yes	22,804 (64.3%)
Employment status	
- Not employment	7073 (19.9%)
- Employed	28,385 (80.1%)

Table S3: Study population characteristics at baseline for FINRISK with complete confounder information in main model 3 (N = 10,224)

Variable	Mean (SD)
Age at baseline	47.9 (13.2)
Number of cigarette equivalents/day (current)	3.8 (7.8)
Years of regular smoking	8.6 (12.2)
Alcohol consumption[a]	0.9 (1.3)
BMI (kg/m^2)	26.4 (4.6)
Average income (3km) (EUR)	22,954 (5459)
	N (%)
Gender	
- Women	5501 (53.8%)
- Men	4723 (46.2%)
Calendar year	
- 1992	2783 (27.2%)
- 1997	2941 (28.8%)
- 2002	2418 (23.7%)
- 2007	2082 (20.4%)
Smoking status	
- Current	2638 (25.8%)
- Former	2947 (28.8%)
- Never	4639 (45.4%)
Marital status	
- Single	1611 (15.8%)
- Married/living with partner	7170 (70.1%)
- Divorced/separated	1100 (10.8%)
- Widowed	343 (3.4%)
Educational level	
- Low	3167 (31.0%)
- Medium	5291 (51.8%)
- High	1766 (17.3%)
Environmental tobacco smoke at work and/or home	
- No	8322 (81.4%)
- Yes	1902 (18.6%)
Intake of fruit	
- Daily	6783 (66.3%)
- Weekly	2639 (25.8%)
- Seldom	592 (5.8%)
- Never	210 (2.1%)
Intake of vegetables	
- Daily	6973 (68.2%)
- Weekly	2550 (24.9%)
- Seldom	488 (4.8%)
- Never	213 (2.1%)
Occupational class	
- Blue collar	1528 (14.9%)
- White collar	5435 (53.2%)
- Students/housewives/retired/ unemployed	3261 (31.9%)
Employment status	
- Employed/Self-employed	7073 (69.2%)
- Unemployed	621 (6.1%)
- Homemaker/housewife	347 (3.4%)
- Retired	2183 (21.4%)
Area indicator	
- Helsinki and Vantaa	4935 (48.3%)
- Turku area	5289 (51.7%)

[a] Number of glasses of alcoholic drink during last week

Table S4: Study population characteristics at baseline for SNAC-K with complete confounder information in main model 3 (N = 2401)

Variable	**Mean (SD)**
Age at baseline	70.3 (8.1)
Number of cigarette equivalents/day (lifetime average)	7.1 (9.5)
Years of regular smoking	9.8 (15.2)
BMI (kg/m^2)	26.0 (4.1)
Average income (neighborhood) (SEK)	352,638 (26,928)
	N (%)
Gender	
- Women	1441 (60.0%)
- Men	960 (40.0%)
Calendar year	
- 2001	512 (21.3%)
- 2002	691 (28.8%)
- 2003	798 (33.2%)
- 2004	400 (16.7%)
Smoking status	
- Current	378 (15.7%)
- Former	960 (40.0%)
- Never	1063 (44.3%)
Marital status	
- Single	305 (12.7%)
- Married/living with partner	1301 (54.2%)
- Divorced/separated	364 (15.2%)
- Widowed	431 (18.0%)
Educational level	
- Low	509 (21.2%)
- Medium	1039 (43.3%)
- High	853 (35.5%)
Environmental tobacco smoke at work	
- No	810 (33.7%)
- Yes	1591 (66.3%)
Environmental tobacco smoke at home	
- No	
- Yes	1094 (45.6%)
	1307 (54.4%)
Occupation class	
- Blue collar	387 (16.1%)
- White collar	2014 (83.9%)
Employment status	
- Other	1714 (71.4%)
- Employed	687 (28.6%)
Alcohol consumption	
- Daily	524 (21.8%)
- Weekly	643 (26.8%)
- Seldom	1060 (44.1%)
- Never	174 (7.2%)

Table S5: Study population characteristics at baseline for SALIA with complete confounder information in main model 3 (N = 4352)

Variable	Mean (SD)
Age at baseline	54.5 (0.6)
Number of cigarette equivalents/day (current)	2.6 (6.6)
Years of regular smoking	4.4 (10.5)
Average income (postal code area) (EUR)	973.6 (69.1)
	N (%)
Gender	
- Women	4352 (100%)
- Men	0 (0%)
Calendar year	
- 1985-1987	1667 (38.3%)
- 1990-1994	2685 (61.7%)
Smoking status	
- Current	729 (16.7%)
- Former	379 (8.7%)
- Never	3244 (74.5%)
Educational level	
- Low	1255 (28.8%)
- Medium	2094 (48.1%)
- High	1003 (23.0%)
Environmental tobacco smoke at work and/or home	
- No	2141 (49.2%)
- Yes	2211 (50.8%)
Occupational exposure to dust	
- No	3923 (90.1%)
- Yes	429 (9.9%)

Table S6: Study population characteristics at baseline for SALT / Twin gene with complete confounder information in main model 3 (N = 5473)

Variable	Mean (SD)
Age at baseline	58.0 (9.9)
Number of cigarette equivalents/day (lifetime average)	8.5 (9.7)
Years of regular smoking	16.7 (17.3)
BMI (kg/m^2)	28.6 (4.1)
	N (%)
Gender	
- Women	3050 (55.7%)
- Men	2423 (44.3%)
Calendar year	
- 1998	262 (4.8%)
- 1999	1467 (26.8%)
- 2000	1410 (25.8%)
- 2001	1177 (21.5%)
- 2002	1157 (21.1%)
Smoking status	
- Current	1295 (23.7%)
- Former	2059 (37.6%)
- Never	2119 (38.7%)
Marital status	
- Single	784 (14.3%)
- Married/living with partner	3723 (68.0%)
- Divorced/separated	612 (11.2%)
- Widowed	354 (6.5%)
Educational level	
- Low	1179 (21.5%)
- Medium	2360 (43.1%)
- High	1934 (35.3%)
Individual level socioeconomic status	
- Low	1643 (30.0%)
- Medium	2842 (51.9%)
- High	988 (18.1%)
Mean income (municipality level) (SEK)	
- Quartile 1	1528 (27.9%)
- Quartile 2	2366 (43.2%)
- Quartile 3	221 (4.0%)
- Quartile 4	1358 (24.8%)

Table S7: Study population characteristics at baseline for SDPP with complete confounder information in main model 3 (N = 7408)

Variable	Mean (SD)
Age at baseline	47.1 (5.0)
Number of cigarette equivalents/day (lifetime average)	8.5 (8.8)
Years of regular smoking	12.3 (12.4)
Alcohol consumption[a]	1.3 (1.9)
BMI (kg/m^2)	25.6 (4.0)
Average income (municipality) (SEK)	277,069 (18,711)
	N (%)
Gender	
- Women	4570 (61.7%)
- Men	2838 (38.3%)
Calendar year	
- 1992	292 (3.9%)
- 1993	1741 (23.5%)
- 1994	805 (10.9%)
- 1996	1815 (24.5%)
- 1997	2378 (32.1%)
- 1998	377 (5.1%)
Smoking status	
- Current	1928 (26.0%)
- Former	2711 (36.6%)
- Never	2769 (37.4%)
Marital status	
- Single/living alone	1217 (16.4%)
- Married/living with partner	6191 (83.6%)
Educational level	
- Low	1892 (25.5%)
- Medium	3321 (44.8%)
- High	2195 (29.6%)
Occupation class	
- Worker/blue collar	2451 (33.1%)
- White collar	4957 (66.9%)
Employment status	
- Not employment	606 (8.2%)
- Employed	6802 (91.8%)
Intake of fruit	
- Daily/weekly	6845 (92.4%)
- Seldom	482 (6.5%)
- Never	81 (1.1%)

[a] Number of glasses of alcoholic drink per day.

Table S8: Study population characteristics at baseline for 60-yr/IMPROVE with complete confounder information in main model 3 (N = 3612)

Variable	**Mean (SD)**
Age at baseline	60.4 (0.1)
Number of cigarette equivalents/day (lifetime average)	8.0 (9.1)
Years of regular smoking	15.2 (16.4)
Alcohol consumption (g/day)	8.9 (9.7)
BMI (kg/m^2)	26.8 (4.2)
Average income (municipality) (SEK)	290,838 (46,103)
	N (%)
Gender	
- Women	1897 (52.5%)
- Men	1715 (47.5%)
Calendar year	
- 1997	757 (21.0%)
- 1998	2772 (76.7%)
- 1999	83 (2.3%)
Smoking status	
- Current	761 (21.0%)
- Former	1371 (38.0%)
- Never	1480 (41.0%)
Environmental tobacco smoke at work and/or home	
- No	1898 (52.5%)
- Yes	1714 (47.5%)
Marital status	
- Single	161 (4.5%)
- Married/living with partner	2587 (71.6%)
- Divorced/separated	617 (17.1%)
- Widowed	247 (6.8%)
Educational level	
- Low	995 (27.5%)
- Medium	1596 (44.2%)
- High	1021 (28.3%)
Occupation class	
- Blue collar	820 (22.7%)
- Low white collar	1977 (54.7%)
- High white	815 (22.6%)
Employment status	
- Employed/Self-employed	1857 (51.4%)
- Unemployed	351 (9.7%)
- Homemaker/housewife	276 (7.6%)
- Retired	1128 (31.2%)
Intake of fruit	
- Daily	2318 (64.2%)
- Weekly	1015 (28.1%)
- Seldom/never	279 (1.7%)
Intake of vegetables	
- Daily	476 (13.2%)
- Weekly	3085 (85.4%)
- Seldom/never	51 (1.4%)

Table S9: Study population characteristics at baseline for E3N with complete confounder information in main model 3 (N = 14,313)

Variable	Mean (SD)
Age at baseline	53.0 (6.7)
Intake of fruit (g/day)	242.0 (164.7)
Intake of vegetables (g/day)	242.0 (126.8)
Alcohol consumption (g/day)	12.0 (15.1)
BMI (kg/m^2)	22.8 (3.2)
Unemployment rate (regional scale)	9.4 (1.0)
	N (%)
Gender	
- Women	14,313 (100%)
- Men	0 (0%)
Calendar year	
- 1993	10,751 (75.1%)
- 1994	2257 (15.8%)
- 1995	917 (6.4%)
- 1996	388 (2.7%)
Smoking status	
- Current	2364 (16.5%)
- Former	4886 (34.1%)
- Never	7063 (49.3%)
Educational level	
- Low	710 (5.0%)
- Medium	856 (6.0%)
- High	12,747 (89.0%)

Table S10: Study population characteristics at baseline for EPIC-PROSPECT with complete confounder information in main model 3 (N = 15,670)

Variable	Mean (SD)
Age at baseline	57.7 (6.0)
Number of cigarette equivalents/day (lifetime average)	5.7 (7.4)
Years of regular smoking	15.2 (16.5)
Intake of fruit (g/day)	231.6 (139.2)
Intake of vegetables (g/day)	136.3 (52.5)
Alcohol consumption (g/day)	9.0 (12.4)
BMI (kg/m^2)	25.5 (4.1)
Percentage of people with low income (municipality)	35.9 (2.7)
Percentage of people with low income (neighborhood)	35.8 (7.2)
	N (%)
Gender	
- Women	15,670 (100%)
- Men	0 (0%)
Calendar year	
- 1993	1354 (8.6%)
- 1994	4071 (26.0%)
- 1995	4023 (25.7%)
- 1996	4102 (26.2%)
- 1997	2120 (13.5%)
Smoking status	
- Current	3454 (22.0%)
- Former	5166 (33.0%)
- Never	7050 (45.0%)
Marital status	
- Single	888 (5.7%)
- Married/living with partner	12,046 (76.9%)
- Divorced/separated	1252 (8.0%)
- Widowed	1484 (9.5%)
Educational level	
- Low	3478 (22.2%)
- Medium	9685 (61.8%)
- High	2507 (16.0%)

Table S11: Study population characteristics at baseline for EPIC-Oxford with complete confounder information in main model 3 (N = 38,941)

Variable	Mean (SD)
Age at baseline	45.8 (13.7)
Number of cigarette equivalents/day (lifetime average)	5.0 (8.3)
Years of regular smoking	6.7 (11.2)
Intake of fruit (g/day)	259.9 (204.5)
Intake of vegetables (g/day)	281.0 (156.4)
Alcohol consumption (g/day)	9.1 (11.7)
BMI (kg/m^2)	24.0 (3.9)
Carstairs index 2001 (continuous)	-1.5 (2.3)
	N (%)
Gender	
- Women	30,178 (77.5%)
- Men	8763 (22.5%)
Calendar year	
- 1993	311 (0.8%)
- 1994	5345 (13.7%)
- 1995	7009 (18.0%)
- 1996	13,399 (34.4%)
- 1997	7854 (20.2%)
- 1998-2001	5023 (12.9%)
Smoking status	
- Current	4016 (10.3%)
- Former	10,294 (26.4%)
- Never	24,631 (63.3%)
Marital status	
- Single	6336 (16.3%)
- Married/living with partner	27,554 (70.8%)
- Divorced/separated	3474 (8.9%)
- Widowed	1577 (4.0%)
Educational level	
- Low	14,194 (36.5%)
- Medium	9391 (24.1%)
- High	15,356 (3.9%)
Employment status	
- Employed/self-employed	28,230 (72.5%)
- Unemployed	958 (2.5%)
- Stay at home	4593 (11.8%)
- Retired	5160 (13.3%)

Table S12: Study population characteristics at baseline for SAPALDIA with complete confounder information in main model 3 (N = 3473)

Variable	Mean (SD)
Age at baseline	41.1 (11.8)
Number of cigarette equivalents/day (lifetime average)	11.5 (14.5)
Years of regular smoking	10.7 (12.4)
BMI (kg/m^2)	23.7 (4.0)
Average educational level (neighborhood)[a]	3.2 (0.3)
	N (%)
Gender	
- Women	1807 (52.0%)
- Men	1666 (48.0%)
Calendar year	
- 1991	3473 (100%)
Smoking status	
- Current	1259 (36.3%)
- Former	740 (21.3%)
- Never	1474 (42.4%)
Marital status	
- Single	1214 (35.0%)
- Married/living with partner	1885 (54.3%)
- Divorced/separated	305 (8.8%)
- Widowed	69 (2.0%)
Educational level	
- Low	522 (15.0%)
- Medium	2222 (64.0%)
- High	729 (21.0%)
Environmental tobacco smoke at home	
- No	3000 (86.4%)
- Yes	473 (13.6%)
Environmental tobacco smoke at work	
- No	3163 (91.1%)
- Yes	310 (8.9%)
Employment status	
- Employed	2931 (84.4%)
- Unemployed	54 (1.6%)
- Stay at home or retired	488 (14.0%)

[a] Average of a 7-categories (1-7) level of education variable, calculated for participants living within the same neighborhood zone

Table S13: Study population characteristics at baseline for EPIC-MORGEN with complete confounder information in main model 3 (N = 16,446)

Variable	Mean (SD)
Age at baseline	43.9 (10.9)
Number of cigarette equivalents/day (lifetime average)	10.4 (11.1)
Years of regular smoking	14.3 (13.7)
Intake of fruit (g/day)	171.9 (129.2)
Intake of vegetables (g/day)	126.6 (51.8)
Alcohol consumption (g/day)	12.7 (18.0)
BMI (kg/m^2)	25.2 (4.0)
Percentage of people with low income (neighborhood)	41.6 (7.4)
	N (%)
Gender	
- Women	8946 (54.4%)
- Men	7500 (45.6%)
Calendar year	
- 1993	3566 (21.7%)
- 1994	2948 (17.9%)
- 1995	3568 (21.7%)
- 1996	3365 (20.5%)
- 1997	2999 (18.2%)
Smoking status	
- Current	5923 (36.0%)
- Former	4762 (29.0%)
- Never	5761 (35.0%)
Marital status	
- Single	3669 (22.3%)
- Married/living with partner	11,118 (67.6%)
- Divorced/separated	1311 (8.0%)
- Widowed	348 (2.1%)
Educational level	
- Low	1954 (11.9%)
- Medium	10,752 (65.4%)
- High	3740 (22.7%)

Table S14: Study population characteristics at baseline for KORA with complete confounder information in main model 3 (N = 8399)

Variable	Mean (SD)
Age at baseline	49.5 (13.8)
Number of cigarette equivalents/day (lifetime average)	9.2 (13.3)
Years of regular smoking	12.0 (14.2)
Alcohol consumption (g/day)	16.3 (22.3)
BMI (kg/m^2)	27.2 (4.6)
Percentage of people with low income (5km grid)	28.2 (18.4)
	N (%)
Gender	
- Women	4270 (50.8%)
- Men	4129 (49.2%)
Calendar year	
- 1994-1995	4299 (51.2%)
- 1999-2001	4100 (48.8%)
Smoking status	
- Current	2183 (26.0%)
- Former	2546 (30.3%)
- Never	3670 (43.7%)
Marital status	
- Single	872 (10.4%)
- Married/living with partner	6356 (75.7%)
- Divorced/separated	635 (7.6%)
- Widowed	536 (6.4%)
Educational level	
- Low	1059 (12.6%)
- Medium	6270 (74.7%)
- High	1070 (12.7%)
Environmental tobacco smoke at home	
- No	6390 (76.1%)
- Yes	2009 (23.9%)
Environmental tobacco smoke at work	
- No	6328 (75.3%)
- Yes	2071 (24.7%)
Employment status	
- Employed/self-employed	4894 (58.3%)
- Unemployed	273 (3.3%)
- Stay at home	1170 (13.9%)
- Retired	2062 (24.6%)
Intake of fruit	
- Daily	4995 (59.5%)
- Weekly	2547 (30.3%)
- Seldom/never	857 (10.2%)
Intake of vegetables	
- Daily	3953 (47.1%)
- Weekly	3821 (45.5%)
- Seldom/never	625 (7.4%)

Table S15: Study population characteristics at baseline for VHM&PP with complete confounder information in main model 3 (N = 117,824)

Variable	Mean (SD)
Age at baseline	41.9 (14.9)
BMI (kg/m^2)	24.8 (4.3)
Average income (municipality) (EUR)	25,119 (1273)
	N (%)
Gender	
- Women	66,042 (56.1%)
- Men	51,782 (43.9%)
Calendar year	
- 1985-1989	58,490 (49.6%)
- 1990-1994	26,393 (22.4%)
- 1995-1999	18,414 (15.6%)
- 2000-2005	14,527 (12.3%)
Smoking status	
- Current	28,255 (24.0%)
- Former	7233 (6.1%)
- Never	82,336 (69.9%)
Marital status	
- Single	20,134 (17.1%)
- Married/living with partner	80,572 (68.4%)
- Divorced/separated	8962 (7.6%)
- Widowed	8156 (6.9%)
Occupational class	
- White collar	66,348 (56.3%)
- Blue collar	40,961 (34.8%)
- Others (mainly self-employed)	10,515 (8.9%)
Employment status	
- Employed/self-employed	81,705 (69.3%)
- Unemployed	4126 (3.5%)
- Retired	31,993 (27.2%)

Table S16: Study population characteristics at baseline for EPIC-Turin with complete confounder information in main model 3 (N = 7261)

Variable	Mean (SD)
Age at baseline	50.4 (7.5)
Number of cigarette equivalents/day (lifetime average)	7.2 (8.2)
Years of regular smoking	17.6 (16.3)
Intake of fruit (g/day)	318.2 (182.2)
Intake of vegetables (g/day)	181.8 (100.2)
Alcohol consumption (g/day)	18.1 (20.3)
BMI (kg/m^2)	25.3 (3.8)
	N (%)
Gender	
- Women	3,461 (47.7%)
- Men	3,800 (52.3%)
Calendar year	
- 1993	457 (6.3%)
- 1994	1264 (17.4%)
- 1995	2318 (31.9%)
- 1996	1541 (21.2%)
- 1997	1432 (19.7%)
- 1998	251 (3.5%)
Smoking status	
- Current	1830 (25.2%)
- Former	2339 (32.2%)
- Never	3092 (42.6%)
Marital status	
- Not married (single, widowed, separated, divorced)	1045 (14.4%)
- Married	6216 (85.6%)
Educational level	
- Low	3168 (43.6%)
- Medium	3104 (42.7%)
- High	989 (13.6%)
Deprivation index (quintiles) (census block)	
- I (less deprived)	1876 (25.8%)
- II	1659 (22.8%)
- III	1350 (18.6%)
- IV	1411 (19.4%)
- V (more deprived)	965 (13.3%)

Table S17: Study population characteristics at baseline for SIDRIA-Turin with complete confounder information in main model 3 (N = 5054)

