ADDITIONAL MATERIALS AVAILABLE ON THE HEI WEBSITE

Research Report 208

Mortality and Morbidity Effects of Long-Term Exposure to Low-Level $PM_{2.5}$, BC, NO_2, and O_3: An Analysis of European Cohorts in the ELAPSE Project

Brunekreef et al.

Additional Materials 1. Appendix 1: ELAPSE Methods

These Additional Materials were not formatted or edited by HEI. This document was part of the HEI Low-Exposure Epidemiology Studies Review Panel's review process.

Correspondence may be addressed to Dr. Brunekreef, Utrecht University, PO Box 80176 3508 TD Utrecht, The Netherlands; e-mail: *b.brunekreef@uu.nl*.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83467701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

Appendix 1: ELAPSE Methods

TABLE OF CONTENTS

Contents

Appendix M1: Details on Administrative cohorts

The English administrative cohort, Dutch national cohort, Rome longitudinal cohort, Swiss national, Norwegian national, Danish national and Belgian national cohort have been analyzed in addition to the ESCAPE cohorts with full individual-level data. This adds over 35 million subjects to the analysis of the well-characterized ESCAPE cohorts. These administrative cohorts were selected because they were recruited relatively recently and major efforts in setting up the cohorts including compliance with national privacy regulations have already been dealt with. Because of these privacy regulations and the size of the cohorts, they will be analyzed individually. In analyses published so far, these cohorts used different approaches for exposure assessment, data analysis and indirect confounder control. The current study will analyze all administrative cohorts using the same exposure assessment method and analytical methods. Using the exposure assessment methodology of the current study, we will expand the pollutants studied (*e.g.*, add PM2.5, BC and O_3 for the Dutch cohort) and refine the spatial scale (from 1*1 km or 10*10 km in some of the rural areas to essentially address level). Follow-up time will be increased, increasing power even further. Finally, for the three cohorts in Norway, Switzerland and Denmark, air pollution exposure assignment and epidemiological analysis will be conducted for the first time.

The **Dutch Environmental Longitudinal Study (DUELS)** is described in detail by (Fischer et al. 2015). Briefly, population statistics based on digital municipal registers are combined by Statistics Netherlands into a longitudinal file for each individual registered in the municipal registration. Changes in demographic attributes (e.g. death, address, marital status, emigration, region of origin) are updated yearly. In these files, the individual identification number is replaced by an encrypted unique identification number. This identification number is used to enrich the individual files with information from other central data sources available at Statistics Netherlands, including data from tax records. In the paper by Fischer et al. (2015), we selected all Dutch inhabitants of ≥30 years of age on 1/1/2004, living at the same residential address since 1/1/1999. For the ELAPSE study, we shifted the baseline to 1/1/2008 and follow-up to 1/1/2013 (*i.e.*, 5 years).

The **Belgian 2001 Census cohort** is based on the entire Belgian population officially residing in Belgium in 2001. Emigration and mortality follow-up data is available for the period from 01/10/2001-31/12/2011 (10.24 years). Data were made available by the Belgian statistical office (Statbel). Additionally, geocoding of the residential addresses at baseline, and linkage of air pollution exposure data was performed by Statbel. Geographical coverage was almost complete with 98.7% of individuals included. Available individual covariates at baseline are: age (birth date), sex, marital status, country of origin, education level, occupational status, and residential history. Available area-level SES variables consisted of mean income, unemployment rate, low education level, and ethnicity.

The **Swiss National Cohort (SNC)** is a national longitudinal research platform linking census data with birth, mortality, and emigration data. The SNC was approved by the Ethics Committees of the Cantons of Zurich and Bern. Due to mandatory participation, nearly all persons residing in Switzerland at the time of the 1990 and 2000 censuses are represented; an estimated 98.6% residents participated in 2000. For each person, the SNC contains an individual (*e.g.*, sex, date of birth, occupation), household (e.g. type of household, socio-economic position (SEP)), and building (*e.g.*, type of building, number of floors, geographical coordinate) record. Prior to 2010 the SNC was based on a probabilistic linkage. In 2010, Switzerland replaced the classic door-to-door census system with the registry-based census repeated each year. As such, a deterministic linkage with a unique pseudo ID (SNC-ID), based on the social security number but cannot be traced back to it, is now used. In this new framework, data on education, occupation, employment or religion is only collected in an annual structural enquiry of a random sample of about 250,000 people per year. Swiss TPH received the latest SNC data (for 1990-2014) in October 2017 with all the necessary permissions to conduct analyses.