Variable	Mean (SD)
Age at baseline	44.2 (6.2)
Number of cigarette equivalents/day (current)	9.3 (10.2)
Years of regular smoking	11.3 (10.6)
	N (%)
Gender	
- Women	2620 (51.8%)
- Men	2434 (48.2%)
Calendar year	
- 1999	5054 (100%)
Smoking status	
- Current	2110 (41.7%)
- Former	1047 (20.7%)
- Never	1897 (37.5%)
Marital status	
- Married/living with partner	4820 (95.4%)
- Single/divorced/separated/ widowed	234 (4.6%)
Educational level	
- Low	884 (17.5%)
- Medium	3604 (71.3%)
- High	566 (11.2%)
Environmental tobacco smoke at home	
- No	4389 (86.8%)
- Yes	665 (13.2%)
Occupational class	
- Blue collar	2120 (41.9%)
- White collar	1529 (30.3%)
- Other	1405 (27.8%)
Employment status	
- Employed	3649 (72.2%)
- Unemployed	351 (6.9%)
- Homemaker/housewife/retired	1054 (20.9%)
Deprivation index (quintiles) (census block)	
- I (less deprived)	878 (17.4%)
- II	1049 (20.8%)
- III	931 (18.4%)
- IV	1097 (21.7%)
- V (more deprived)	1099 (21.7%)

Table S18: Study population characteristics at baseline for SIDRIA-Rome with complete confounder information in main model 3 (N = 9177)

Variable	Mean (SD)
Age at baseline	44.3 (6.0)
Number of cigarette equivalents/day (current)	10.1 (10.5)
Years of regular smoking	11.7 (10.4)
	N (%)
Gender	
- Women	4848 (52.8%)
- Men	4329 (47.2%)
Calendar year	
- 1999	9177 (100%)
Smoking status	
- Current	3898 (42.5%)
- Former	2106 (22.9%)
- Never	3173 (34.6%)
Marital status	
- Married/living with partner	9177 (100%)
Educational level	
- Low	4121 (44.9%)
- Medium	3681 (40.1%)
- High	1375 (15.0%)
Occupation class	
- Non-manual	4783 (52.1%)
- Manual	1179 (12.8%)
- Worker unspecified	521 (5.7%)
- Unemployed	392 (4.3%)
- Housewife	2302 (25.1%)
Index of socioeconomic position (census block)	
- 1 (=High)	1703 (18.6%)
- 2	1684 (18.4%)
- 3	1667 (18.2%)
- 4	1797 (19.6%)
- 5 (=Low)	2326 (25.3%)

Table S19: Study population characteristics at baseline for EPIC-Athens with complete confounder information in main model 3 (N = 4192)

Variable	Mean (SD)
Age at baseline	49.4 (11.7)
Number of cigarette equivalents/day (lifetime average)	1.7 (15.0)
Years of regular smoking	10.8 (13.1)
Intake of fruit (g/day)	402.6 (258.2)
Intake of vegetables (g/day)	609.5 (288.6)
Alcohol consumption (g/day)	9.2 (14.5)
BMI (kg/m^2)	27.5 (4.5)
	N (%)
Gender	
- Women	2306 (55.0%)
- Men	1886 (45.0%)
Calendar year	
- 1994	1582 (37.7%)
- 1995	1100 (26.2%)
- 1996	367 (8.8%)
- 1997	457 (10.9%)
- 1998	278 (6.6%)
- 1999	408 (9.7%)
Smoking status	
- Current	1707 (40.7%)
- Former	830 (19.8%)
- Never	1655 (39.5%)
Marital status	
- Single	394 (9.4%)
- Married/living with partner	3270 (78.0%)
- Divorced/separated	266 (6.3%)
- Widowed	262 (6.3%)
Educational level	
- Low	990 (23.6%)
- Medium	1753 (41.8%)
- High	1449 (34.6%)
Occupation class	
- Blue collar	493 (11.8%)
- White collar	1990 (47.5%)
- Other	1709 (40.8%)
Employment status	
- Employed/self-employed	2804 (66.9%)
- Unemployed	28 (0.7%)
- Homemaker/housewife	669 (16.0%)
- Retired	691 (16.5%)
Educational level (municipality level)	
- 1: Low (primary)	214 (5.1%)
- 2: Medium (secondary)	3277 (78.2%)
- 3: High (higher)	701 (16.7%)

Online supplement 4: Description of exposure to particle components concentrations at participant addresses in each cohort

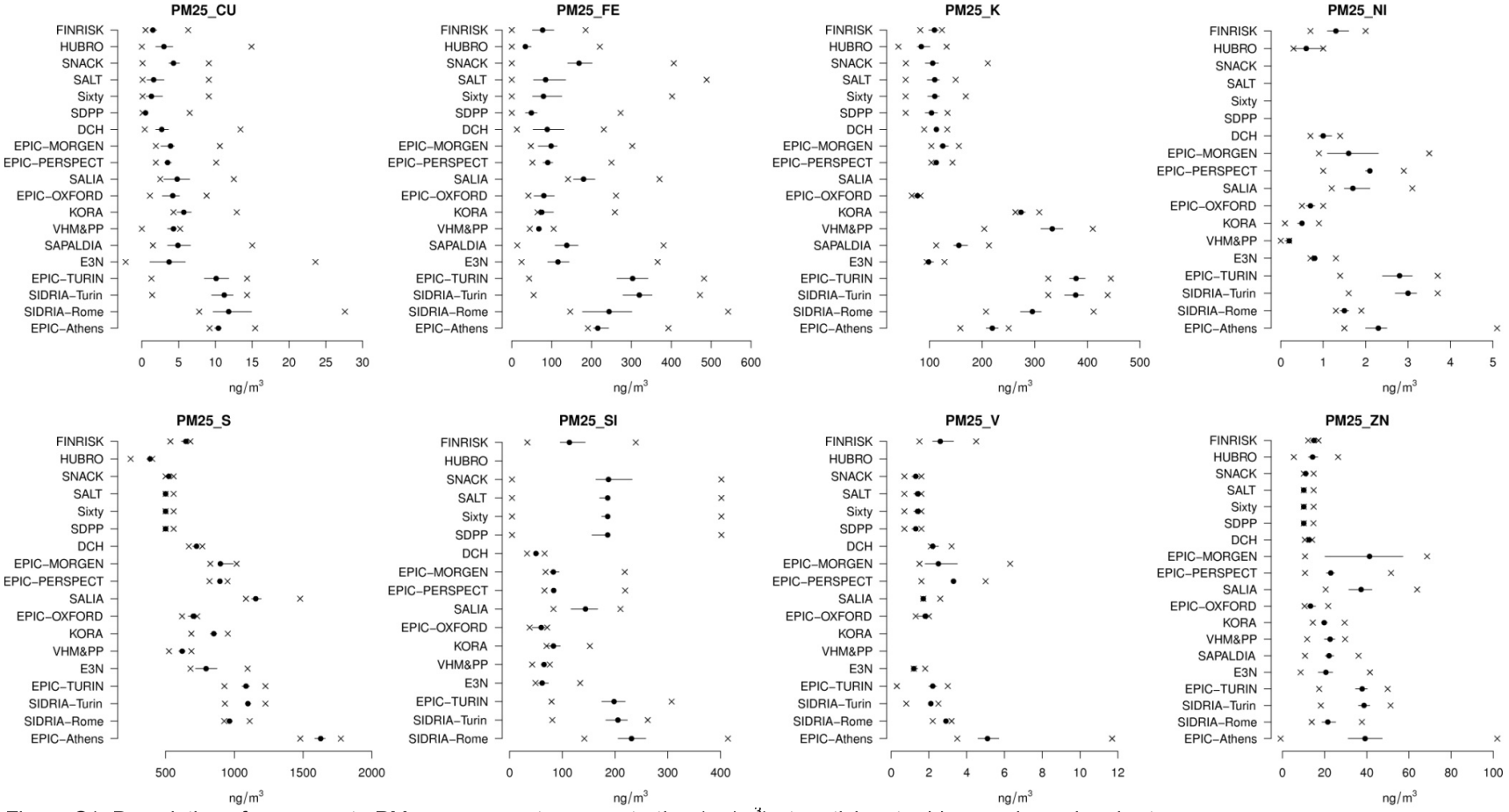


Figure S1: Description of exposure to $PM_{2.5}$ components concentration (ng/m^3) at participant addresses in each cohort.
The solid circle and bars shows the median and 25%, 75% percentile of elemental constituents of $PM_{2.5}$ and PM_{10}; the x shows the minimum and maximum values.

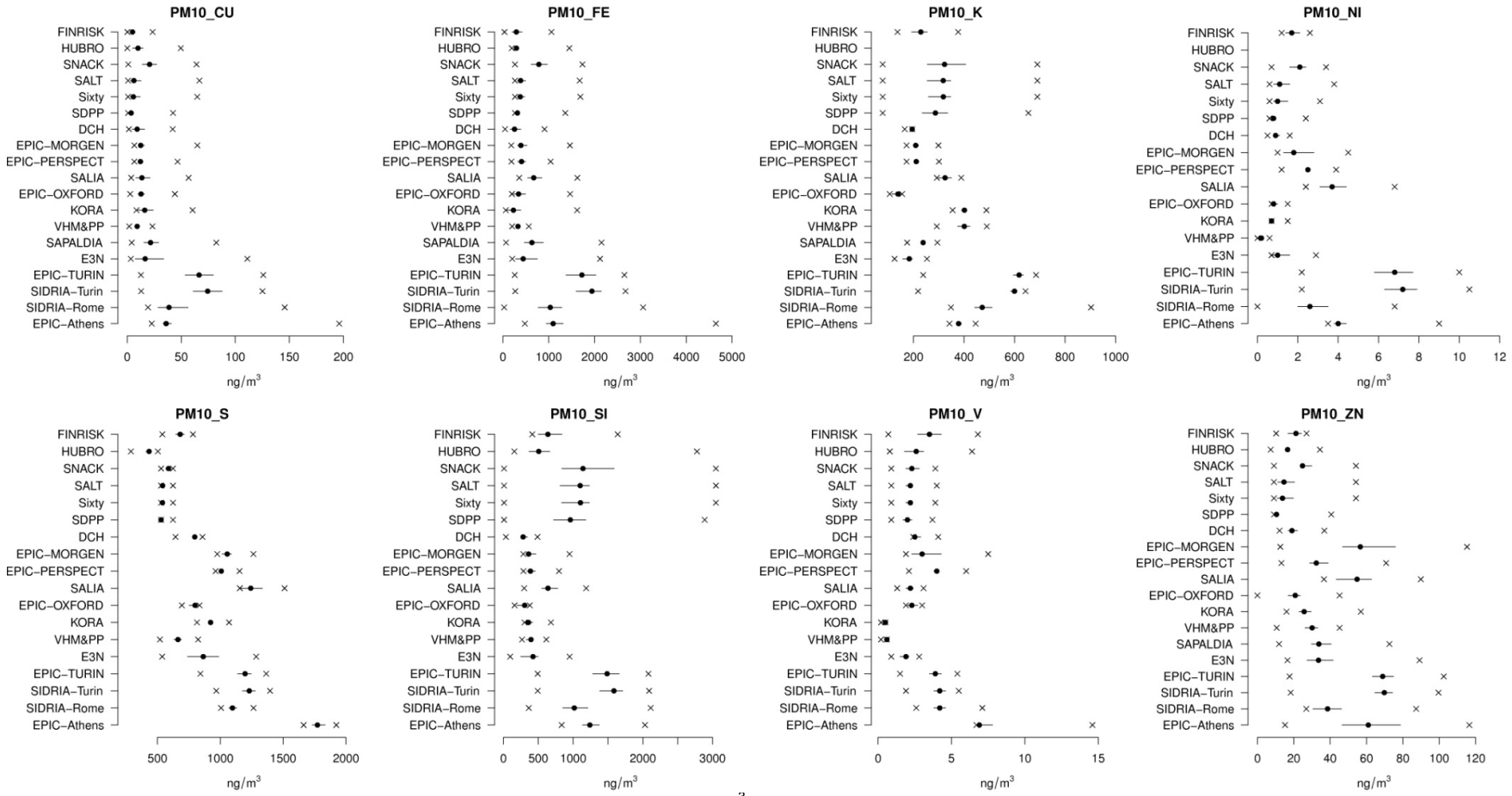


Figure S2: Description of exposure to PM_{10} components concentration (ng/m^3) at participant addresses in each cohort.
The solid circle and bars shows the median and 25%, 75% percentile of elemental constituents of $PM_{2.5}$ and PM_{10}; the x shows the minimum and maximum values.

Online supplement 5: Correlations between elemental constituent in $PM_{2.5}$ and PM_{10} at participant addresses in each cohort

Table S1 Correlations between PM constituents and corresponding $PM_{2.5}$ or PM_{10} mass concentrations

	Cu		Fe		K		Ni		S		Si		V		Zn	
Cohorts	PM2.5	PM10	PM2.5	PM10	PM2.5	PM10	PM2.5	PM10	PM2.5	PM10	PM2.5	PM10	PM2.5	PM10	PM2.5	PM10
FINRISK	0.31	0.74	0.48	0.71	0.97	0.71	0.17	0.74	0.86	0.79	0.19	0.81	0.09	0.61	0.81	0.60
SNAC-K	0.57	0.12	0.51	0.85	0.44	1.00		-0.18	0.65	0.28	0.59	1.00	0.41	0.70	0.59	0.83
SALT	0.48	0.22	0.45	0.71	0.36	1.00		0.00	0.51	0.30	0.50	1.00	0.34	0.73	0.45	0.64
Sixty	0.54	0.31	0.51	0.76	0.41	1.00		0.11	0.53	0.37	0.56	1.00	0.36	0.76	0.49	0.68
SDPP	0.41	0.17	0.45	0.45	0.26	1.00		0.17	0.26	0.23	0.22	1.00	0.32	0.93	0.27	0.45
DCH	0.54	0.72	0.56	0.75	0.43	0.30	0.58	0.72	0.52	0.68	0.58	0.76	0.57	0.67	0.66	0.69
EPIC_MORGEN	0.69	0.76	0.48	0.92	0.48	0.78	0.02	0.85	0.55	0.41	0.54	0.94	0.00	0.79	0.40	-0.37
EPIC_PROSPECT	0.72	0.47	0.47	0.81	0.44	0.74	-0.08	0.56	0.27	0.29	0.32	0.84	-0.06	0.20	0.42	0.60
SALIA	0.86	0.79	0.72	0.93		0.49	0.76	0.89	0.70	0.72	0.32	0.47	0.45	0.11	0.61	0.43
EPIC_Oxford	0.67	0.53	0.68	0.45	0.61	0.37	0.77	0.45	0.37	0.16	0.61	0.37	0.61	0.39	0.69	0.53
KORA	0.31	0.47	0.44	0.57	0.42	0.35	0.41	0.38	0.39	0.46	0.50	0.47		0.00	0.64	0.66
VHM_PP	0.58	0.61	0.31	0.47	0.87	0.63	0.39	0.57	0.58	0.83	0.05	0.37		0.03	0.33	0.69
SAPALDIA	0.73	0.67	0.85	0.91	0.89	0.95									0.57	0.91
SIDRIA_Turin	0.32	0.69	0.59	0.82	-0.03	0.30	0.30	0.73	0.33	0.69	0.19	0.65	0.33	0.48	0.40	0.65
SIDRIA_Rome	0.71	0.74	0.70	0.76	0.50	0.60	0.24	0.14	0.67	0.18	0.53	0.38	0.08	0.56	0.63	0.72
EPIC_Athens	0.36	0.23	0.56	0.22	0.27	0.30	0.20	0.03	0.36	0.17		0.09	0.20	-0.02	0.31	0.09

Online supplement 6: Sensitivity analysis

Table S1 Association[a] between CVD mortality and exposure to PM constituents: Results from random-effects meta-analyses (HRs and 95% CIs) (using additional hypertension confounder)

Exposure		Cohorts	Model3+ prevalent hypertension and physical activity[b]	Model3[b]
Cu	$PM_{2.5}$	18	0.90(0.75-1.09)	0.90(0.76-1.07)
	PM_{10}	18	0.92(0.80-1.07)	0.93(0.82-1.06)
Fe	$PM_{2.5}$	18	0.97(0.85-1.12)	0.98(0.87-1.11)
	PM_{10}	18	0.97(0.84-1.12)	0.95(0.84-1.09)
K	$PM_{2.5}$	17	0.97(0.94-1.01)	0.98(0.94-1.02)
	PM_{10}	17	0.99(0.93-1.05)	1.00(0.93-1.08)
Ni	$PM_{2.5}$	14	0.95(0.72-1.27)	0.97(0.78-1.21)
	PM_{10}	17	1.00(0.82-1.21)	1.01(0.88-1.16)
S	$PM_{2.5}$	18	1.05(0.93-1.20)	1.08(0.95-1.22)
	PM_{10}	18	1.07(0.89-1.28)	1.09(0.90-1.32)
Si	$PM_{2.5}$	16	1.15(0.90-1.47)	1.17(0.93-1.47)
	PM_{10}	18	1.01(0.90-1.13)	1.01(0.90-1.13)
V	$PM_{2.5}$	15	0.99(0.77-1.29)	1.00(0.80-1.24)
	PM_{10}	18	0.98(0.78-1.22)	1.00(0.85-1.17)
Zn	$PM_{2.5}$	18	1.04(0.84-1.28)	1.04(0.87-1.24)
	PM_{10}	18	0.99(0.84-1.16)	1.00(0.86-1.16)

[a]Effects are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn.

[b]the same cohorts were used in these two models.

Table S2 Association[a] between CVD mortality and exposure to PM constituents: Results from random-effects meta-analyses (HRs and 95% CIs)) (using additional hypertension, prevalent diabetes, and cholesterol confounder)

Exposure		Cohorts	Model3+ prevalent hypertension, physical activity ,prevalent diabetes, and cholesterol level[b]	Model3[b]
Cu	$PM_{2.5}$	18	0.91(0.74-1.12)	0.92(0.73-1.15)
	PM_{10}	18	0.92(0.78-1.07)	0.94(0.75-1.17)
Fe	$PM_{2.5}$	18	0.97(0.84-1.11)	1.01(0.84-1.21)
	PM_{10}	18	0.96(0.83-1.11)	0.94(0.76-1.16)
K	$PM_{2.5}$	17	0.98(0.94-1.02)	0.98(0.94-1.02)
	PM_{10}	17	0.98(0.92-1.03)	1.01(0.89-1.14)
Ni	$PM_{2.5}$	14	0.96(0.70-1.31)	0.89(0.63-1.27)
	PM_{10}	17	0.97(0.79-1.19)	0.97(0.80-1.16)
S	$PM_{2.5}$	18	1.05(0.92-1.20)	1.07(0.94-1.22)
	PM_{10}	18	1.06(0.86-1.29)	1.08(0.84-1.39)
Si	$PM_{2.5}$	16	1.15(0.90-1.47)	1.25(0.90-1.72)
	PM_{10}	18	0.99(0.89-1.11)	1.01(0.83-1.24)
V	$PM_{2.5}$	15	0.98(0.74-1.28)	0.94(0.69-1.28)
	PM_{10}	18	0.97(0.82-1.16)	0.97(0.73-1.29)
Zn	$PM_{2.5}$	18	1.04(0.85-1.28)	1.05(0.84-1.31)
	PM_{10}	18	0.98(0.84-1.15)	1.00(0.82-1.22)

[a]Effects are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn.

[b] the same cohorts were used in these two models.

Table S3 Association[a] between CVD mortality and exposure to PM constituents: Results from random-effects meta-analyses (HRs and 95% CIs) (using additional noise continuous variable confounder)

Exposure		Cohorts	Model3+noise (continuous variable)[b]	Model3[b]
Cu	$PM_{2.5}$	11	0.80(0.62-1.02)	0.93(0.71-1.21)
	PM_{10}	11	0.88(0.75-1.04)	1.00(0.84-1.18)
Fe	$PM_{2.5}$	11	0.90(0.8-1.02)	1.00(0.87-1.15)
	PM_{10}	11	0.87(0.75-1.03)	0.97(0.85-1.11)
K	$PM_{2.5}$	10	1.15(0.94-1.42)	1.07(0.89-1.28)
	PM_{10}	10	1.03(0.94-1.14)	1.02(0.89-1.18)
Ni	$PM_{2.5}$	8	0.81(0.66-0.99)	0.88(0.69-1.11)
	PM_{10}	10	0.81(0.66-1.00)	0.94(0.79-1.12)
S	$PM_{2.5}$	11	1.18(0.90-1.54)	1.21(0.96-1.53)
	PM_{10}	10	1.08(0.75-1.55)	1.17(0.84-1.62)
Si	$PM_{2.5}$	10	1.13(0.95-1.35)	1.25(0.95-1.65)
	PM_{10}	10	1.00(0.87-1.15)	1.03(0.89-1.20)
V	$PM_{2.5}$	9	0.84(0.57-1.23)	0.98(0.68-1.42)
	PM_{10}	10	0.94(0.69-1.27)	0.98(0.79-1.21)
Zn	$PM_{2.5}$	11	1.13(0.91-1.39)	1.08(0.81-1.43)
	PM_{10}	11	1.05(0.89-1.23)	1.07(0.92-1.24)

[a]Effects are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn.
[b] the same cohorts were used in these two models.

Table S4 Association[a] between CVD mortality and exposure to PM constituents: Results from random-effects meta-analyses (HRs and 95% CIs) (using additional noise category variable confounder)

Exposure		Cohorts	Model3+noise (categorical variable)[b]	Model3[b]
Cu	$PM_{2.5}$	11	0.81(0.61-1.09)	0.93(0.71-1.21)
	PM_{10}	11	0.90(0.73-1.12)	1.00(0.84-1.18)
Fe	$PM_{2.5}$	11	0.92(0.81-1.04)	1.00(0.87-1.15)
	PM_{10}	11	0.88(0.74-1.06)	0.97(0.85-1.11)
K	$PM_{2.5}$	10	1.16(0.94-1.42)	1.07(0.89-1.28)
	PM_{10}	10	1.02(0.92-1.13)	1.02(0.89-1.18)
Ni	$PM_{2.5}$	8	0.82(0.66-1.02)	0.88(0.69-1.11)
	PM_{10}	10	0.82(0.67-1.01)	0.94(0.79-1.12)
S	$PM_{2.5}$	11	1.21(0.93-1.59)	1.21(0.96-1.53)
	PM_{10}	10	1.09(0.73-1.64)	1.17(0.84-1.62)
Si	$PM_{2.5}$	10	1.16(0.92-1.47)	1.25(0.95-1.65)
	PM_{10}	10	0.99(0.84-1.17)	1.03(0.89-1.20)
V	$PM_{2.5}$	9	0.84(0.60-1.19)	0.98(0.68-1.42)
	PM_{10}	10	0.92(0.71-1.21)	0.98(0.79-1.21)
Zn	$PM_{2.5}$	11	1.15(0.92-1.44)	1.08(0.81-1.43)
	PM_{10}	11	1.09(0.94-1.26)	1.07(0.92-1.24)

[a]Effects are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn.
[b] the same cohorts were used in these two models.

Table S5 Adjusted association between CVD mortality and PM constituents: Results from random-effects meta-analyses from cohorts with LOOCV R^2 higher than 0.5 (HRs and 95%-CIs)[a]

Exposure		N[b]	Model 3 (LOOCV R^2<0.5)[c]	N[b]	Model3 (LOOCV R^2>0.5)[c]	P_{het}[d]
Cu	$PM_{2.5}$	0	NA	19	0.90(0.77-1.07)	NA
	PM_{10}	0	NA	18	0.95(0.84-1.08)	NA
Fe	$PM_{2.5}$	2	1.16(0.88-1.53)	17	0.97(0.85-1.10)	0.25
	PM_{10}	2	0.65(0.50-0.85)	17	1.01(0.90-1.14)	0.01
K	$PM_{2.5}$	10	0.97(0.81-1.16)	8	0.99(0.85-1.17)	0.82
	PM_{10}	8	0.99(0.73-1.35)	10	0.99(0.94-1.05)	0.98
Ni	$PM_{2.5}$	8	0.95(0.68-1.32)	3	1.02(0.83-1.25)	0.73
	PM_{10}	6	1.02(0.84-1.23)	11	1.00(0.81-1.23)	0.89
S	$PM_{2.5}$	14	1.17(0.94-1.46)	4	1.04(0.89-1.21)	0.36
	PM_{10}	13	1.21(0.87-1.69)	5	1.00(0.84-1.19)	0.32
Si	$PM_{2.5}$	8	1.22(0.72-2.05)	8	1.08(0.89-1.31)	0.66
	PM_{10}	5	0.91(0.71-1.15)	13	1.04(0.91-1.18)	0.34
V	$PM_{2.5}$	12	1.09(0.69-1.73)	3	1.00(0.81-1.25)	0.75
	PM_{10}	4	0.82(0.62-1.07)	14	1.07(0.92-1.25)	0.09
Zn	$PM_{2.5}$	7	0.68(0.32-1.42)	12	1.09(0.90-1.33)	0.22
	PM_{10}	2	1.31(0.79-2.15)	17	0.96(0.84-1.11)	0.25

[a]Effects are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn.

[b]N: number of cohorts in the meta-analysis

[c]Main model with full cohorts compared to main model with LOOCV R^2 of cohorts higher or lower than 0.5.

[d]heterogeneity of HR between cohorts with LOOCV R^2 higher and lower than 0.5

Table S6 Adjusted association between CVD mortality and PM constituents: Results from random-effects meta-analyses from three European regions (HRs and 95%-CIs)[a]

Exposure		N[b]	North[c]	N[b]	West and Central[c]	N[b]	South[c]	P_{het}[d]
Cu	$PM_{2.5}$	7	0.76(0.52-1.10)	8	0.99(0.80-1.22)	4	0.84(0.57-1.25)	0.48
	PM_{10}	7	0.93(0.80-1.07)	8	0.94(0.69-1.27)	4	0.93(0.80-1.10)	0.95
Fe	$PM_{2.5}$	7	0.95(0.85-1.06)	8	1.07(0.81-1.42)	4	0.95(0.70-1.28)	0.47
	PM_{10}	7	0.92(0.81-1.05)	8	0.98(0.72-1.32)	4	0.97(0.80-1.17)	0.74
K	$PM_{2.5}$	7	1.06(0.87-1.28)	7	0.98(0.94-1.02)	4	0.87(0.69-1.10)	0.45
	PM_{10}	6	1.03(0.95-1.12)	8	1.10(0.77-1.56)	4	0.93(0.78-1.10)	0.75
Ni	$PM_{2.5}$	3	0.82(0.54-1.25)	7	0.99(0.65-1.50)	4	1.07(0.69-1.65)	0.74
	PM_{10}	6	0.92(0.67-1.26)	7	0.92(0.71-1.20)	4	1.16(0.93-1.44)	0.37
S	$PM_{2.5}$	7	1.36(0.90-2.05)	7	1.06(0.93-1.22)	4	0.84(0.24-2.94)	0.44
	PM_{10}	7	0.93(0.67-1.29)	7	1.18(0.80-1.73)	4	1.21(0.77-1.89)	0.65
Si	$PM_{2.5}$	6	1.03(0.91-1.16)	6	1.57(0.91-2.73)	2	1.04(0.42-2.58)	0.21
	PM_{10}	7	1.02(0.95-1.10)	7	1.08(0.70-1.66)	4	0.98(0.69-1.38)	0.97
V	$PM_{2.5}$	6	1.23(0.72-2.07)	5	0.82(0.48-1.37)	4	0.93(0.25-2.34)	0.56
	PM_{10}	7	1.05(0.84-1.30)	7	0.82(0.54-1.27)	4	0.88(0.53-1.44)	0.53
Zn	$PM_{2.5}$	7	0.66(0.30-1.45)	8	1.08(0.90-1.29)	4	1.08(0.59-2.01)	0.62
	PM_{10}	7	0.90(0.73-1.11)	8	1.04(0.80-1.37)	4	0.99(0.67-1.46)	0.74

[a]Effects are presented for an increase of 5 ng/m^3 for $PM_{2.5}$ Cu, 20ng/m^3 for PM_{10} Cu, 100ng/m^3 for $PM_{2.5}$ Fe, 500ng/m^3 for PM_{10} Fe, 50ng/m^3 for $PM_{2.5}$ K, 10ng/m^3 for PM_{10} K, 1ng/m^3 for $PM_{2.5}$ Ni, 2ng/m^3 for PM_{10} Ni, 200ng/m^3 for $PM_{2.5}$ S, 200ng/m^3 for PM_{10} S, 100ng/m^3 for $PM_{2.5}$ Si, 500ng/m^3 for PM_{10} Si, 2ng/m^3 for $PM_{2.5}$ V, 3ng/m^3 for PM_{10} V, 10ng/m^3 for $PM_{2.5}$ Zn, 20ng/m^3 for PM_{10} Zn.

[b]N: number of cohorts in each region

[c]North: FINRISK, HUBRO, SNAC-K, SALT, Sixty, SDPP, DCH; West and Central: EPIC-MORGEN, EPIC_PROSPECT, SALIA, EPIC-Oxford, KORA, VHM&PP, SAPALDIA, E3N; South, EPIC-Turin, SIDRIA-Turin, SIDRIA-Rome, EPIC-Athens

[d]heterogeneity of HR between regions

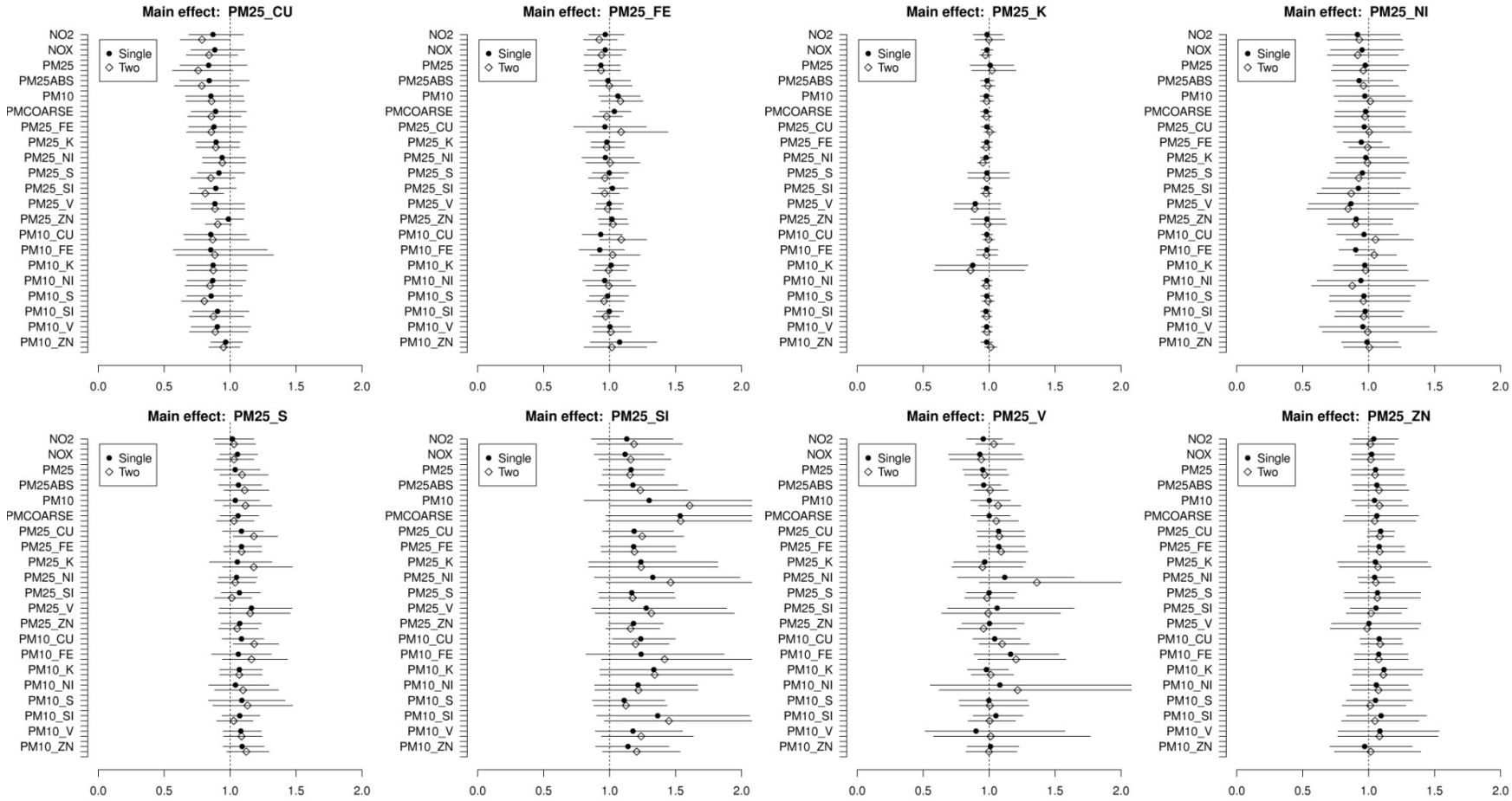


Figure S1 Two-pollutant model with main effect of 8 components in $PM_{2.5}$ adjusting all the NO_2/NOx and PM metrics while restricting the analyses to the cohorts with correlation coefficients between the two pollutants less than 0.7. To be comparable with the results of the two-pollutant model, single pollutant model included the same number of cohorts as those in the two-pollutant model.

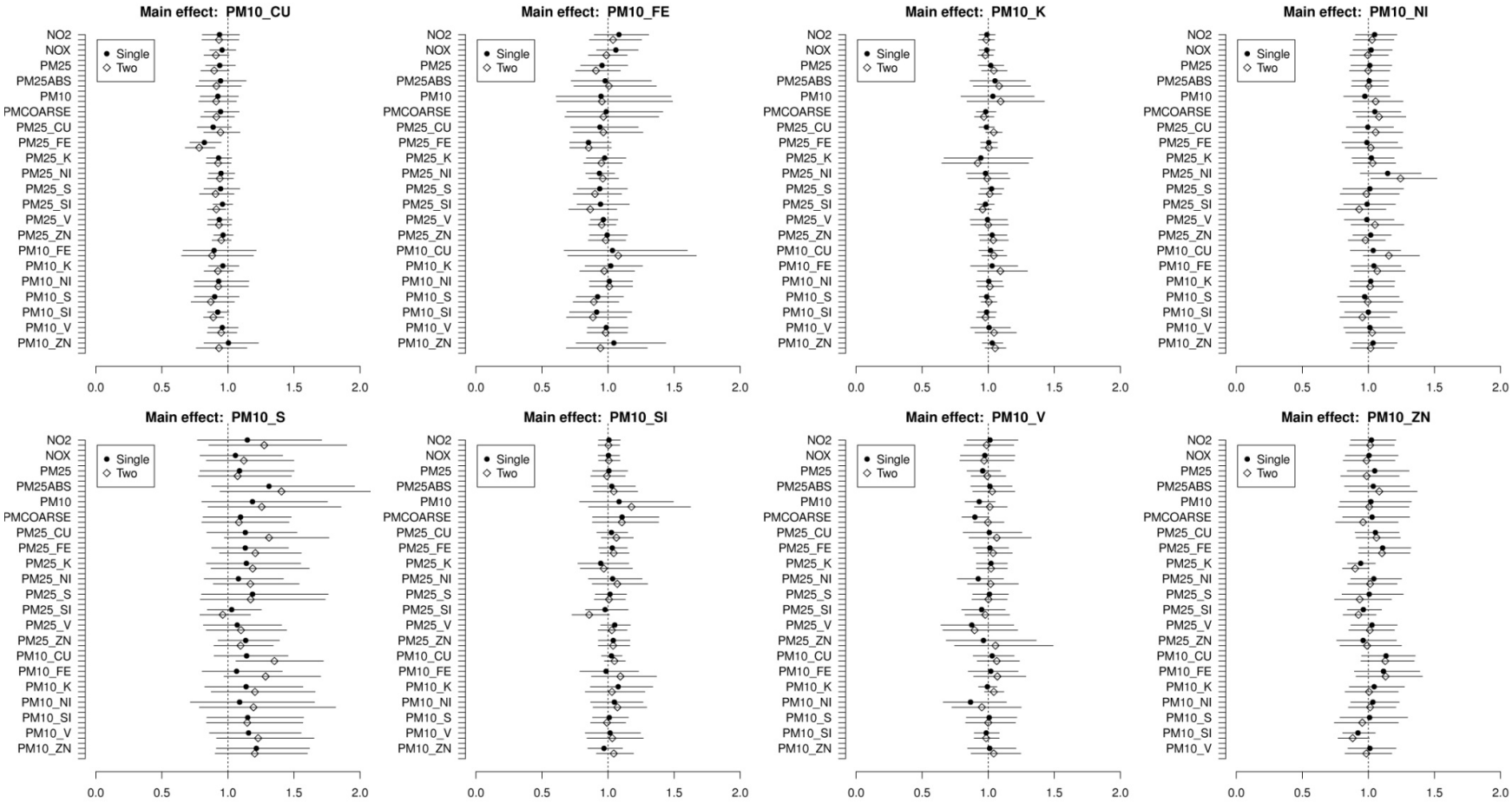


Figure S2 Two-pollutant model with main effect of 8 components in PM_{10} adjusting all the NO_2/NOx and PM metrics while restricting the analyses to the cohorts with correlation coefficients between the two pollutants less than 0.7. To be comparable with the results of the two-pollutant model, single pollutant model included the same number of cohorts as those in the two-pollutant model.

Chapter 7

General Discussion

The objective of this thesis is to develop and evaluate LUR models for air pollutants such as NO_2 and particulate constituents as well as to investigate the associations between constituents and cardiovascular mortality. The research is conducted within the framework of the European Study of Cohorts for Air Pollution Effects (ESCAPE) study which aims at qualifying chronic health effects of outdoor air pollution. This study has benefited from the large dataset collected by a standard protocol in exposure assessment and the large population of cohorts in multiple cities of Europe. With the unique dataset, we are able to gain insight into the performances of LUR models and the effects of constituents of ambient particles on cardiovascular mortality in a number of European study areas.

Main findings

Evaluation is an essential part of LUR model development. In chapter 2, we made use of a large study including 144 sites with NO_2 concentration data in the Netherlands. Hold-out validation (HV) R^2 was substantially lower than the corresponding cross-validation R^2, especially for the smallest training sets. LUR model performance for NO_2 varied with the number of training sites: HV increased and LOOCV decreased with larger training datasets. Truly independent evaluation data are especially useful when LUR models are developed from small training sets. In our study, models based on as few as 24 training sites across the Netherlands, achieved acceptable hold out validation R^2s of, on average, 0.60.

In view of this, we evaluated LUR models in all 20 ESCAPE study areas with PM and NO_2 data for NO_2 and traffic-related particulate matter (PM) components in chapter 3. True HV was not possible as in each area PM was measured at 20 sites only. We used the NO_2 measurements at the 20 sites without PM measurements as an approximation, exploiting the high correlation between concentrations of PM metrics (absorbance and Cu) and NO_2. Our results confirm that the predictive ability of LUR models based on relatively small training sets is overestimated by the LOOCV R^2s. Nevertheless, in most areas LUR models still explained a substantial fraction of the variation of concentrations measured at independent sites.

In chapter 4, we attempted to develop European and regional LUR models by combining 23 and 17 ESCAPE study areas for NO_2 and PM, respectively, and to evaluate the model performances. These combined models are based on large numbers of training sites. The European models explained 56% of the concentration variability across all sites for NO_2 (based on 960 sites), and 86% and 70% for $PM_{2.5}$ and $PM_{2.5}$ absorbance, respectively (based on 356 sites). The prediction ability of the European models was only slightly overestimated by model R^2. For NO_2 and $PM_{2.5}$ absorbance, these models predicted spatial variations in areas not used for model building well. For $PM_{2.5}$, prediction R^2s were moderate for intra-urban variation. These results suggested that it is possible to include new study populations from areas where local measurements were never conducted but relevant predictor variables included in the LUR models are available.

In chapter 5 we developed LUR models for eight a priori selected particle elements; copper (Cu), iron (Fe), potassium (K), nickel (Ni), sulphur (S), silicon (Si), vanadium (V) and zinc (Zn) for all study areas using a specially written

standardized script. Good models were developed for Cu, Fe and Zn in both fractions (PM_{10} and $PM_{2.5}$) explaining on average between 67 and 79% of the concentration variance (R^2) with, however, a large variability between areas. Traffic variables were the dominant predictors, reflecting non-tailpipe emissions. Models for V and S in the PM_{10} and $PM_{2.5}$ fractions and Si, Ni and K in the PM_{10} fraction performed moderately with median R^2 ranging from 50 to 61%. Si, NI and K models for $PM_{2.5}$ performed poorest with median R^2 below 50%. Lack of specific predictor variables e.g. on wood burning emissions for K and small spatial variation (e.g. for S) likely contributed to poorer models.

With the estimates of particle composition concentrations calculated using the LUR models developed in chapter 5, we assessed the association with cardiovascular mortality in 19 cohorts across Europe, adjusting for major confounders. We found no significant association between long-term exposure to 8 elemental constituents of particles and total cardiovascular mortality in a joint analysis of these cohorts. Most of the hazard ratios were close to unity, with the exception of $PM_{2.5}$ Si and S in $PM_{2.5}$ and PM_{10} (chapter 6). The HR for Si in $PM_{2.5}$ was 1.17 (95% CI: 0.93-1.47) per 100ng/m^3 and for S in $PM_{2.5}$ and PM_{10} the HR was 1.08(95% CI: 0.95-1.22) and 1.09(95% CI: 0.90-1.32) per 200ng/m^3 respectively.