The **Rome cohort** includes 30+ year old subjects who filled in the Census questionnaire and were resident in Rome at 20 October 2001 (census reference day) were identified and followed up using the Health Information System (Cesaroni et al. 2013). The cohort is part of the National Statistical Program for the years 2011-2017 and was approved by the Italian Data Protection Authority.

The **English CPRD cohort** is constructed from data provided by the Clinical Practice Research Datalink (https://www.cprd.com). CPRD is a large, validated, and nationally representative database containing anonymized patient data from UK primary care. It includes a full longitudinal medical record for each patient consulting their family practitioner including information on diagnoses made within the practice. Where patient consent is given, the data are linked to death registrations and hospital admissions. Access to data in CPRD for the ELAPSE study requires three separate, consecutive approvals covering the legal basis for the linkage of primary care records to air pollution data; the scientific basis for the project including feasibility, quality and public health value; and the integrity, security and management (including linkage) of the data. The first application, to the Confidential Advisory Group, is the responsibility of CPRD, as is the application to NHS Digital (formerly HSCIC) for the actual linkage. The application for scientific approval (ISAC) is made by St George's, University of London.

NORCOHORT is the administrative national cohort for Norway**.** The size and age range of the cohort is 2.7 million adults, 30 years or older at baseline. The follow-up period is 2001 to 2016. The cohort links data from various registries, all requiring separate approvals, including cancer registry, mortality registry, CVDNOR health register, Norwegian Patient registry and from Statistics Norway Address history, several SES variables on individual and area level (county, municipality, neighborhood, small neighborhood), noise data.

The **Danish National Cohort** includes 5.3 M subjects of all ages as of the year 2000, with follow-up until 2015. This data is stored in Statistics Denmark's Computer for linking to geocoded address data on individual level for the entire population. We have obtained permission from all relevant Danish authorities to utilize information on the entire Danish population to study adverse effect of air pollution parameters on various health outcomes. These data obtained include personal level information on sex, birthdate, vital status, obstetric factors, somatic health outcomes, non-accidental and cause-specific mortality, social and socioeconomic information including education, income, and labor market affiliation.

Appendix M2: Local exposure models

In this section the local models listed in main report Table 2 are being described in more detail. ESCAPE models are land use regression models developed for individual study areas, hence based upon a limited number of monitoring locations, but using local predictor data and based on actual points, not 100x100 m grids. Details are in (Beelen et al. 2013; Eeftens et al. 2012). Local models were available from other sources and use a diversity of models. Some models have previously been especially applied in the administrative cohorts. Methods for the Duels, the Rome and UK administrative cohort can be found in (Carey et al. 2013; Cesaroni et al. 2013; Fischer et al. 2015). Below more detailed discussion is provided for the administrative cohorts and briefly the pooled cohort. Methods differ per cohort, which is a larger concern for the pooled cohort.

In **Switzerland** the dispersion model Pollumap is used. Pollumap is a Gaussian plume atmospheric dispersion model estimating NO_2, PM2.5 and O_3 concentrations (2010) at a 200x200 m resolution across Switzerland (FOEN. NO2 Ambient Concentrations in Switzerland. Modelling Results for 2005, 2010, 2015, Environmental Studies No. 1123; Federal Office for the Environment (FOEN): Berne, Switzerland, 2011; pp. 1–68).