Land use regression model development

Although there is no rigorous rule for the number of sampling sites needed for LUR modeling, our studies suggested that modeling with small number of sampling sites may overestimate the prediction ability of LUR model to independent locations. These findings were consistent with the results of other studies in Spain and Canada[1, 2]. In most of the ESCAPE study areas, LUR models still explained a substantial fraction of the variation of concentrations measured at independent sites. Our study followed the strict ESCAPE measurement and modeling criteria using a standardized protocol for air pollution measurements, modeling procedures and data collection in exposure assessment in all study areas, aiming at representing spatially harmonized and physically interpretable distribution of cohort exposures across study areas. However, since LUR modeling is an empirical technique, there is no gold standard method for model development[3, 4]. Table 1 summarizes current main techniques to derive LUR models. Many studies used various strategies including basic backward, forward or stepwise linear regression techniques to develop a parsimonious model from a large predictor variables dataset to maximizing percentage of explained spatial variability (model R^2)[4]. Within the ESCAPE project, we used a supervised forward stepwise procedure to select influential predictors from a large set of a priori defined predictor variables. Each variable had an a priori hypothesized effect of direction based on physical principles in order to generate an interpretable model structure and to reduce

the over fitting risk[3]. This strategy has been widely used for LUR model development in many studies[5-8]. The models generally explained modest to large spatial concentration contrasts. Other studies used an automatic forward variable selection approach without any restrictions on expected sign and significance of

Table 1 Summary of current modeling strategies in the frame work of LUR modeling

Methods	Selected studies	Advantages	Limitations
MLR without restrictions[a]	Crouse et al. (2009)[9]	• Model explains largest variability	• High over-fitting risk • Difficult to interpret model structure • Spatial autocorrelation in residuals • Generalisability
MLR with constraints[a]	Eeftens et al. (2012)[12] Beelen et al. (2012)[3] de Hoogh et al. (2013)[13]	• Interpretable model structure • Model explains variability from modest to high • Generalisability	• Over-fitting risk • Spatial autocorrelation in residuals
MLR+[b]IDW[a]	Hoek et al. (2001)[11] Briggs et al. (2005) Beelen et al. (2007)[7]	• Interpretable model structure • Model explains variability from modest to high • Physically logical procedures • Generalisability • Relatively low over-fitting risk	• Not suitable in small scale • Complex computation
MLR+coKriging[a]	Mercer et al (2011)[12] Bergen et al. (2013)[14] Li et al. (2012)[15]	• Interpretable model structure • Model explains variability from modest to high • Physically logical procedures • Generalisability • Low spatial autocorrelation in residuals	• Not suitable in small scale • Complex computation
ADRESS[c]	Su et al. 2009a[16] Su et al. 2009b[17]	• Interpretable model structure • Model explains variability from modest to high • Physically logical procedures • Low spatial autocorrelation in residuals	• Complex computation • Not easy in practical work
DSA[d]	Beckermann et al. 2013b[18] Beckermann et al. 2013a[19]	• Interpretable model structure • Low over-fitting risk • Generalisability	• Complex computation • Low model R^2 • Modeling procedure difficult to interpret • Require relatively large sampling sites • Spatial autocorrelation in residuals

[a]MLR: multiple linear regression;
[b]IDW: inverse distance weight;
[c]ADDRESS: A Distance Decay Regression Selection Strategy;
[d]DSA: Deletion/Substitution/Addition algorithm

variable coefficients as well as on the effects of multicollinearity between variables with the purpose to maximize the model R^2 [9]. Although the simple method usually produced high model R^2, the prediction ability of the models may be overestimated as shown in chapters 2 and 3.

Several studies in the Netherlands and UK established a hierarchical approach to model air pollution concentrations assuming ambient concentrations as a combination of rural and urban background, and within-urban local contributions from traffic[7, 10, 11]. Concentrations at predefined rural background sites were interpolated using inverse distance weight (IDW) approach in the first stage to generate a regional background concentration of air pollution. The residuals of the urban and street sites were then explained by multiple regression models in the following stages using urban background and traffic variables respectively. Since this approach relied on more theoretical knowledge with efforts to explain regional transport and local emissions separately it may increase the likelihood of transferring the developed models elsewhere[4]. Further, concentrations at urban and regional background sites may be better predicted compared to an overall model that may be dominated by traffic predictors. However, the approach is limited by the availability of a sufficient number of rural sites and is usually not possible in small regions or at city-wide scale due to homogenous regional distribution in spatial coverage. Instead of using smoothed- regional variable, Brauer et al. (2003) incorporated an indicator variable in the Dutch model by classifying regions according to influential factors such as population, meteorology etc. In ESCAPE, the multilevel strategy was not applied but we did evaluate whether the addition of regional estimates at the end of modeling with GIS predictors improved the model R^2 in larger study areas such as the Netherlands.

Instead of IDW, several studies used Kriging in a two-stage LUR framework[14, 15, 20]. This approach incorporated a spatial smooth term and was powerful to reduce spatial autocorrelation in the residuals. Several studies have documented consistently that the two-stage model slightly outperformed and was more robust than multiple linear regression models[20]. In ESCAPE, spatial autocorrelation in residuals has been investigated in the model development using Moran's I and was generally very low and insignificant for area-specific LUR models across study areas. However, since significant modest spatial autocorrelation has been reported in a Canadian national model for $PM_{2.5}$, further research is needed to assess the spatial autocorrelation within the European models and to develop a two-stage model for comparison.

Concerning sampling approach and selection of predictor variables for spatial modeling, Su et al. (2009a,b)[16, 17] used a location-allocation algorithm for site selection and a programmable distance decay regression selection strategy to decide appropriate buffers of variables which resulted in 86% and 85% of the explained variance in measured NO_2 and NOx. This method took population distribution into account and required intensive computation based on geographic data to select qualified sites from existing monitoring sites[21]. In ESCAPE, we conducted a relatively pragmatic strategy in monitoring sites selection by taking different site types into account. We allocated 20%-40% of the sites to (busy) street locations expected to exhibit the largest concentration contrasts for model

development. We also used different buffer sizes for the predictor variables largest radius of 5km based on observed air pollutants' dispersion behaviors in urban environments. Our models explained large fractions of variability which were at least comparable with, or even better in some study areas than the model R^2 produced by Su et al. (2009)[16, 17].

Recently, a stricter algorithm, Deletion/Substitution/Addition (DSA), has been introduced under the frame work of LUR methodology[18, 19]. The DSA approach uses an aggressive covariate search algorithm to fit a generalized linear model. Instead of maximizing model R^2, this approach seeks to search for the best predictive model by asymptotically optimizing cross-validation R^2 and tests nearly all the covariate combination of both polynomial and interaction terms[18]. Since the DSA algorithm seeks to balance the needs to maximize predictive power and to minimize over-fitting, this algorithm has been reported with both good estimates and performance (LOOCV R^2=0.79 for $PM_{2.5}$[18] and 0.71 for NO_2[19]). Basagana recently compared the performance of the DSA algorithm with that of the ESCAPE standard procedure with Spanish NO_2 data, showing that even though model R^2 of the LUR model was less inflated by using the DSA than the ESCAPE procedure, the prediction ability (HV R^2) of the DSA was, however, lower than the ESCAPE procedure especially when the number of training sites are larger than 30 (HV: MLR with constraints 32%-51% versus 29%-47%, N=30-120)[2].

In summary, the performance of modeling procedures in the ESCAPE study has been well evaluated. Development of LUR model in the future should consider less empirical strategy (i.e. more theoretical based restriction). For instance, predictor variables should be more specific for emission sources. Buffer selections of predictor variables should rely on dispersion patterns or spatial distribution characteristics to be more representative in model structures. More sampling sites are needed if available to minimize over-fitting risk. Although the ESCAPE LUR models explained fairly large fractions of spatial contrasts based on 40 and 20 sites for NO_2 and PM respectively, we realize that there is still room to improve the model performances especially for PM if more sampling sites were available. However, this is always a trade-off between better quality of modeling and more study areas with limited funding and resources in multi-cities study such as ESCAPE. Further research is needed to compare models with different approaches such as the two-stage modeling with kriging and the DSA algorithm.

Comparison between LUR and dispersion models

This thesis has evaluated the land use regression (LUR) technique and explored its performance in terms of prediction ability. Dispersion modeling is another main approach that has been extensively applied to characterize small-scale spatial variability of air pollution[22]. For many years, these two approaches have been implemented in fine scale exposure assessment for epidemiological studies[23].

Dispersion models generally rely on dispersion theory (e.g. Gaussian plume equation)[24] and utilize data on emission, meteorological conditions and topographical data to simulate the physic-chemical processes of transport and atmospheric chemistry when estimating outdoor air pollution concentrations[22]. Like

LUR models, recent dispersion models have been used in conjunction with GIS to provide more realistic and higher spatial-resolved predictions than the conventional ones. Currently, comparison of LUR and dispersion model has not been widely conducted. In general, LUR modeling has advantages in its relatively simple inputs, its straightforward constructing procedures, and its capability of fine scale predictions, but often suffers from lack of temporal and larger spatial scale data[4]. Dispersion modeling is more advanced in covering larger periods of time whenever multi-year records of emissions and meteorological conditions are available, but is subject to relatively low spatial resolution in meteorological data and absence of accurate emission data on a small scale[25]. In this section we will compare the performances of the two kinds of models from previous studies supplemented with part of the study thesis.

Spatial aspects

A growing number of studies compared the two techniques with emphasis on the ability of spatial predictions. To achieve a fair comparison, an independent validation air pollution data set is usually required aside from the sites used for modeling which could be either Hold-out Validation (HV)-based[26, 27] or routine monitoring sites[28]. As documented in chapters 2 and 3, LUR model R^2 can be inflated by 20-30% for training sets of 24-40 sites. Published comparisons have been summarized in table 2. Comparisons were mostly conducted in European cities with one exception in Vancouver, Canada. Locally developed dispersion models were employed in each study area due to different characteristics and data quality in local emissions, meteorology and topography. Emission data of dispersion models were less complete for $PM_{2.5}$ in most of the studies hence comparisons have been mainly on NO_2 prediction.

Overall, few comparisons have been performed in different settings using different dispersion models at different spatial scales. The CMAQ and FARM model (Table 2) did not model small-scale variation but only the urban background. Furthermore, even fewer studies had the ability to compare the models at a sufficient number of independent sites. Comparisons at a small number of validation sites (e.g. 10) can be difficult to interpret. In some studies, LUR model outperformed localized dispersion model, i.e. correlations between predictions and observations at external sites were larger for LUR model than for the dispersion model. In two studies in the Netherlands the opposite was found. The number of training sites for LUR modeling in the comparison studies is modest to high (40-497). The difference between model R^2 and HV R^2 is generally small, though some models even predicted higher HV R^2 than the model R^2. This is attributable to the use of small validation sets with respect to corresponding large training sets, for instance 72 versus 8 sites in Huddersfield[29] and 116 versus 14 sites in Vancouver[28]. Our systematic evaluation in chapter 2 (Figure 1) has demonstrated higher HV R^2 than model R^2 occasionally from hundreds iterative computations[30], suggesting that predictions may be highly variable with small validation sets. The findings of the Dutch studies can be explained from several reasons (e.g. different sampling periods, quality of important predictor variables), one of which is the use of a large-area LUR model on a smaller-area, e.g. a Dutch national LUR model applied

Table 2 Performance of LUR and dispersion models at the same validations sites from different studies

Study	Area	Pollutant	Dispersion model	LUR Model R^2	N Model	N Validation	LUR HV R^2	Dispersion Validation R^2	Agreement R^2 [d]
Brigg et al (2000)[27]	Huddersfield, UK	NO_2	CALINE-3	0.61	72	8	0.82	0.63	NA
Cyrys et al (2005)[30]	Munich, Germany	NO_2	$IMMIS^{net}$	0.62	40[a]	NA[a]	NA	0.46	0.69
Marshall et al (2008)[26]	Vancouver, Canada	NO	CMAQ	0.62	116	14	0.56	0.21	0.16
		NO_2		0.53			0.58	0.50	0.24
Beelen et al (2010)[25]	Rotterdam, Netherlands	NO_2	URBIS	0.88	44	16	0.22	0.59	0.30
Dijkema et al (2011)[29]	Amsterdam, Netherlands	NO_2 (large-area)	CAR	0.87	60	62	0.48	0.57	NA
		NO_2 (city-specific)		0.72	62	13	0.57	0.74	NA
Gulliver et al (2011)[24]	London, UK	PM_{10}	ADMS-Urban	0.52	52	52[b]	0.47	0.28	NA
Liu et al (2012)[23]	Swiss cities	NO_2 (1993)	PolluMap	0.84	497[a]	NA[a]	NA	0.71	NA
		NO_2 (2003)		0.80	348[a]	NA[a]	NA	0.64	NA
Cesaroni et al (2013)[31]	Rome, Italy	$PM_{2.5}$	FARM	NA	NA	18[c]	NA	0.41	NA
Wang et al (chapter 4)	Rome, Italy	$PM_{2.5}$	NA	0.56	336	20[c]	0.54	NA	NA

[a]No validation sets were available, all sites were used for model development

[b]Sites were divided equally into four groups three of which were used for modeling to validate the rest dataset.

[c]18 validation sites from a total of 20 ESCAPE measurement sites were used for validation of dispersion model by Cesaroni et al. 2013[31]; whereas a full set of 20 sites were used in Wang et al. (chapter 4).

[d]Agreement: correlation coefficient of predictions between dispersion model and LUR model

to the Rijnmond area (1404 km^2), the Netherlands[27].This could be deduced from the observations in our study (chapter 2 table 2) that the models that were based on similar number of training sites as Beelen et al (2010)[27] and Dijkema et al (2011)[31] were still able to predict at least 67% HV R^2s at the validation sites throughout the Netherlands[30]. The aforementioned problem may be even worse for particulate matter (PM) and some of their compositions (e.g. S and K), which have clearly shown larger between-area than within-area variability in the ESCAPE multi-cities in chapter 4 and in Tsai et al. (in preparation), than for traffic-related pollutants e.g. NO_2. Our European model basically predicted little variations for $PM_{2.5}$ concentrations in some central European cities which is partially due to lack of representative local predictor variables. Gulliver et al. (2011)[26] recommended including information on site type which may partially overcome such problem. It is worth noting that comparisons in Munich and Swiss cities were not ideal because independent validation sets were not available for LUR models in these studies[25, 32]. However, our study in chapter 3 (Figure 2) exhibited that the ESCAPE Munich model can predict 62% variations at independent sites even based on 20 training sites. Meanwhile, our European NO_2 model based on 480 sites showed that the HV R^2 was almost equal to the model R^2 (model R^2 0.57 versus HV R^2 0.54) (chapter 4 table S3). These results suggested that the predictions of the LUR models in Cyrys et al. (2005) and Liu et al. (2012) were reliable.

The ESCAPE study produced a rich empirical air pollution dataset and a large number of city-specific LUR models for air pollutants which can be used to conduct more comparisons with dispersion models. A comparison between ESCAPE published LUR models and local dispersion models in available study areas is ongoing (de Hoogh, in preparation). One of the foreseenable problems is the selection of validation sites. Use of grouped jack-knife analysis seems infeasible especially for PM due to the small number of sites (20 per area). Using the AirBase regulatory monitoring sites for validation may avoid the above problem but the interpretation of results may be difficult as the validation sites may not be representative to population exposures and there are few sites within the individual cities/study areas that were included in ESCAPE. Alternatively, validating air pollution at external independent areas has been implemented for the European LUR models in chapter 4 (transferability) and therefore is potentially available to conduct reasonable comparison with local dispersion models. As an example in table 2, a Rome $PM_{2.5}$ dispersion model reported by Cesaroni et al (2013)[33] presented 41% agreement at a subset of 18 of the total 20 ESCAPE sampling sites (2 regional sites outside Rome were excluded) compared to 54% predicted by our European $PM_{2.5}$ model for all 20 sites.

Temporal aspects

Dispersion models are powerful tools to produce historical exposures if relevant emission and meteorological data are available, hence it has been suggested to assess retrospective individual exposure of air pollution[22]. Downs et al. (2007)[34] assessed PM_{10} exposures between 1990 and 2002 using a Swiss dispersion model to assess effects of reduced exposure to PM_{10} and decline in lung function in adults.

Land use regression models are usually constrained to short temporal coverage of purpose-designed monitoring campaigns which requires additional steps to extrapolate predictions over time[4]. Recent epidemiological studies increasingly relied on LUR modeling to predict historical annual or seasonal exposure to ambient air pollution using recent/current air pollution data[35-37]. In our study, the LUR models used for exposure assessment were based on air pollution measurements in the period 2008-2011 while cohort studies included in ESCAPE started in the past (1985-2007 with most studies starting in the mid-90s). Thus evidence was needed to validate the extrapolation of LUR model estimates back in time. So far, five studies investigated the performance of back-extrapolation of LUR model estimates for NO_2 in the Netherlands[38], Great Britain[39], Rome[40], Oslo[41] and Vancouver[42]. These studies show that for periods from 3 to up to 18 years NO_2 spatial contrasts often remained the same, even with a decrease in concentrations over time. This indicates that LUR modeling is able to provide reliable estimates for historical exposure. However, this finding can only be generalized to cities or countries with limited reconstruction activities and a stable development history, and therefore is applicable to many European countries but not to all. Furthermore, it is not clear whether this finding can be generalized to more complex particulate mixtures.

LUR models often model annual average concentrations. Recently, more sophisticated spatiotemporal LUR models have been developed in the United States[18, 43] and Canada[44] which incorporated temporally resolved data from continuous regulatory monitoring networks and satellites. This hybrid approach applied multistage weighted and smoothing techniques over space and time and was capable to estimate national scale spatiotemporal variability of air pollution. Several epidemiological studies have adopted this approach to estimate both long-term and short-term associations between hospitalization and mortality in concurrent period[44, 45]. A disadvantage of the updated LUR models we just referred to stems from the relatively coarse spatial resolution (10km x 10km) which misses small-scale variability thus is not suitable for within-city comparisons such as the ESCAPE study.

An alternative approach is to combine dispersion and LUR in one module. Few recent studies in Europe developed local spatiotemporal models as a combination of regression based and local dispersion models in the context of assessing health effects of long-term exposure to air pollution in a large cohort[25, 46]. A logical next step is therefore to develop local and European-wide spatiotemporal models taking the advantages of dispersion model and the European AirBase regulatory network.

Validity of LUR model in predicting personal exposures

Determinants of residential outdoor air pollution

Although this thesis documented a reasonable prediction ability of LUR models of concentrations for the pollutants NO_2, $PM_{2.5}$ absorbance, and the Cu, Fe and Zn content of PM, it should be noted that these estimates reflect residential outdoor concentrations. Residential outdoor concentrations are often used as a surrogate for personal exposure, the ideal exposure for epidemiological studies. Some

studies consistently showed that residential indoor concentrations serve as an important indicator for personal exposure given the large fraction of time people spend at home[47, 48]. Indoor concentrations are affected by infiltration of outdoor air and indoor sources. For pollutants with strong indoor sources, residential outdoor exposure alone may not be a good proxy for personal exposure. For instance, a large-scale population study in six European cities (EXPOLIS study) illustrated a fair amount of indoor $PM_{2.5}$ and NO_2 concentration contributions from indoor smoking, and gas stove and fuel heating respectively[49]. A study in 14 U.S. cities found that short-term effects of PM_{10} presented stronger association with hospital admission with lower usage of air conditioning, explained by the strong influence of infiltration rates[50]. Van Roosbroeck et al. found no differences between the personal exposures to $PM_{2.5}$ and NO_2 of adults living on high- or low-traffic streets because of the influence on exposures of time-activity patterns, indoor sources, and ventilation[51]. They further suggested that the effect of exposure to outdoor air pollution on respiratory and other health effects may be substantially attenuated when exposure measurement used outdoor concentration in children's schools instead of personal exposure[52].

Association between LUR outdoor predictions and personal exposure

Despite a fair number of studies documenting often moderate to strong associations between temporal variation of outdoor, indoor and personal exposure, the validity of LUR modeling regarding to personal exposure has been scarcely investigated. Only four studies attempted to validate the accuracy of estimates of LUR modeling by personal exposure monitoring.

One study was conducted in Vancouver, Canada with aims to characterize the difference between personal exposure to air pollutants and intra-urban concentration estimates using a LUR model and to identify other determinants associated with personal exposure[53]. This study included 62 non-smoking pregnant women with 1-3 48-hour personal measurements for NO, NO_2, fine particle and soot annual estimates from LUR modeling have been corrected by continuous measurements to generate monthly estimates to compare with simultaneous personal exposure. The authors found that the LUR outdoor predictions were only associated with NO and NO_2 personal exposure while use of gas stove has significant impact on the personal exposure to all the pollutants comparing with using electronic stove

Another study in Hamilton, Canada recruited a panel of 33 elderly adults for 72-hour personal exposures with concurrent residential indoor and outdoor measurements in three seasons[54]. In addition, a monitoring network-based LUR model was previously developed of which the estimates of annual average NO_2 concentrations were subsequently compared with the short- term personal measurements. The authors found that personal NO_2 exposure was consistently and strongly associated with contemporaneously collected indoor (Pearson correlation R=0.86) and outdoor (Pearson correlation R=0.72) concentrations, whereas personal exposure was not associated with annual estimates of NO_2 predicted by the LUR model. The authors explained it as the influence by temporal differences in the concentrations given the short time scale of personal

measurement. Another important factor is the use of regulatory monitoring sites for the development of an NO_2 LUR model which may not be representative to residential outdoor exposure. As the study showed, the NO_2 predictive values of the LUR model were only weakly associated with outdoor NO_2 concentrations (Pearson correlation R=0.33). This study illustrated the challenges of obtaining a sufficiently large temporal and spatial coverage when comparing personal measurements with estimates from LUR models.

A recent study in Manchester, UK included 85 non-smoking pregnant women with a 48-hour personal measurement[55]. The ESCAPE NOx LUR model has been adjusted by temporal trends to obtain monthly average of outdoor exposures. The authors found that the associations between LUR model outdoor predictions and personal exposures became remarkedly stronger after monthly adjustment (R^2=0.09 versus 0.59) which supported the importance of comparing personal exposure with model prediction at the same time scale.

More recently, the VE^{3}SPA project has been done in Europe to validate the effects of LUR models from the ESCAPE project in estimating residential outdoor concentrations on long-term personal exposures for NO_2, $PM_{2.5}$ (including soot)[56] and elemental constituents (In preparation). In the first report focusing on NO_2, $PM_{2.5}$ and soot, three panels of 15 subjects following a school child or elderly time activity pattern from Helsinki (Finland), Utrecht (The Netherlands) and Barcelona (Spain) living in three different types of areas (semi-urban, urban background and near traffic residences) participated. In this study six 96-hour indoor, outdoor and personal measurements over three seasons were collected to increase the temporal coverage of personal exposure monitoring. Annual average air pollution concentrations were calculated from individual measurements by adjusting temporal trends from a reference site with continuous measurements. These averages were compared with annual average concentrations of air pollutants predicted by LUR models based on the ESCAPE purpose-designed sites. Soot was the only component of which the estimated concentrations were significantly associated with outdoor, indoor and personal exposures across the three study areas (agreement: outdoor>indoor> personal). Interestingly, elevated and significant correlations for all modeled pollutants were observed in a pooled analysis. In contrast to previous studies, the agreements between modeled and outdoor measured were considerably high for NO_2 and soot particles, confirming the findings of LUR performances in chapter 4. This study suggested the use of LUR modeled traffic-related pollutants with less indoor sources as proxy for personal exposure.

These studies have suggested that modeling outdoor concentrations without taking personal activity and household information into account may generate slight to large exposure error from true personal exposure depending on how these pollutants are related to these factors. This error may reduce the statistical power to detect true causal effects between air pollution and health outcomes.

The predictive power of personal exposure is related to a number of factors, including personal time-activity patterns, ventilation characteristics, type of indoor source, and residential outdoor exposure. Few studies attempted to use LUR-style models to predict indoor concentrations or personal exposures. Finding no

association between personal NO_2 exposure and outdoor predictions by LUR model, Sahsuvaroglu and his colleagues (2009) combined illustrative household and personal activity factors in the original model to explore personal NO_2 exposures for these participants, showing that the LUR model partially predicted NO_2 exposures in the presence of time-activity diaries (e.g. gas stove, windows or time indoors for NO_2; model R^2=0.32-0.33)[54]. Another study in Vancouver on exploring LUR predictions for personal exposures of pregnant women found consistently that the effects of outdoor estimates by LUR on personal exposure did not significantly differ with and without personal (e.g. gas stove) adjustment for all pollutants except soot, and thus suggested use of outdoor concentrations estimates as proxies for exposure in epidemiological studies[53]. Recent studies made efforts to establish a multi-stage framework to incorporate personal estimates from separate models for indoor and outdoor exposure[57, 58]. This integration included modeling outdoor concentration by LUR, simulating indoor-outdoor infiltration factor according to I/O measurements and available questionnaire and finally estimating personal exposure by weighing indoor versus outdoor time activity patterns of participants.

In the ESCAPE study, we modeled residential outdoor concentrations assuming that indoor and, more importantly, personal concentrations were correlated with outdoor concentrations. This has been validated by the VE^{3}SPA study which suggested to use intra-urban LUR model for especially soot in epidemiology.

Health effects at between- and within-area level

Associations between air pollution and health effects can be studied either by using between or within study area exposure contrasts or both depending on the study interests and the characteristics of air pollutants in space. Within the ESCAPE study, cohort-specific analyses were conducted within individual study areas with emphasis on exposure to traffic related pollutants. The effect estimates therefore indicated the health effects due to exposure in different locations in a study area, which usually consisted of an urban area with surrounding more rural areas. This design has benefited from a growing body of evidence that gradients of exposure to air pollutants were substantial in a small spatial scale especially for traffic-related pollutants[59, 60]. In the ESCAPE study, a majority of the pollutants showed substantial within-city variation supporting the needs of exploring their effects linked to CVD mortality and other endpoints[13, 61, 62]. For the traffic-related pollutants (e.g. NO_2, soot, Cu and Fe), within-city variances were larger than between-city variances. Several studies confirmed the importance of with-city variability as a risk factor for mortality due to air pollution[63, 64]. Within-city effects of $PM_{2.5}$ on CVD mortality were larger than between-city effects in one study[65]. For constituents of particles, differences in the compositions of PM emitted from nearby emission sources may trigger different responses of health effects related to CVD within a city[66]. Regardless of characterizing spatial heterogeneity of air pollutants within a city may result in severe exposure errors with both classical and Berkson types[67, 68]. Such exposure misclassifications may diminish the risk effects toward null[69]. On the contrary, assessing exposure with refined spatial resolution within a

city may be more powerful to probe the true deleterious effects on CVD mortality. Nevertheless if there are large contrasts in exposure across cities, the design of comparing health effects across cities remains powerful. For $PM_{2.5}$ and K, S in PM, the within-city variability was much smaller than the between-city variability[13], suggesting that these metrics tended to vary in a large spatial scale[70]. Therefore, the between city design is also needed.

A recent review on CVD mortality risk linked to $PM_{2.5}$ found an significant effect (RR 1.15 95%CI 1.04, 1.27) per 10 μg/m^3 based on a series of large cohorts studies[23]. However, most of the studies were conducted in the U.S. and Canada. Geographic differences in CVD risk due to PM have been reported between North American and European cities[71]. In the ESCAPE study with 19 cohorts, no associations were found between PM and CVD mortality whereas effect estimates varied between study areas[72], suggesting the differences in mobile (e.g. diesel emission in Europe) and stationary (e.g. coal combustions) sources in different regions of Europe. Although the results of area-specific studies were inconsistent across study areas[73-75], elevated risks of CVD mortality were consistently related to V and Ni in both $PM_{2.5}$ and PM_{10} in the large U.S. national time-series study[76, 77]. Long-term studies have not yet been done except the California Teacher Study (CTS) which indicated that constituents derived from combustion of fossil fuel as well as those of crustal origin were associated with cardiopulmonary and ischemic heart disease (IHD) mortality[78]. In contrast, our study found no associations between all the eight elements (copper, iron, potassium, nickel, sulfur, silicon, vanadium and zinc) and total CVD mortality based on 19 European cohorts. The inconsistency can possibly be explained by the differences in exposure assessment, characteristics of participants and confounders adjustments between European and American cohorts. Because of the limited number of long-term studies on CVD mortality in relation to PM constituents, the role of important constituents remains less clear. However, toxicological studies have identified several transition metals (copper, iron, vanadium, nickel, and zinc) that are likely to promote inflammation (e.g. TNF-α, PMN) and oxidative stress[79-81]. Sources such as non-exhaust emissions (copper and zinc) are important for pulmonary toxicity[81]. Because a constituent is emitted from several sources and a single source emits several pollutants[82], the null effect of our study results could be interpreted as that the power of single constituents may be weak to represent the effect of its relevant sources whereas it is unclear whether a joint set of similar trace markers can increase the power to detect such effect. One source apportionment study exhibited that road dust, minerals and fuel oil from grouped constituents were associated with daily CVD mortality[83]. It is valuable to investigate the adverse CVD health effects linked to sources of pollution in the future.

Lessons learned from the ESCAPE study and perspectives in China

Mortality risks attributable to ambient air pollution has become one of the major concerns and challenges for Chinese public health administrations. Although associations between air pollution and CVD, respiratory and total non-accidental

mortality has been widely documented by short-term studies in Chinese mega-cities[84], mortality effects of long-term air pollution exposure have been understood less mainly due to lack of efficient quantitative techniques in exposure assessment. The ESCAPE study has implemented a successful method to quantify spatial variability of air pollution for cohort studies. Several lessons are learned in air pollution exposure assessment from this thesis and the ESCAPE study.

Firstly, exposure modeling of intra-urban spatial variability of air pollution is needed for cohort studies. So far, exposure assessment in China has merely relied on a central station or a small number of monitoring stations in a city. For instance, a prospective cohort (the China National Hypertension Follow-up Survey) recruited 158,666 adults, of which only 70,947 subjects (44.7%) were eligible for the air pollution survey due to inadequate ambient monitoring stations[85]. Mesoscale dispersion models are often used in Chinese studies for air pollution simulation in regional and city scales. These dispersion models have relatively coarse spatial resolution (finest resolution of 4km by 4km) as well as modest agreement with spatial measurements. For instance, a common used WRF-CMAQ (Weather Research and Forecasting-Community Multiscale Air Quality) dispersion model has been evaluated in the CAREBeijing-2006 (Campaign for Air Quality Research in Beijing and the Surrounding Region) which presented agreement of 0.7 (R) with measured aerosol optical depth (AOD) values[86]. Dispersion modeling predictions are often difficult to implement for multiple pollutants because input sources are possibly incomplete[86]. In this thesis, we have demonstrated that LUR modeling is a powerful approach to predict small-scale variability of air pollution in many European cities. Hence this modeling framework can potentially be applied in meta-cities of China. LUR modeling has been developed in a few studies in Tianjin and Jinan cities[87-89]. The model explained 62% and 64% for NO_2 and PM_{10} in Jinan(N=14) and 89% and 84% for NO_2 and PM_{10} in Tianjin (N=20) respectively, suggesting that it is feasible to develop LUR models in Chinese cities with relatively good model R^2. In contrast to the ESCAPE study, these studies relied on routine monitoring sites rather than purpose-decided sampling sites.

Secondly, the necessity of a standard protocol has been clarified, which needs to be kept throughout the exposure assessment for multi-city studies. The intention is to minimize possible errors that may lead to heterogeneity in study results. A universal protocol was followed in ESCAPE study throughout the whole air pollution sampling, data collections, modeling procedures and statistical analyses, which reduced heterogeneity of estimated effects across study areas. In China, substantial heterogeneity of effect estimates of air pollution on various mortality outcomes has been found in a recent comprehensive review of the Chinese short-term studies[84] as well as the China Air Pollution and Health Effects multi-cities Study (CAPES)[90]. At least, some of the errors were attributable to the differences in the measurements and statistical analyses in individual studies. Therefore, use of a standard protocol may help to reduce heterogeneity between studies.

Thirdly, sites distribution should be purposely designed to represent the anticipated spatial variation of air pollution at home addresses of participants rather than simply applying routine monitoring sites in the epidemiologic studies. Although

the number of the Chinese national routine monitoring stations for $PM_{2.5}$ have been increased to 1500 throughout the country, most of the station settings were not densely distributed within a city (except a few number of big cities e.g. Beijing N=35) and were not representative to population exposures. Besides, the monitoring data are not easily accessible. Concentration contrasts in mega-cities such as Beijing and Shanghai may be larger than those of the ESCAPE study areas. Given the complexity of city configurations and the characteristics of multiple sources air pollution which are more obvious in these Chinese mega-cities, sites should be more densely distributed surrounding populations with more efforts to capture hotspot concentration variances of air pollution.

Fourthly, predictor variables should be as specific and accurate as possible to explain known sources of air pollutants. In the ESCAPE study, local predictor variables were encouraged to be included complementary to the central dataset. Importantly, it is advisable to incorporate traffic intensity variables if they are locally available. Traffic variables have been once reported for modeling in Tianjin, China[88]. Whereas, traffic intensity data are generally difficult to obtain in other Chinese cities which will constrain the prediction ability of LUR models. An alternative approach includes providing manually traffic counts on the nearest road of sampling sites in urban settings[91]. This however is not feasible for studies involving large study populations. Moreover, the problem of accumulation effect of air pollution in street canyon is more worthy of attention in many Chinese cities due to rapid urbanization and high buildings compared to some of the European study areas included in ESCAPE. The obstruction of air flow in street canyon can be quantified by GIS based approaches such as the SkyView factor (a measure of the total fraction of visible sky) which are able to improve the prediction ability of the LUR model[92]. A difference from the Chinese situation is that many residents are living in high rise apartment buildings in the city center. Since the concentration of air pollution decreases from lower to higher elevation[93], an appropriate algorithm is needed to take this factor into account. Another encountered challenge is the rapid undergoing urbanization in many Chinese cities since the past decades which may violate the assumption of stability over time for LUR models for long-term health study. Therefore, use of LUR modeling is more suitable in prospective than in retrospective applications. Developing LUR models combining a temporal term by means of satellite data or national network might overcome such shortcomings in regional or national scale studies in the future.

Fifthly, understanding and assessing the role of exposure measurement error in health effects assessment are central issues for the design and implementation of health effect cohort studies[94]. This is essentially important in Chinese cities where exposure measurement errors in air pollution epidemiology come in various dominant forms such as data quality, fewer predictor resources and more complex city structure than those in the ESCAPE study areas. Furthermore, exposure to indoor emissions may be more obvious in Chinese homes than in western homes[95]. Therefore, validity of LUR models are needed especially for pollutants with significant indoor sources. Assessing the predictive power of LUR model should rely on agreement with measurements at independent validation sets.

Sixthly, it is promising to explore transferability of LUR model based on large

multi-cities dataset to other Chinese areas. The uneven urbanization facilitates highly centered available resources in mega-cities in the east coast. The situation is similar to the United States where national monitoring networks are centered on the east and west coasts. It is desirable to develop a countrywide LUR model based on accessible monitoring sites of big cities to predict air pollution concentrations in small towns where no measurements are available. Moreover, developing a national-wide LUR incorporation with satellite data have been demonstrated as a useful tool to greatly improve the model (R^2)[18, 43] and therefore is applicable to the areas with less monitoring resources.

Seventhly, modeling long-term exposure to PM compositions is possible if relevant predictor variables are available and accessible. This allows identifying source-related health effects which is helpful for policy makers to regulate air pollution. In China, short-term studies have suggested that $PM_{2.5}$ constituents from the combustion of fossil fuel may have an important influence on the mortality outcomes attributable to $PM_{2.5}$[96]. However, long-term study on health effects of air pollution is still a big challenge due to lack of exposure measures. There is a need to explore a proper algorithm of LUR modeling for PM compositions in China.

Overall, LUR modeling is a promising technique to satisfy the urgent needs of assessing long-term health effects of exposure to ambient air pollution in China. The experiences from the ESCAPE project will provide precious guidance to facilitate its adaption in China. Nevertheless, developing a local satisfactory LUR model is still challenging and needs more practical experiences in China.

Conclusion

Our study is among the first empirical LUR studies that has systematically compared results of LOOCV and HV evaluations using training and test sets of varying sample sizes and is also the first that evaluated particulate matter LUR models in such a large number of study areas (N=20) within the framework of the ESCAPE study. Our study with the Dutch dataset suggested that truly independent evaluation data are especially useful when LUR models are developed from small training sets where we have shown the adjusted model and LOOCV R^2s to deviate most from the hold out validation R^2s.

This was supported with the ESCAPE models showing that for a wide range of study areas and pollutants including NO_2, $PM_{2.5}$ absorbance and PM_{10} Cu, model and LOOCV R^2 from land use regression models based on relatively small training sets overestimate predictive ability in independent test sets. Despite of this, in most areas LUR models still explained a substantial fraction of the spatial variation measured at independent sites.

European LUR models for NO_2 and $PM_{2.5}$ absorbance were found to have reasonable power to predict spatial variations of these components in areas not used for model building. For $PM_{2.5}$, prediction R^2s were moderate for intra-urban variation.

For the first time, we have attempted to develop land use regression models for the spatial variation of long term ambient concentrations of elemental composition in the PM_{10} and $PM_{2.5}$ fractions. Good LUR models were developed for Cu, Fe and

Zn in both PM_{10} and $PM_{2.5}$ fractions. Moderate models were developed for S, Si, Ni, V, K for PM_{10} and V and S for $PM_{2.5}$. Models for the elements K, Ni and Si in the $PM_{2.5}$ fraction performed poorest.

Our study provided the first result that no associations were found between long-term exposure to these 8 elemental constituents of particles (Cu, Fe, K, Ni, S, Si, V and Zn in $PM_{2.5}$ and PM_{10}) and total cardiovascular mortality based on 19 European cohorts. Most of the hazard ratios were close to unity.

References

1. Johnson, M.; Isakov, V.; Touma, J. S.; Mukerjee, S.; Ozkaynak, H., Evaluation of land-use regression models used to predict air quality concentrations in an urban area. *Atmospheric Environment* **2010,** *44*, (30), 3660-3668.
2. Basagana, X.; Rivera, M.; Aguilera, I.; Agis, D.; Bouso, L.; Elosua, R.; et al., Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment* **2012,** *54*, 634-642.
3. Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; et al., Development of NO_2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. *Atmospheric Environment* **2013,** *72*, 10-23.
4. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; et al., A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* **2008,** *42*, (33), 7561-7578.
5. Aguilera, I.; Sunyer, J.; Fernandez-Patier, R.; Hoek, G.; -Alfaro, A. A.; Meliefste, K.; et al., Estimation of outdoor NOx, NO_2, and BTEX exposure in a cohort of pregnant women using land use regression modeling. *Environmental Science & Technology* **2008,** *42*, (3), 815-821.
6. Henderson, S. B.; Beckerman, B.; Jerrett, M.; Brauer, M., Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology* **2007,** *41*, (7), 2422-2428.
7. Beelen, R.; Hoek, G.; Fischer, P.; van den Brandt, P. A.; Brunekreef, B., Estimated long-term outdoor air pollution concentrations in a cohort study. *Atmospheric Environment* **2007,** *41*, (7), 1343-1358.
8. Hoek, G.; Beelen, R.; Kos, G.; Dijkema, M.; van der Zee, S. C.; Fischer, P. H.; et al., Land Use Regression Model for Ultrafine Particles in Amsterdam. *Environmental Science & Technology* **2011,** *45*, (2), 622-628.
9. Crouse, D. L.; Goldberg, M. S.; Ross, N. A., A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in Montreal, Canada. *Atmospheric Environment* **2009,** *43*, (32), 5075-5084.
10. Briggs, D., The role of GIS: Coping with space (and time) in air pollution exposure assessment. *Journal of Toxicology and Environmental Health-Part a-Current Issues* **2005,** *68*, (13-14), 1243-1261.
11. Hoek, G.; Fischer, P.; Van Den Brandt, P.; Goldbohm, S.; Brunekreef, B., Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J Expo Anal Environ Epidemiol* **2001,** *11*, (6), 459-69.
12. Eeftens, M.; Beelen, R.; de Hoogh, K.; Bellander, T.; Cesaroni, G.; Cirach, M.; et al., Development of Land Use Regression Models for $PM_{2.5}$, $PM_{2.5}$ Absorbance, PM_{10} and PMcoarse in 20 European Study Areas; Results of the ESCAPE Project. *Environmental Science & Technology* **2012,** *46*, (20), 11195-11205.
13. de Hoogh, K.; Wang, M.; Adam, M.; Badaloni, C.; Beelen, R.; Birk, M.; et al., Development of land use regression models for particle composition in twenty study areas in europe. *Environ Sci Technol* **2013,** *47*, (11), 5778-86.
14. Bergen, S.; Sheppard, L.; Sampson, P. D.; Kim, S. Y.; Richards, M.; Vedal, S.; et al., A National Prediction Model for PM Component Exposures and Measurement Error-Corrected Health Effect Inference. *Environ Health Perspect* **2013**.