Belgian Local exposure assessment was performed by the Belgian Interregional Environment Agency (IRCEL-CELINE). Annual mean concentrations of ambient air pollution for $PM_{2.5}$, NO_2, O_3 warm season and BC in micrograms per cubic meter (µg/m^3) were assigned at the level of the georeferenced residential address of the participant at baseline (01/10/2001). The estimates for the different pollutants were obtained through the interpolation of air quality measurements derived from fixed measuring stations and the use of a layer of land cover data (CORINE Land Cover) (Hooyberghs et al. 2006). These metrics were combined with a dispersion model using emissions from point and line sources and meteorological data (Lefebrve and Vrankxs 2013). The data was modelled on high-resolution grids of 25 m x 25 m. Further details regarding the applied model chain can be consulted in the following technical report by Lefebvre and Vranckx (Lefebrve and Vrankxs 2013). The estimates for $PM_{2.5}$, NO_2, O_3 warm season and BC were equivalently available for the year 2010.

In **the Netherlands**, the local model consists of two components. First, large scale concentrations maps of among others PM2.5, NO_2 and EC at a spatial resolution of 1x1 km are made, using the Operational Priority Substances (OPS) dispersion model and measurements from the National Air Monitoring network. Next, the local traffic contribution is added to these background grids at a much higher resolution using dedicated models (Velders 2016).

For **Norway**, the local model uses data from the project "Nasjonalt beregningsverktøy for luftkvalitet" (National calculation tools for air quality; http://www.luftkvalitet-nbv.no). Main executers were the Norwegian Institute for Air Research (NILU) (www.nilu.no) and the Norwegian Meteorological Institute (www.met.no) with contributions from NIPH among others. NILU provided air pollutant concentration maps for PM10, PM2.5, NO_2, but only for

one specific year (2015). The population included for the NILU exposures are the 15 largest cities in Norway, which cover a large part of the population. For these we get a resolution below 1x1 km, for the rest of the country 1x1 km. The emission inventory has been updated recently as a basis for the modelling.

For **Rome**, a PM2.5 a dispersion model and a LUR model for NO_2 based on a large survey in 2007 is used (Cesaroni et al. 2012). The surface predicted by the 2007 LUR model agreed closely with the surface from a LUR based on a 1996 survey in Rome (Cesaroni et al. 2012).

For **Denmark**, Urban background air pollution concentrations (UBAP) have been generated for the entire Danish area on 1km x 1km using the chemistry-transport modelling system AirGIS (http://envs.au.dk/en/knowledge/air/models/airgis) (Hertel et al. 2001; Ketzel et al. 2011; Ketzel M. et al. 2003). This Urban Background air pollution data has been transferred to Statistics Denmark's computer and linked to address geocoded data on individual level for the entire population. For 40% of the addresses in Denmark, there is significant traffic at the address (>500 vehicles per day at a road in the immediate vicinity of the address) to provide a local pollution contribution. For these addresses, the street pollution model OSPM was run (Berkowicz et al. 2008). This data are merged with the urban background air pollution data to form a new combined exposure database. This data was transferred to Statistics Denmark's Computer for linking to geocoded address data on individual level for the entire population.

For the **UK cohort** the dispersion model methods described in (Carey et al. 2013) have been used.

For the **pooled cohort,** local models were supplied for the CEANS cohort, based upon dispersion models for Stockholm, which have been extensively used in epidemiological studies on air pollution by Karolinska Institute. For the Danish DCH cohort, NO_2 from the Airgis model used for the Danish national cohort is available. For the Danish DNC, data on PM2.5, O_3 and NO_2 were available based on the Airgis model. For the German HNR cohort data on all pollutants were available from local dispersion models applied extensively in work the HNR cohort. For the German KORA cohort, a refined LUR model was available.