15. Li, L. F.; Wu, J.; Wilhelm, M.; Ritz, B., Use of generalized additive models and cokriging of spatial residuals to improve land-use regression estimates of nitrogen oxides in Southern California. *Atmospheric Environment* **2012,** *55*, 220-228.
16. Su, J. G.; Jerrett, M.; Beckerman, B., A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. *Science of the Total Environment* **2009,** *407*, (12), 3890-3898.
17. Su, J. G.; Jerrett, M.; Beckerman, B.; Wilhelm, M.; Ghosh, J. K.; Ritz, B., Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. *Environmental Research* **2009,** *109*, (6), 657-670.
18. Beckerman, B. S.; Jerrett, M.; Serre, M.; Martin, R. V.; Lee, S. J.; van Donkelaar, A.; et al., A Hybrid Approach to Estimating National Scale Spatiotemporal Variability of $PM_{2.5}$ in the Contiguous United States. *Environ Sci Technol* **2013,** *47*, (13), 7233-41.
19. Beckerman, B. S.; Jerrett, M.; Martin, R. V.; van Donkelaar, A.; Ross, Z.; Burnett, R. T., Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in California. *Atmospheric Environment* **2013,** *77*, 172-177.
20. Mercer, L. D.; Szpiro, A. A.; Sheppard, L.; Lindstrom, J.; Adar, S. D.; Allen, R. W.; et al., Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Atmospheric Environment* **2011,** *45*, (26), 4412-4420.
21. Kanaroglou, P. S.; Jerrett, M.; Morrison, J.; Beckerman, B.; Arain, M. A.; Gilbert, N. L.; et al., Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. *Atmospheric Environment* **2005,** *39*, (13), 2399-2409.
22. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; et al., A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology* **2005,** *15*, (2), 185-204.
23. Hoek, G.; Krishnan, R. M.; Beelen, R.; Peters, A.; Ostro, B.; Brunekreef, B.; et al., Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environ Health* **2013,** *12*, (1), 43.
24. Bellander, T.; Berglind, N.; Gustavsson, P.; Jonson, T.; Nyberg, F.; Pershagen, G.; et al., Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environmental Health Perspectives* **2001,** *109*, (6), 633-639.
25. Liu, L. J. S.; Tsai, M. Y.; Keidel, D.; Gemperli, A.; Ineichen, A.; Hazenkamp-von Arx, M.; et al., Long-term exposure models for traffic related NO_2 across geographically diverse areas over separate years. *Atmospheric Environment* **2012,** *46*, 460-471.
26. Gulliver, J.; de Hoogh, K.; Fecht, D.; Vienneau, D.; Briggs, D., Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. *Atmospheric Environment* **2011,** *45*, (39), 7072-7080.
27. Beelen, R.; Voogt, M.; Duyzer, J.; Zandveld, P.; Hoek, G., Comparison of the performances of land use regression modelling and dispersion modelling in estimating small-scale variations in long-term air pollution concentrations in a Dutch urban area. *Atmospheric Environment* **2010,** *44*, (36), 4614-4621.
28. Marshall, J. D.; Nethery, E.; Brauer, M., Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmospheric Environment* **2008,** *42*, (6), 1359-1369.
29. Briggs, D. J.; de Hoogh, C.; Guiliver, J.; Wills, J.; Elliott, P.; Kingham, S.; et al., A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of the Total Environment* **2000,** *253*, (1-3), 151-167.
30. Wang, M.; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B., Systematic Evaluation of Land Use Regression Models for NO_2. *Environ Sci Technol* **2012,** *46*, (8), 4481-4489.
31. Dijkema, M. B.; Gehring, U.; van Strien, R. T.; van der Zee, S. C.; Fischer, P.; Hoek, G.; et al., A comparison of different approaches to estimate small-scale spatial variation in outdoor NO concentrations. *Environ Health Perspect* **2011,** *119*, (5), 670-5.
32. Cyrys, J.; Hochadel, M.; Gehring, U.; Hoek, G.; Diegmann, V.; Brunekreef, B.; et al., GIS-based estimation of exposure to particulate matter and NO_2 in an urban area: stochastic versus

dispersion modeling. *Environ Health Perspect* **2005,** *113*, (8), 987-92.
33. Cesaroni, G.; Badaloni, C.; Gariazzo, C.; Stafoggia, M.; Sozzi, R.; Davoli, M.; et al., Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ Health Perspect* **2013,** *121*, (3), 324-31.
34. Downs, S. H.; Schindler, C.; Liu, L. J.; Keidel, D.; Bayer-Oglesby, L.; Brutsche, M. H.; et al., Reduced exposure to PM_{10} and attenuated age-related decline in lung function. *N Engl J Med* **2007,** *357*, (23), 2338-47.
35. Brauer, M.; Lencar, C.; Tamburic, L.; Koehoorn, M.; Demers, P.; Karr, C., A cohort study of traffic-related air pollution impacts on birth outcomes. *Environ Health Perspect* **2008,** *116*, (5), 680-6.
36. Gehring, U.; Wijga, A. H.; Brauer, M.; Fischer, P.; de Jongste, J. C.; Kerkhof, M.; et al., Traffic-related air pollution and the development of asthma and allergies during the first 8 years of life. *Am J Respir Crit Care Med* **2010,** *181*, (6), 596-603.
37. Ryan, P. H.; Lemasters, G. K.; Biswas, P.; Levin, L.; Hu, S.; Lindsey, M.; et al., A comparison of proximity and land use regression traffic exposure models and wheezing in infants. *Environ Health Perspect* **2007,** *115*, (2), 278-84.
38. Eeftens, M.; Beelen, R.; Fischer, P.; Brunekreef, B.; Meliefste, K.; Hoek, G., Stability of measured and modelled spatial contrasts in NO(2) over time. *Occup Environ Med* **2011,** *68*, (10), 765-70.
39. Gulliver, J.; de Hoogh, K.; Hansell, A.; Vienneau, D., Development and Back-Extrapolation of NO_2 Land Use Regression Models for Historic Exposure Assessment in Great Britain. *Environ Sci Technol* **2013,** *47*, (14), 7804-11.
40. Cesaroni, G.; Porta, D.; Badaloni, C.; Stafoggia, M.; Eeftens, M.; Meliefste, K.; et al., Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. *Environmental Health* **2012,** *11*.
41. Madsen, C.; Gehring, U.; Haberg, S. E.; Nafstad, P.; Meliefste, K.; Nystad, W.; et al., Comparison of land-use regression models for predicting spatial NO(x) contrasts over a three year period in Oslo, Norway. *Atmospheric Environment* **2011,** *45*, (21), 3576-3583.
42. Wang, R. R.; Henderson, S. B.; Sbihi, H.; Allen, R. W.; Brauer, M., Temporal stability of land use regression models for traffic-related air pollution. *Atmospheric Environment* **2013,** *64*, 312-319.
43. Kloog, I.; Nordio, F.; Coull, B. A.; Schwartz, J., Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal $PM_{(2.5)}$ exposures in the Mid-Atlantic states. *Environ Sci Technol* **2012,** *46*, (21), 11913-21.
44. Hystad, P.; Demers, P. A.; Johnson, K. C.; Brook, J.; van Donkelaar, A.; Lamsal, L.; et al., Spatiotemporal air pollution exposure assessment for a Canadian population-based lung cancer case-control study. *Environ Health* **2012,** *11*, 22.
45. Kloog, I.; Ridgway, B.; Koutrakis, P.; Coull, B. A.; Schwartz, J. D., Long- and short-term exposure to $PM_{2.5}$ and mortality: using novel exposure models. *Epidemiology* **2013,** *24*, (4), 555-61.
46. Molter, A.; Lindley, S.; de Vocht, F.; Simpson, A.; Agius, R., Modelling air pollution for epidemiologic research - Part I: A novel approach combining land use regression and air dispersion. *Science of the Total Environment* **2010,** *408*, (23), 5862-5869.
47. Levy, J. I., Impact of residential nitrogen dioxide exposure on personal exposure: An international study. *Journal of the Air & Waste Management Association* **1998,** *48*, (6), 553-560.
48. Zota, A.; Adamkiewicz, G.; Levy, J. I.; Spengler, J. D., Ventilation in public housing: implications for indoor nitrogen dioxide concentrations. *Indoor Air* **2005,** *15*, (6), 393-401.
49. Lai, H. K.; Bayer-Oglesby, L.; Colvile, R.; Gotschi, T.; Jantunen, M. J.; Kunzli, N.; et al., Determinants of indoor air concentrations of $PM_{2.5}$, black smoke and NO_2 in six European cities (EXPOLIS study). *Atmospheric Environment* **2006,** *40*, (7), 1299-1313.
50. Janssen, N. A.; Schwartz, J.; Zanobetti, A.; Suh, H. H., Air conditioning and source-specific particles as modifiers of the effect of PM(10) on hospital admissions for heart and lung disease. *Environ Health Perspect* **2002,** *110*, (1), 43-9.
51. Van Roosbroeck, S.; Hoek, G.; Meliefste, K.; Janssen, N. A. H.; Brunekreef, B., Validity of residential traffic intensity as an estimate of long-term personal exposure to traffic-related air pollution among adults. *Environmental Science & Technology* **2008,** *42*, (4), 1337-1344.
52. Van Roosbroeck, S.; Li, R.; Hoek, G.; Lebret, E.; Brunekreef, B.; Spiegelman, D., Traffic-related

outdoor air pollution and respiratory symptoms in children: the impact of adjustment for exposure measurement error. *Epidemiology* **2008,** *19*, (3), 409-16.
53. Nethery, E.; Teschke, K.; Brauer, M., Predicting personal exposure of pregnant women to traffic-related air pollutants. *Science of the Total Environment* **2008,** *395*, (1), 11-22.
54. Sahsuvaroglu, T.; Su, J. G.; Brook, J.; Burnett, R.; Loeb, M.; Jerrett, M., Predicting personal nitrogen dioxide exposure in an elderly population: integrating residential indoor and outdoor measurements, fixed-site ambient pollution concentrations, modeled pollutant levels, and time-activity patterns. *J Toxicol Environ Health A* **2009,** *72*, (23), 1520-33.
55. Hannam, K.; McNamee, R.; De Vocht, F.; Baker, P.; Sibley, C.; Agius, R., A comparison of population air pollution exposure estimation techniques with personal exposure estimates in a pregnant cohort. *Environ Sci Process Impacts* **2013,** *15*, (8), 1562-72.
56. Montagne, D.; Hoek, G.; Nieuwenhuijsen, M.; Lanki, T.; Pennanen, A.; Portella, M.; et al., Agreement of Land Use Regression Models with Personal Exposure Measurements of Particulate Matter and Nitrogen Oxides Air Pollution. *Environ Sci Technol* **2013**.
57. Cohen, M. A.; Adar, S. D.; Allen, R. W.; Avol, E.; Curl, C. L.; Gould, T.; et al., Approach to estimating participant pollutant exposures in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Environ Sci Technol* **2009,** *43*, (13), 4687-93.
58. Levy, J. I.; Clougherty, J. E.; Baxter, L. K.; Houseman, E. A.; Paciorek, C. J., Evaluating heterogeneity in indoor and outdoor air pollution using land-use regression and constrained factor analysis. *Res Rep Health Eff Inst* **2010**, (152), 5-80; discussion 81-91.
59. Lebret, E.; Briggs, D.; van Reeuwijk, H.; Fischer, P.; Smallbone, K.; Harssema, H.; et al., Small area variations in ambient NO_2 concentrations in four European areas. *Atmospheric Environment* **2000,** *34*, (2), 177-185.
60. Weijers, E. P.; Khlystov, A. Y.; Kos, G. P. A.; Erisman, J. W., Variability of particulate matter concentrations along roads and motorways determined by a moving measurement unit. *Atmospheric Environment* **2004,** *38*, (19), 2993-3002.
61. Eeftens, M., Tsai, M., Ampe, C., Anwande, B., Beelen, R.M.J., Cesaroni, G., et al., Variation of $PM_{2.5}$, PM_{10}, $PM_{2.5}$ absorbance and PMcoarse concentrations between and within 20 European study areas – results of the ESCAPE project. *Atmospheric Environment* **2012**.
62. Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R.M.J., et al., Variation of NO_2 and NO_x concentrations between and within 36 European study areas: results from the ESCAPE study. *Atmospheric Environment* **2012**.
63. Jerrett, M.; Burnett, R. T.; Ma, R.; Pope, C. A., 3rd; Krewski, D.; Newbold, K. B.; et al., Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* **2005,** *16*, (6), 727-36.
64. Hoek, G.; Brunekreef, B.; Goldbohm, S.; Fischer, P.; van den Brandt, P. A., Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *Lancet* **2002,** *360*, (9341), 1203-9.
65. Miller, K. A.; Siscovick, D. S.; Sheppard, L.; Shepherd, K.; Sullivan, J. H.; Anderson, G. L.; et al., Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med* **2007,** *356*, (5), 447-58.
66. Dvonch, J. T.; Kannan, S.; Schulz, A. J.; Keeler, G. J.; Mentz, G.; House, J.; et al., Acute effects of ambient particulate matter on blood pressure: differential effects across urban communities. *Hypertension* **2009,** *53*, (5), 853-9.
67. Bell, M. L.; Ebisu, K.; Peng, R. D., Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research. *J Expo Sci Environ Epidemiol* **2011,** *21*, (4), 372-84.
68. Sheppard, L.; Burnett, R. T.; Szpiro, A. A.; Kim, S. Y.; Jerrett, M.; Pope, C. A., 3rd; et al., Confounding and exposure measurement error in air pollution epidemiology. *Air Qual Atmos Health* **2012,** *5*, (2), 203-216.
69. Armstrong, B. G., Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med* **1998,** *55*, (10), 651-6.
70. Gilliland, F.; Avol, E.; Kinney, P.; Jerrett, M.; Dvonch, T.; Lurmann, F.; et al., Air pollution exposure assessment for epidemiologic studies of pregnant women and children: lessons learned from the Centers for Children's Environmental Health and Disease Prevention Research. *Environ Health Perspect* **2005,** *113*, (10), 1447-54.

71. Stanek, L. W.; Sacks, J. D.; Dutton, S. J.; Dubois, J. J. B., Attributing health effects to apportioned components and sources of particulate matter: An evaluation of collective results. *Atmospheric Environment* **2011,** *45*, (32), 5655-5663.
72. Beelen, R.; Stafoggia, M.; Raaschou-Nielsen, O.; Andersen, Z. J.; Xun, W.; Katsouyanni, K.; et al., Long-term exposure to air pollution and cardiovascular mortality: an analysis of 22 European cohorts within the ESCAPE project. *Submitted* **2013**.
73. Ostro, B.; Feng, W. Y.; Broadwin, R.; Green, S.; Lipsett, M., The effects of components of fine particulate air pollution on mortality in california: results from CALFINE. *Environ Health Perspect* **2007,** *115*, (1), 13-9.
74. Zhou, J.; Ito, K.; Lall, R.; Lippmann, M.; Thurston, G., Time-series analysis of mortality effects of fine particulate matter components in Detroit and Seattle. *Environ Health Perspect* **2011,** *119*, (4), 461-6.
75. Ito, K.; Mathes, R.; Ross, Z.; Nadas, A.; Thurston, G.; Matte, T., Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. *Environ Health Perspect* **2011,** *119*, (4), 467-73.
76. Bell, M. L.; Ebisu, K.; Peng, R. D.; Samet, J. M.; Dominici, F., Hospital admissions and chemical composition of fine particle air pollution. *Am J Respir Crit Care Med* **2009,** *179*, (12), 1115-20.
77. Lippmann, M.; Ito, K.; Hwang, J. S.; Maciejczyk, P.; Chen, L. C., Cardiovascular effects of nickel in ambient air. *Environ Health Perspect* **2006,** *114*, (11), 1662-9.
78. Ostro, B.; Lipsett, M.; Reynolds, P.; Goldberg, D.; Hertz, A.; Garcia, C.; et al., Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study. *Environ Health Perspect* **2010,** *118*, (3), 363-9.
79. Happo, M. S.; Hirvonen, M. R.; Halinen, A. I.; Jalava, P. I.; Pennanen, A. S.; Sillanpaa, M.; et al., Chemical compositions responsible for inflammation and tissue damage in the mouse lung by coarse and fine particulate samples from contrasting air pollution in Europe. *Inhal Toxicol* **2008,** *20*, (14), 1215-31.
80. Happo, M. S.; Hirvonen, M. R.; Halinen, A. I.; Jalava, P. I.; Pennanen, A. S.; Sillanpaa, M.; et al., Seasonal variation in chemical composition of size-segregated urban air particles and the inflammatory activity in the mouse lung. *Inhal Toxicol* **2010,** *22*, (1), 17-32.
81. Gerlofs-Nijland, M. E.; Dormans, J. A.; Bloemen, H. J.; Leseman, D. L.; John, A.; Boere, F.; Kelly, F. J.; Mudway, I. S.; et al., Toxicity of coarse and fine particulate matter from sites with contrasting traffic profiles. *Inhal Toxicol* **2007,** *19*, (13), 1055-69.
82. Mostofsky, E.; Schwartz, J.; Coull, B. A.; Koutrakis, P.; Wellenius, G. A.; Suh, H. H.; et al., Modeling the association between particle constituents of air pollution and health outcomes. *Am J Epidemiol* **2012,** *176*, (4), 317-26.
83. Ostro, B.; Tobias, A.; Querol, X.; Alastuey, A.; Amato, F.; Pey, J.; et al., The effects of particulate matter sources on daily mortality: a case-crossover study of Barcelona, Spain. *Environ Health Perspect* **2011,** *119*, (12), 1781-7.
84. Shang, Y.; Sun, Z.; Cao, J.; Wang, X.; Zhong, L.; Bi, X.; et al., Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. *Environ Int* **2013,** *54*, 100-11.
85. Cao, J.; Yang, C.; Li, J.; Chen, R.; Chen, B.; Gu, D.; et al., Association between long-term exposure to outdoor air pollution and mortality in China: a cohort study. *J Hazard Mater* **2011,** *186*, (2-3), 1594-600.
86. Matsui, H.; Koike, M.; Takegawa, N.; Kita, K.; Miyazaki, Y.; Hu, M.; et al., Spatial and temporal variations of aerosols around Beijing in summer 2006: Model evaluation and source apportionment. *Journal of Geophysical Research* **2009,** *114*, D00G13.
87. Chen, L.; Du, S. Y.; Bai, Z. P.; Kong, S. F.; You, Y.; Han, B.; et al., Application of land use regression for estimating concentrations of major outdoor air pollutants in Jinan, China. *Journal of Zhejiang University-Science A* **2010,** *11*, (11), 857-867.
88. Chen, L.; Wang, Y. M.; Li, P. W.; Ji, Y. Q.; Kong, S. F.; Li, Z. Y.; et al., A land use regression model incorporating data on industrial point source pollution. *Journal of Environmental Sciences-China* **2012,** *24*, (7), 1251-1258.
89. Chen, L.; Bai, Z. P.; Kong, S. F.; Han, B.; You, Y.; Ding, X.; et al., A land use regression for predicting NO_2 and PM_{10} concentrations in different seasons in Tianjin region, China. *Journal of Environmental Sciences* **2010,** *22*, (9), 1364-1373.

90. Chen, R.; Kan, H.; Chen, B.; Huang, W.; Bai, Z.; Song, G.; et al., Association of particulate air pollution with daily mortality: the China Air Pollution and Health Effects Study. *American Journal of Epidemiology* **2012,** *175*, (11), 1173-81.
91. Rivera, M.; Basagana, X.; Aguilera, I.; Agis, D.; Bouso, L.; Foraster, M.; et al., Spatial distribution of ultrafine particles in urban settings: A land use regression model. *Atmospheric Environment* **2012,** *54*, 657-666.
92. Eeftens, M.; Beekhuizen, J.; Beelen, R.; Wang, M.; Vermeulen, R.; Brunekreef, B.; et al., Quantifying urban street configuration for improvements in air pollution models. *Atmospheric Environment* **2013,** *72*, 1-9.
93. Maletto, A.; McKendry, I. G.; Strawbridge, K. B., Profiles of particulate matter size distributions using a balloonborne lightweight aerosol spectrometer in the planetary boundary layer. *Atmospheric Environment* **2003,** *37*, 661-670.
94. Jerrett, M.; Finkelstein, M., Geographies of risk in studies linking chronic air pollution exposure to health outcomes. *J Toxicol Environ Health A* **2005,** *68*, (13-14), 1207-42.
95. Zhang, Y.; Mo, J.; Weschler, C. J., Reducing health risks from indoor exposures in rapidly developing urban china. *Environ Health Perspect* **2013,** *121*, (7), 751-5.
96. Cao, J.; Xu, H.; Xu, Q.; Chen, B.; Kan, H., Fine particulate matter constituents and cardiopulmonary mortality in a heavily polluted Chinese city. *Environ Health Perspect* **2012,** *120*, (3), 373-8.

Summary

Exposure assessment is one of the key issues for health effect estimates in environmental epidemiology. Recent interest has increased in exposure modeling incorporating Geographic Information System (GIS) data to capture small-scale spatial variability in air pollution concentrations. Land use regression (LUR) modeling is one of the most popular models due to the high resolution mapping technique. Even although LUR technique has been studied extensively in the past decades, performances of LUR models associated with prediction ability to outdoor exposure have not been well explored. Furthermore, transferability of LUR models from city to city has been investigated, but little is known about the performance of models based on large numbers of monitoring sites covering a large area.

Ambient $PM_{2.5}$ and PM_{10} (particle dynamic diameter <2.5 μm and 10μm) represent a heterogeneous mixture of constituents from diverse sources e.g. fossil fuel combustion, biomass burning and human activity. However, it is still unclear which PM constituents are associated with higher risks. Cardiovascular (CVD) mortality contributes a majority of the all-cause mortality in many studies. Several studies showed evidence of acute effects of PM components on CVD mortality, but results were inconsistent. Very few studies have assessed health effects related to long-term exposure to elemental composition. Lack of spatially resolved elemental composition measurement data and a lack of models for elemental composition have contributed to this gap.

The multi-center European Study of Cohorts for Air Pollution Effects (ESCAPE) aims at quantifying long-term impacts of air pollution on diverse health outcomes.

This thesis is within the framework of the ESCAPE project and with specific aims:

1. To evaluate the performances of LUR models in terms of model fit and prediction ability
2. To develop LUR models for particle compositions
3. To estimate associations between long-term exposure to particle compositions and cardiovascular mortality

Evaluation is an essential part of Land Use Regression (LUR) model development. In chapter 2, we developed LUR models for nitrogen dioxide (NO_2) using measurements conducted at 144 sampling sites in the Netherlands. Sites were randomly divided into training datasets with a size of 24, 36, 48, 72, 96, 108 and 120 sites. LUR models were evaluated using (1) internal "leave-one-out-cross-validation (LOOCV)" within the training datasets and (2) external "hold-out" validation (HV) against independent test datasets. LUR model performance for NO_2 varies with the number of training sites. Hold-out validation (HV) R^2 was lower than the corresponding cross-validation R^2, especially for the smallest training sets. Our study suggested that truly independent evaluation data are especially useful when LUR models are developed from small training sets where we have shown the adjusted model and LOOCV R^2s to deviate most from the hold out validation R^2s. In our specific application, models based on as few as 24 training sites, however, achieved acceptable hold out validation R^2s of, on average, 0.60.

Following on the findings in chapter 2, LUR models have been further developed for NO_2, $PM_{2.5}$ absorbance and Copper (Cu) in PM_{10} based on 20 sites in each of the 20 study areas of the ESCAPE project with PM data (Chapter 3). Models were

evaluated with LOOCV and "hold-out evaluation (HEV)" using the correlation of predicted NO_2 or PM concentrations with measured NO_2 concentrations at the 20 additional NO_2 sites in each area. Our results confirm that the predictive ability of LUR models based on relatively small training sets is overestimated by the LOOCV R^2s. Nevertheless, in most areas LUR models still explained a substantial fraction of the variation of concentrations measured at independent sites.

In chapter 4, we evaluated LUR models for NO_2 and Particulate Matter ($PM_{2.5}$, $PM_{2.5}$ absorbance) by combining standardized measurement data from 17 (PM) and 23 (NO_2) ESCAPE study areas across 14 European countries for PM and NO_2. Additionally, we investigated the transferability of the models by successively excluding each study area from model building and applying developed models to these excluded areas to test the model performance. Using a large dataset from 23 European study areas, we were able to develop LUR models for NO_2 and PM metrics that predicted measurements made at independent sites and areas reasonably well. For $PM_{2.5}$, prediction R^2s were moderate for intra-urban variation. This finding is useful for assessing exposure in health studies conducted in areas where no measurements were conducted.

In chapter 5, LUR models for eight a priori selected elements; copper (Cu), iron (Fe), potassium (K), nickel (Ni), sulphur (S), silicon (Si), vanadium (V) and zinc (Zn) were developed. Good models were developed for Cu, Fe and Zn in both fractions (PM_{10} and $PM_{2.5}$) explaining on average between 67 and 79% of the concentration variance (R^2) with a large variability between areas. Traffic variables were the dominant predictors, reflecting non-tailpipe emissions. Models for V and S in the PM_{10} and $PM_{2.5}$ fractions and Si, Ni and K in the PM_{10} fraction performed moderately with R^2 ranging from 50 to 61%. Si, NI and K models for $PM_{2.5}$ performed poorest with R^2 under 50%. The LUR models are used to estimate exposures to elemental composition in the health studies involved in ESCAPE.

Chapter 6 focuses on the association between long-term exposure to these eight elemental constituents (Cu, Fe, K, Ni, S, Si, V and Zn) of particles and total cardiovascular mortality within the framework of the ESCAPE and TRANSPHORM projects. Residential annual average exposure to elemental constituents within $PM_{2.5}$ and PM_{10} was estimated by LUR models. Cohort-specific analyses were conducted using Cox proportional hazards models with a standardized protocol. Random-effects meta-analysis was used to combine pooled effect estimates for 19 European cohorts. Our study provided the first result that no associations were found between long-term exposure to these eight elemental constituents of particles and total cardiovascular mortality. Most of the hazard ratios were close to unity.

Samenvatting

Blootstellingsschatting is één van de belangrijkste kwesties voor de schatting van gezondheidseffecten in milieu-epidemiologie. Recent is de belangstelling toegenomen om Geografisch Informatie Systeem (GIS) data op te nemen in de blootstellingsmodellering om zo kleinschalige ruimtelijke variabiliteit van luchtverontreinigingsconcentraties vast te stellen. Land Use Regressie (LUR) modellering is een van de meest populaire modellen vanwege de hoge resolutie van deze zogeheten mapping techniek. Hoewel de LUR techniek uitgebreid bestudeerd is in de afgelopen decennia, is het toepassen van de LUR modellen voor het voorspellen van buitenlucht concentraties niet goed onderzocht. Daarnaast is de toepassing van LUR modellen van stad tot stad onderzocht, maar er is weinig bekend over de prestaties van de modellen gebaseerd op een groot aantal meetlocaties in een groot gebied.

$PM_{2.5}$ en PM_{10}, deeltjes in de omgeving met een zogeheten dynamische diameter kleiner dan 2.5 µm en 10 µm, vormen een heterogeen mengsel van bestanddelen uit diverse bronnen zoals verbranding van fossiele brandstoffen, biomassaverbranding en menselijke activiteiten. Het is echter nog onduidelijk welke elementen in fijn stof geassocieerd zijn met een hoger gezondheidsrisico. Cardiovasculaire (CVD) sterfte draagt sterk bij aan de totale mortaliteit in veel studies. Verschillende studies toonden acute effecten van PM componenten op CVD sterfte, maar de resultaten waren inconsistent. Zeer weinig studies hebben de gezondheidseffecten van lange-termijn blootstelling aan elementen in fijn stof onderzocht. Gebrek aan meetgegevens van PM componenten en een gebrek aan modellen om PM componenten te schatten hebben bijgedragen aan deze kloof.

De multi-center European Study of Cohorts for Air Pollution Effects (ESCAPE) is gericht op het kwantificeren van de effecten op lange termijn van luchtverontreiniging op diverse gezondheidsuitkomsten.

Dit proefschrift is geschreven in het kader van het ESCAPE-project en heeft de specifieke doelstellingen:

1. De prestaties van LUR modellen evalueren in termen van model fitting en het voorspellend vermogen
2. LUR modellen ontwikkelen voor PM componenten
3. Schattingen maken voor de associatie tussen langdurige blootstelling aan PM componenten en cardiovasculaire mortaliteit

Evaluatie is een essentieel onderdeel van LUR modelontwikkeling. In hoofdstuk 2 hebben we LUR modellen ontwikkeld voor stikstofdioxide (NO_2), gebruikmakend van metingen uitgevoerd op 144 meetpunten in Nederland. De meetpunten werden willekeurig verdeeld in datasets met een omvang van 24, 36, 48, 72, 96, 108 en 120 meetpunten. LUR modellen werden geëvalueerd met behulp van (1) interne "leave-one-out-cross-validatie (LOOCV)" binnen de datasets en (2) externe "hold-out" validatie (HV) in vergelijking met onafhankelijke test-datasets. LUR model prestaties voor NO_2 varieerden met het aantal meetpunten. De R^2 voor hold-out validatie was lager dan de overeenkomstige cross-validatie R^2, vooral voor de kleinste data sets. Onze studie suggereert dat werkelijk onafhankelijke evaluatiegegevens vooral waardevol zijn wanneer LUR modellen worden ontwikkeld voor kleine data sets, voor deze wijken de model en LOOCV R^2s het meeste af van de hold-out validatie R^2s. In onze studie hadden modellen op basis van 24 meetpunten echter aanvaardbare hold-out validatie R^2s van gemiddeld

0,60.

Naar aanleiding van de bevindingen in hoofdstuk 2 zijn de LUR modellen verder ontwikkeld voor NO_2, $PM_{2.5}$ absorptie en koper (Cu) in $PM_{10,}$ gebaseerd op 20 locaties in elk van de 20 studiegebieden met PM-gegevens binnen het ESCAPE-project (hoofdstuk 3). Modellen werden geëvalueerd met LOOCV en "hold-out evaluatie (HEV)" gebruikmakend van de correlatie van de voorspelde NO_2 of PM-concentraties met gemeten NO_2-concentraties op de 20 extra NO_2 locaties in elk studiegebied. Onze resultaten bevestigen dat het voorspellend vermogen van LUR modellen op basis van relatief kleine data sets wordt overschat door de LOOCV R^2s. Niettemin verklaren de LUR modellen in de meeste gebieden nog een aanzienlijke fractie van de variatie van de concentraties gemeten op onafhankelijke sites.

In hoofdstuk 4 hebben we LUR modellen geëvalueerd voor NO_2 en fijn stof ($PM_{2.5}$, $PM_{2.5}$-absorptie) door gestandaardiseerde meetgegevens te combineren van 17 (PM) en 23 (NO_2) ESCAPE studiegebieden in 14 Europese landen. Daarnaast hebben we de bruikbaarheid van de modellen onderzocht door achtereenvolgens elk studiegebied afzonderlijk uit te sluiten in de modelbouw en vervolgens de ontwikkelde modellen toe te passen op de uitgesloten gebieden om zo de prestaties van het model te testen. Met behulp van een grote dataset van 23 Europese studie gebieden waren we in staat om voor NO_2 en PM LUR modellen te ontwikkelen die metingen op onafhankelijke sites en gebieden redelijk goed voorspellen. Voor $PM_{2.5}$ waren de R^2s redelijk voorspellend voor de intra-stedelijke variant. Deze bevinding is nuttig voor het schatten van blootstelling in gezondheidsstudies in gebieden waar geen metingen werden uitgevoerd.

In hoofdstuk 5 worden LUR modellen voor acht a priori geselecteerde elementen, koper (Cu), ijzer (Fe), kalium (K), nikkel (Ni), zwavel (S), silicium (Si), vanadium (V) en zink (Zn) ontwikkeld. Goede modellen konden worden ontwikkeld voor Cu, Fe en Zn in beide fracties (PM_{10} en $PM_{2,5}$). Deze verklaren gemiddeld tussen 67% en 79% van de gemeten variantie met een grote variabiliteit tussen gebieden. Verkeersvariabelen waren de dominante voorspellers, wat de niet-uitlaat emissies van deze elementen weerspiegelt. Modellen voor V en S in de PM_{10} en $PM_{2,5}$ en voor Si, Ni en K in de PM_{10}-fractie gaven een matige R^2 van 50 tot 61%. Si, Ni en K-modellen voor $PM_{2.5}$ gaven de slechtste R^2 van onder de 50%. De LUR modellen worden gebruikt om de blootstelling aan PM componenten te schatten in de gezondheidsstudies in het ESCAPE-project.

Hoofdstuk 6 richt zich op de relatie tussen langdurige blootstelling aan deze acht PM componenten (Cu, Fe, K, Ni, S, Si, V en Zn) en totale cardiovasculaire mortaliteit zoals onderzocht in het kader van de ESCAPE en TRANSPHORM projecten. Residentiële jaargemiddelde blootstelling aan PM componenten werd geschat door LUR modellen. Cohort-specifieke analyses werden uitgevoerd met behulp van Cox proportionele hazard modellen met een gestandaardiseerd protocol. Random-effecten meta-analyse werd gebruikt om gepoolde effect schattingen te maken door 19 Europese cohorten te combineren. Onze studie heeft als eerste de associaties tussen langdurige blootstelling aan deze acht PM componenten en totale cardiovasculaire mortaliteit onderzocht. Er werden geen associaties gevonden, de meeste hazard ratio's waren dicht bij 1.

总结

暴露评估分析是环境流行病学健康效应研究的重要组成部分。近年来，越来越多的研究开始运用暴露评估模型与地理信息系统(Geographical Information System, GIS)相结合的方法模拟较小区域内的大气污染物空间分布。其中，利用地理信息系统的回归模型(Land Use Regression, LUR)由于其高空间分辨率而被广泛运用于健康效应研究。尽管如此， 关于 LUR 模型对环境暴露浓度的预测能力以及该模型结构是否适用于建模区域以外的区域的研究还相对很少。

大气环境中的 $PM_{2.5}$ 和 PM_{10}(颗粒物动力学直径小于 2.5 和 10 微米)是一种由非均相混合物所组成的颗粒物，其成分来源复杂，主要包括化石，生物质燃烧以及人为源排放等。然而，关于颗粒物成分的具体危害性了解还较少。大量研究表明心血管疾病死亡率占总死亡率比重最大。尽管少数研究证实颗粒物成分与心血管疾病的急性死亡率之间存在相关性，然而研究结果却很难统一。颗粒物成分致人群慢性健康效应研究则相对更少，主要原因在于缺少高空间分辨率的颗粒物成分数据以及合适的暴露评估模型。

ESCAPE 项目是一项欧洲多城市的空气污染健康效应的大型队列研究，其研究目的为定量描述空气污染物致多种慢性疾病的健康效应。

本论文研究主要基于 ESCAPE 项目框架内，其主要研究目的为：

1. 评估 LUR 模型的拟合及预测能力
2. 建立颗粒物成分的 LUR 模型
3. 研究颗粒物成分及心血管疾病的慢性死亡率相关性

模型性能评估是建模的核心。在第二章节，我们基于荷兰地区的 144 个采样点建立了二氧化氮(NO2)的 LUR 模型。采样点被随机分配用于建模和评估两部分，建模采样点数量由小到大依次为 24,36,48,72,96,108 和 120，剩余部分用于模型预测性能评估。LUR 模型性能评估方法分为：(1)以建模采样点为主的单点去除交叉验证(Leave-One-Out-Cross-Validation, LOOCV)；(2)以评估采样点为主的独立验证(Hold-out Valiation, HV)。结果显示： NO_2 模型性能随建模采样点数量变化而变化。HV R2 低于相同采样点数目的 LOOCV R^2，并且当建模采样点数量较少时其差距更加显著。该研究证明了采用独立的评估采样点能更加真实地反映 LUR 模型的预测能力。尤其当建模采样点数量较少时，HV R^2 的随机变化将更大。然而，即使 LUR 模型只基于 24 个采样点，其平均预测能力 HV R2 仍然高于 60%。

基于第二章节结论，我们在第三章节分别以 ESCAPE 项目框架下的 20 个城市(每个城市 20 个采样点)为基础，建立了 NO_2, $PM_{2.5}$ 的炭黑成分以及 PM_{10} 的铜元素为主的 LUR 模型。模型评估采用类似的单点去除交叉验证(LOOCV)和独立验证(HEV)方法。不同之处在于由于没有额外的颗粒物模型的评估采样点，我们采用 NO_2 的额外采样点替代评估。该研究证实了第二章节的结论，即 LUR 模型的预测性能(LOOCV R^2)在建模采样点数量较少时将被高估。然而，LUR 模型在大多数城市仍然具有较高的预测性。

在第四章节，我们合并了 14 个国家分别为 17 和 23 个城市的颗粒物和 NO_2 采样点以及相应的地理信息变量数据，建立和评估了基于该数据的 $PM_{2.5}$，炭黑和 NO_2 的 LUR 模型，并进一步研究了该模型在未建模区域的预测能力。该方法主要通过依次排除单个城市建模，并运用该模型预测被排除建模城市的污染物浓度。结果表明，

由于该研究基于多城市的大样本量数据，模型能够较好地预测没有参与建模的观测点及城市的污染物浓度。但对于 $PM_{2.5}$，大尺度模型预测城市区域内部空间分布能力弱于其他两种污染物。该模型可以帮助扩大健康研究范围至没有污染物数据的区域。

在第五章节，我们在 20 个城市分别建立了 $PM_{2.5}$ 和 PM_{10} 的颗粒物元素成分(铜、铁、钾、镍、硫、硅、钒和锌)的 LUR 模型。铜、铁、锌模型较好的拟合和解释了 67%至 79%的空间分布但不同城市的模型解释度差别较大。车流量作为最主要的模型变量反映了非机动车尾气排放的污染物变化。$PM_{2.5}$ 和 PM_{10} 中的钒、硫成分以及 PM_{10} 中的硅、镍、钾成分模型拟合解释了 50%至 61%的污染物空间变化。$PM_{2.5}$ 中的硅、镍和钾空间拟合解释度则少于 50%。该颗粒物成分模型将用于 ESCAPE 项目框架下的相关健康研究。

第六章节主要研究颗粒物八种金属元素成分(铜、铁、钾、镍、硫、硅、钒和锌) 长期暴露及心血管疾病死亡率的相关性。该研究由ESCAPE和TRANSPHORM项目资助。我们运用第五章节建立的LUR模型预测人群居住地室外颗粒物元素成分的年均浓度。队列研究分析采用ESCAPE统一制定的生存分析模型(Cox)并运用随机效应的Meta-analysis方法综合评估19个队列的分析结果。研究结果显示，颗粒物成分长期暴露与人群心血管疾病死亡率没有显著相关。

Affiliation of contributors

Affiation of contributors

Bert Brunekreef
Institute for Risk Assessment Sciences (IRAS), Division of Environmental Epidemiology, Utrecht University, Utrecht, The Netherlands
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

Gerard Hoek, Rob Beelen, Kees Meliefste, Marloes Eeftens
Institute for Risk Assessment Sciences (IRAS), Division of Environmental Epidemiology, Utrecht University, Utrecht, The Netherlands

Michiel L Bots
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

Paul Fischer, Bas Bueno-de-Mesquita
National Institute of Public Health and the Environment, Bilthoven, the Netherlands

Menno Keuken
TNO, Netherlands Organisation for Applied Scientific Research, Utrecht, The Netherlands

David Olsson, Lars Modig
Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden

Charlotte Eriksson, Claes-Göran Östensson, Göran Pershagen, Laura Fratiglioni, Michal Korek, Nancy Pedersen, Tom Bellander, Ulf De Faire
Karolinska Institutet, Stockholm, Sweden

Christian Madsen, Per Nafstad, Per Schwarze, Wenche Nystad
Norwegian Institute of Public Health, Oslo, Norway

Tarja Yli-Tuomi, Timo Lanki
Department of Environmental Health, National Institute for Health and Welfare, Kuopio, Finland

Jaakko Kukkonen
Finnish Meteorological Institute, Helsinki, Finland

Kirsten Eriksen, Mette Sørensen, Ole Raaschou-Nielsen, Zorana Andersen
Danish Cancer Society Research Center, Copenhagen, Denmark

Thomas Becker
Department of Environmental Science, Aarhus University, Denmark

Audrius Dedele, Regina Grazuleviciene
Vytautas Magnus University, Kaunas, Lithuania

Timothy Key
Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom

Gabriele Nagel, Gioia Mosler, Kees de Hoogh, Paolo Vineis, Wei Xun
Imperial College London, London, United Kingdom

Anna Mölter, Sarah Lindley
The University of Manchester, Manchester, England, United Kingdom

Ulrich Quass
Air Quality & Sustainable Nanotechnology, IUTA Institüt für Energie- und Umwelttechnik e.V., Duisburg, Germany

Annette Peters, Joachim Heinrich, Josef Cyrys, Kathrin Wolf, Matthias Birk
Institute of Epidemiology I and II, Helmholtz Research Center Munich, Neuherberg, Germany

Barbara Hoffmann, Dorothea Sugiri, Gudrun Weinmayr
IUF Leibniz Research Institute for Environmental Medicine, University of Düsseldorf, Düsseldorf, Germany

Christophe Declercq, Morgane Stempfelet
French Institute for Public Health Surveillance, Saint-Maurice, France

Evi Dons
Flemish Institute for Technological Research, Environmental Risk and Health unit, Mol, Belgium

Alex Ineichen, Danielle Vienneau, Harisch Phuleria, Martin Adam, Martina Ragettli, Ming-Yi Tsai, Nicole Probst-Hensch, Nino Künzli
Department of Epidemiology and Public Health, Swiss Tropical & Public Health Institute, Basel, Switzerland; University of Basel, Basel, Switzerland

Gabriele Nagel, Hans Concin
Agency for Preventive and Social Medicine, Bregenz, Austria

Mihaly J. Varró, Orsolya Udvardy, Gizella Nádor
National Institute of Environmental Health, Budapest, Hungary

Audrey de Nazelle, Mark Nieuwenhuijsen, Marta Cirach, Xavier Basagana
Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; CIBER Epidemiología y Salud Pública (CIBERESP), Spain; IMIM (Hospital del Mar Research Institute), Barcelona, Spain

Andrea Ranzi, Claudia Galassi
AOU San Giovanni Battista - CPO Piedmont, Turin, Italy

Chiara Badaloni, Francesco Forastiere, Giulia Cesaroni, Massimo Stafoggia
Epidemiology Department, Lazio Regional Health Service, Rome, Italy

Enrica Migliore
Department of Public Health and Community Medicine, University of Verona, Italy

George Grivas
School of Chemical Engineering, National Technical University of Athens, Athens, Greece

Alexandros Gryparis, Evangelia Samoli, Klea Katsouyanni, Konstantina Dimakopoulou, Xanthi Pedeli
Department of Hygiene, Epidemiology & Medical Statistics, National and Kapodistrian University of Athens, Medical School, Athens, Greece

Euripides Stephanou, Minas Iakovides
Environmental Chemical Processes Laboratory, University of Crete, Heraklion, Greece

Michael Jerrett
School of Public Health, University of California, Berkeley, California, United States

Acknowledgements

Finally, it is time to reap the harvest of my four years' PhD study. Completing a PhD degree is indeed a challenging and passionate progress which cultivates me to be an independent scientist. This would not have been achieved without the aids and supports from many people to whom I would like to express my sincere gratitude.

My first and special gratitude is given to my promotor Prof. Dr. Bert Brunekreef for your excellent supervision, patience, understanding and systematic training during my graduate period at Utrecht University. Your charisma is what first attracted me to the Netherlands and to this academic journey. Thanks for spending time regularly on discussing articles with me at the beginning of my PhD, no matter how tight your schedule was. It was your effort that enabled me to bridging the gap of my backgrounds smoothly. Moreover, it was you who led me to this opportunity to be involved in an excellent European study, where I received comprehensive training in both exposure and epidemiology fields, and where I learnt how international cooperation works. Besides the academic instructions, it was also you who supported me in the scholarship application and especially in settling down in the Netherlands, and showing how wonderful this country is. During the four years, you always answered all my questions patiently, no matter if they were scientific or personal. It is a great honor to receive your supervision.

A special and sincere thanks also goes to my co-promoters Dr. Gerard Hoek and Dr. Rob Beelen, who made sure my graduation process went smoothly and for providing great assistances in writing my thesis. Gerard, thank you for always being there, showing me how to solve scientific problems, and how to working in an international team. My work always benefited from your inspiring thoughts and valuable suggestions. Rob, thank you for your step-by-step instructions, in guiding me in developing land use regression models and conducting statistical analyses. I appreciate your advice on my research and your patience for my endless questions. Your encouragement and suggestions gave me confidence whenever I felt doubt on my progress.

Many thanks to the Chinese Scholarship Council for funding my PhD research. And I would like to take this opportunity to thank my master supervisor Prof. Dr. Tong Zhu at Peking University for leading me to Prof. Dr. Bert Brunekreef and for your valuable advices on my career. Thank you Prof. Dr. Chang-Chuan Chan at National Taiwan University, Prof. Dr. Haidong Kan at Fudan University and Prof. Dr. Ta-yuan Chang at China Medical University for sharing your academic career experiences and for your valuable suggestions on the project and ISEE conferences. I very much enjoyed our discussions.

I would like to particularly acknowledge Kees Meliefste, who guided my lab and field works in my first year; Dr. Roel Vermeulen, who trained me in molecular epidemiology; Ingrid Dahmen, who aided my stay in IRAS and in the Netherlands; Dr. Kees de Hoogh at imperial college London, who collaborated and contributed to chapter 4 of my thesis, and the ESCAPE members, who shared their precious comments to improve my thesis.

I would like to thank my dissertation committee members Prof. Dr. Hans Kromhout, Prof. Dr. Michiel Bots, Prof. Dr. Martin Dijst, Prof. Dr. Michael Brauer at British Columbia University and Dr. Menno Keuken at TNO for your time, interests

and valuable comments.

Many thanks to my IRAS colleagues and friends who collaborated and helped me directly and indirectly. Thanks to Denise Montagne and Aileen Yang for being the paranymphs of my defense. Aileen, thank you for the language assistances and for your hospitality and hosting all those wonderful parties. Denise, thank you for showing me almost everything at IRAS. Marloes and Ming, I am indebted to you for everything you shared with me in terms of, experiences, knowledge and nice food. Thanks also to my former and current officemates Maartje, Monica, Yi, Saskia, Elaine, Natasha, Victoria, Sadegh, Kristel, Tom, Jose, Hanna, Johan, Anke, Virissa, Marisa, Josue and Katrin, for all the laughs, fun, advices and the happy times we shared.

Besides my academic career, I feel lucky to have met many old and new friends here in the Netherlands. I have to mention the “Waranda team” (Chenjuan Jiang, Meng Chen, Yinghuan Kuang, Wufeng Chen, Cong Zhang, Fang Chen, Yafei Liu, Shaoyu Yin, Jingwei Zhang, Junting Zhang etc), who encouraged, helped each other to ease the transition of moving to a new country. Thanks to Guotai, a good friend with a warm heart, for your willingness to share your experiences and progresses with me, and for your support during difficult times. Time flies, now we are departing to different places of the world to pursue our dreams. I will cherish the memories we made in the Netherlands forever.

I would like to thank Junting Zhang, uncle Wu, aunt Li, Hong and Darwin. You made Netherlands the home for me.

Finally and most importantly, I would like to express my deepest and most sincere love to my parents, who are always there for me with your love, support and encouragements!

Curriculum Vitae

Meng Wang was born on 31th July, 1983 in Kunming, China. After graduating his high school (college attached to Yunnan Normal University in Kunming, China) in 2002, he started his undergraduate study in pharmaceutical engineering in Tianjin University (Tianjin, China). He was funded by undergraduate association to start medical solvent extraction research and further on participating in the development of new antirheumatic drug in Tianjin Pharmaceutical Institute. In 2006, he started the MSc programme “Temporal and spatial variation of air pollution in Beijing using mobile *in-situ* laboratory” in Peking University (Beijing, China) under the supervision of Prof. Tong Zhu and obtained his Master's degree in 2009. He participated two international corporative projects: the CAREBeijing-2008 (Campaign of Air Quality Research of Beijing-2008) and the PRIDE-PRD-2008 (Program of Regional Integrated Experiments on Air Quality over the Pearl River Delta-2008) project. Since 2009, he started working as a PhD student in environmental epidemiology at the Institute for Risk Assessment Sciences (IRAS) of Utrecht University (The Netherlands) in Prof. Bert Brunekreef's group. His PhD research is within the framework of the ESCAPE (The European Study of Cohorts for Air Pollution Effects) and the TRANSPHORM (Transport related Air Pollution and Health impacts – Integrated Methodologies for Assessing Particulate Matter) projects which is mainly focusing on the associations between air pollution and health effects in multi-cities of Europe. The results of his PhD study are described in this thesis

List of publications

- **Wang, M**.; Beelen, R.; Stafoggia, M.; Samoli, E.; Andsen, Z.; Nieuwenhuijsen, M.; Fischer, P.; Hoffmann, B.; Raacshou-Nielsen, O.; Vineis, P.; Xun, W.; Dimakopoulou, K.; Sorensen, M.; Krogh, V.; Bueno-de-Mesquita, B.; Amiano, P.; Peeters, P.; Ostensson, C.; Key, T.; Peters, A.; Migliore, E.; Forsberg, B.; Lanki, T.; Pedersen, N.; De Faire, U.; Heinrich, J.; Concin, H.; Badaloni, C.; Fratiglioni, L.; Nagel, G.; Katsouysanni, K.; Hoek, G.; Brunekreef, B.; Long-term exposure to elemental constituents of particulate matter and cardiovascular mortality in 19 European cohorts: results from the ESCAPE and TRANSPHORM projects. Submitted.
- **Wang, M**.; Beelen, R.; Bellander, T.; Birk, M.; cesaroni, G.; Cirach, M.; Cyrys, J.; de Hoogh, K.; Declercq, C.; Dimakopoulou, K.; Eeftens, M.; Eriksen, K.T.; Forastiere, F.; Galassi, C.; Grivas, G.; Heinrich, J.; Hoffmann, B.; Ineichen, A.; Korek, M.; Lanki, T.; Lindley, S.; Modig, L.; Molter, A.; Nafstad, P.; Nieuwenhuijsen, M.; Nystad, W.; Olsson, D.; Rasschou-Nielsen, O.; Ragettli, M.; Ranzi, A.; Stempfelet, M.; Sugiri, D.; Tsai, M.Y.; Udvardy, O.; Varro, M.J.; Vienneau, D.; Weinmayr, G.; Wolf, K.; Tli-Tuomi, T.; Hoek, G.; Brunekreef, B.; Performance of multi-cities land use regression models for NO_2 and fine particles. Environ Health Perspect, under revision.
- Beelen, R.; Stafoggia, M.; Raaschou-Nielsen, O.; Andersen, Z.J.; Xun, W.W; Katsouyanni, K.; Dimakopoulou, K.; Brunekreef, B.; Weinmayr, G.; Hoffmann, B.; Wolf, K.; Samoli, E.; Houthuijs, D.; Nieuwenhuijsen, M.; Oudin, A.; Forsberg, B.; Olsson, D.; Salomaa, V.; Lanki, T.; Yli-Tuomi, T.; Oftedal, B.; Aamodt, G.; Nafstad, P.; De Faire, U.; Pedersen, N.L.; Östenson, C.; Fratiglioni, L.; Penell, J.; Korek, M.; Pyko, A.; Eriksen, K.T.; Tjønneland, A.; Becker, T.; Eeftens, M.; Bots, M.; Meliefste, M.; **Wang, M.**; Mesquita, B.B.; Sugiri, D.; Krämer, U.; Heinrich, J.; de Hoogh, K.; Key, T.; Peters, A.; Cyrys, J.; Concin, H.; Nagel, G.; Ineichen, A.; Schaffner, E.; Probst-Hensch, N.; Dratva, J.; Ducret-Stich, R.; Vilier, A.; Clavel-Chapelon, F.; Stempfelet, M.; Grioni, S.; Krogh, V.; Tsai, M.Y.; Marcon, A.; Ricceri, F.; Sacerdote, C.; Galassi, C.; Migliore, E.; Ranzi, A.; Cesaroni, G.; Badaloni, C.; Forastiere, F.; Tamayo, I.; Amiano, P.; Dorronsoro, M.; Katsoulis, M.; Trichopoulou, A.; Vineis, P.; Hoek, G.; Long-term exposure to air pollution and cardiovascular mortality: an analysis of 22 European cohorts within the ESCAPE project. Epidemiology, accepted.
- Beelen, R.; Raaschou-Nielsen, O.; Stafoggia, M.; Andersen, Z.J.; Weinmayr, G.; Hoffmann, B.; Wolf, K.; Samoli, E.; Fischer, P.; Nieuwenhuijsen, M.; Vineis, P.; Xun, W.W.; Katsouyanni, K.; Dimakopoulou, K.; Oudin, A.; Forsberg, B.; Modig, L.; Havulinna, A.S.; Lanki, T.; Turunen, A.; Oftedal, B.; Nystad, W.; Nafstad, P.; De Faire, U.; Pedersen, N.L.; Östenson, C.G.; Fratiglioni, L.; Penell, J.; Korek, M.; Pershagen, G.; Eriksen, K.T.; Overvad, K.; Ellermann, T.; Eeftens, M.; Peeters, P.H.; Meliefste, K.; **Wang, M**.; Bueno-de-Mesquita, B.; Sugiri, D.; Krämer, U.; Heinrich, J.; de Hoogh, K.; Key, T.; Peters, A.; Hampel, R.; Concin, H.; Nagel, G.; Ineichen, A.; Schaffner, E.; Probst-Hensch, N.; Künzli, N .; Schindler, C.; Schikowski, T.; Adam, M.; Phuleria, H.; Vilier, A.; Clavel-Chapelon, F.; Declercq, C.; Grioni, S.; Krogh, V.; Tsai, M.Y.; Ricceri, F.; Sacerdote, C.; Galassi, C.; Migliore, E.; Ranzi, A.; Cesaroni, G.; Badaloni, C.; Forastiere, F.; Tamayo, I.; Amiano, P.; Dorronsoro, M.; Katsoulis, M.; Trichopoulou, A.; Brunekreef, B.; Hoek, G.; Effects of long-term exposure to air pollution on natural cause mortality: an analysis of 22 European cohorts within the multi-center ESCAPE project. The Lancet, accepted.
- Raaschou-Nielsen, O.; Andersen, Z. J.; Beelen, R.; Samoli, E.; Stafoggia, M.; Weinmayr, G.; Hoffmann, B.; Fischer, P.; Nieuwenhuijsen, M. J.; Brunekreef, B.; Xun, W. W.; Katsouyanni, K.; Dimakopoulou, K.; Sommar, J.; Forsberg, B.; Modig, L.; Oudin, A.; Oftedal, B.; Schwarze, P. E.; Nafstad, P.; Faire, U. D.; Pedersen, N. L.; Östenson, C.; Fratiglioni, L.; Penell, J.; Korek, M.; Pershagen, G.; Eriksen, K.; Sørensen, M.; Tjønneland, A.; Ellermann, T.; Eeftens, M.; Peeters, P. H.; Meliefste, K.; **Wang, M**.; Bueno-de-Mesquita, B.; Key, T. J.; Hoogh, K. d.; Concin, H.; Nagel, G.; Vilier, A.; Grioni, S.; Krogh, V.; Tsai, M.; Ricceri, F.; Sacerdote, C.; Galassi, C.; Migliore, E.; Ranzi, A.; Cesaroni, G.; Radaloni, C.; Forastiere, F.; Tamayo, I.; Amiano, P.; Dorronsoro, M.; Trichopoulou, A.; Bamia, C.; Vineis, P.; Hoek, G.; A ir pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). The Lancet-oncology, 2013, 14 (9), 813-822.
- de Hoogh, K.; **Wang, M.**; Adam, M.; Badaloni, C.; Beelen, R.; Birk, M.; Cesaroni, G.; Cirach, M.; Declercq, C.; Dedele, A.; Dons, E.; de Nazelle, A.; Eeftens, M.; Eriksen, K.; Fischer, P.;

Grazuleviene, R.; Gryparis, A.; Hoffmann, B.; Jerrett, M,; Katsouyanni, K.; Iakovides, M.; Lanki, T.; Lindley, S.; Madsen, C.; Molter, A.; Mosler, G.; Nador, G.; Nieuwenhuijsen, M.; Pershagen, G.; Peters, A.; Phuleria, H.; Probst-Hensch, N.; Raaschou-Nielsen, O.; Quass, U.; Ranzi, A.; Stephanou, E.; Sugiri, D.; Schwarze, P.; Tsai, M.; Yli-Tuomi, T.; Varro, M.; Vienneau, D.; Weinmaryr, G.; Brunekreef, B.; Hoek, G.; Development of land use regression models for particle composition in 20 study areas in Europe. Environ. Sci. Technol. 2013, 47 (11), 5778–5786.

- **Wang, M.**; Beelen, R.; Basagana, X.; Becker, T.; Cesaroni, G; de Hoogh, K.; Dedele, A.; Declercq, C.; Dimakopoulou, K.; Eeftens, M.; Forastiere, F.; Galassi, C.; Grazuleviciene, R.; Hoffmann, B.; Heinrich, J.; Iakovides, M.; Kunzli, N.; Korek, M.; Lindley, S.; Molter, A.; Mosler, G.; Madsen, C.; Nieuwenhuijsen, M.; Phuleria, H.; Pedeli, X.; Raaschou-Nielsen,O.; Ranzi, A.; Sugiri, D.; Stempfelet, M.; Tsai, M.; Lanki, T.; Udvardy, O.; Varro, M.J.; Wolf, K.; Weinmayr, G.; Yli-Tuomi, T.; Hoek, G.; Brunekreef, B. Evaluation of land use regression models for NO_2 and particulate matter in 20 European study areas: the ESCAPE project. Environ. Sci. Technol. 2013, 47 (9), 4357–4364.
- Boogaard, H.; Fisher, P.H.; Janssen, N.A.H.; Kos, G.P.A.; Weijers, E.P.; Cassee, F.R.; van der Zee, S.C.; de Hartog, J.J.; Meliefste, K.; **Wang, M.**; Brunekreef, B.; Hoek, G.; Respiratory effects of a reduction in outdoor air pollution concentration. Epidemiology, 2013, 24 (5), 753-761.
- Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; Tsai, M.; Künzli, N.; Schikowski, T.; Marcon, A.; Eriksen, K.; Raaschou Nielsen, O.; Stephanou, E.G.; Patelarou, E.; Lanki, T.; Yli-Tuomi, T.; Declercq, C.; Falq, G.; Stempfelet, M.; Birk, M.; Cyrys, J.; von Klot, S.; Nádor, G.; János Varró, M.; Dedele, A.; Grazuleviciene, R.; Mölter, A.; Lindley, S.; Madsen, C.; Cesaroni, G.; Ranzi, A.; Badaloni, C.; Hoffmann, B.; Nonnemacher, M.; Krämer, U.; Kuhlbusch, T.; Cirach, M.; de Nazelle, A.; Nieuwenhuijsen, M.J.; Bellander, T.; Korek, M.; Olsson, D.; Strömgren, M.; Dons, E.; Jerrett, M.; Fischer, P.; **Wang, M.**; Brunekreef, B.; de Hoogh, K.; Development of NO_2 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, doi 10.1016/j.atmosenv.2013.02. 037.
- Eeftens, M.; Beekhuizen, J.; Beelen, R.; **Wang, M.**; Vermeulen, R.; Brunekreef, B.; Huss, A.; Hoek, G.; Quantifying urban street configuration for improvements in air pollution models. Atmos Environ 2013, 72, 1-9, 10.1016/j.atmosenv.2013.02.007.
- **Wang, M.**; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B.; A systematic validation of land use regression models for NO_2. Environ. Sci. Technol. 2012, 46, 4481−4489.
- Cyrys, J.; Eeftens, M.; Heinrich, J.; Ampe, C.; Armengaud, A.; Beelen, R.; Bellander, T.; Beregszaszi, T.; Birk, M.; Cesaroni, G.; Cirach, M.; de Hoogh, K.; De Nazelle, A.; de Vocht, F.; Declercq, C.; Dedele, A.; Dimakopoulou, K.; Eriksen, K.; Galassi, C.; Grazuleviciene, R.; Grivas, G.; Gruzieva, O.; Hagenbjörk Gustafsson, A.; Hoffmann, B.; Iakovides, M.; Ineichen, A.; Krämer, U.; Lanki, T.; Lozano, P.; Madsen, C.; Meliefste, K.; Modig, L.; Mölter, A.; Mosler, G.; Nieuwenhuijsen, M. J.; Nonnemacher, M.; Oldenwening, M.; Peters, A.; Pontet, S.; Probst Hensch, N.; Quass, U.; Raaschou Nielsen, O.; Ranzi, A.; Sugiri, D.; Stephanou, E. G.; Taimisto, P.; Tsai, M.; Vaskövi, E.; Villani, S.; **Wang, M.**; Brunekreef, B.; Hoek, G.; Variation of NO_2 and NOx concentrations between and within 36 European study areas: results from the ESCAPE study. Atmos Environ 2012, 62, 374 - 390.
- Boogaard, H.; Janssen, N.A.; Fischer, P.H.; Kos, G.P.; Weijers, E.P.; Cassee, F.R.; van der Zee, S.C.; de Hartog, J.J.; Meliefste, K.; **Wang, M.**; Brunekreef, B.; Hoek, G.; Impact of low emission zones and local traffic policies on ambient air pollution concentrations. Science of the Total Environment, 2012, 435, 132-140.
- **Wang, M.**; Zhu, T., Zhang, J.P.; Zhang, Q.H.; Lin, W.W.; Li, Y.; Wang, Z.F.; Using a mobile laboratory to characterize the distribution and transport of sulfur dioxide in and around Beijing, Atmos. Chem. Phys., 2011, 11, 11631–11645, doi:10.5194/ acp-11-11631-2011.
- **Wang, M.**; Zhu, T.; Zheng, J.; Zhang, R.Y.; Zhang, S.Q.; Xie, X.X.; Han, Y.Q.; Li, Y.; Use of a mobile laboratory to evaluate changes in on-road air pollutants during the Beijing 2008 Summer Olympics, Atmos. Chem. Phys., 2009, 9, 8247–8263, doi:10.5194/acp-9-8247- 2009.