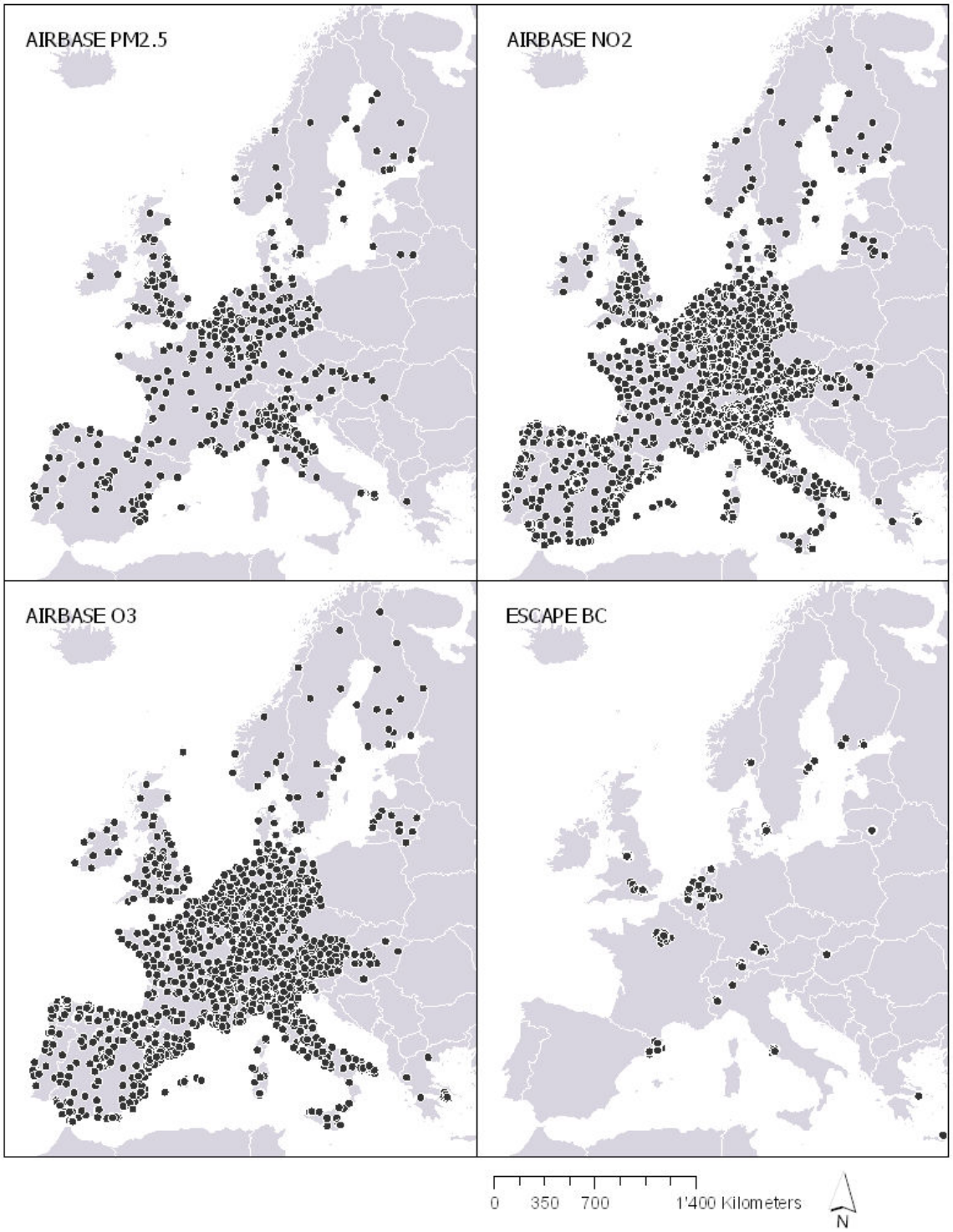


Figure M1 Location of PM2.5, NO_2, O_3 and BC monitoring sites for the year 2010. (DeHoogh et al. 2018; reproduced with permission from Elsevier)

Table M1. Model structure[1] and performance of 2010 LUR models (adapted from (de Hoogh et al. 2018))

Pollutant	Stage	Method	N sites	R^2	RMSE[2]	Full LUR model[3]
$PM_{2.5}$	Training	LUR	543	62.2	3.17	3.19 +13.24*SAT-PM25 +7.08*MACC-PM25 – 3.82* ALT + 2.17*$ALRD_{100}$ - 2.07*NAT_{50} +2.39*POR_{800} +1.41*RES_{200}
		LUR +Kriging		72.2	2.71	
	HOV	LUR		58.7	3.30	
		LUR +Kriging		66.4	2.97	
BC[4]	Training	LUR	436	54.4	0.56	0.99 + 0.85* MACC-PM25 + 0.30* SAT-PM25 + 0.68*$MJRD_{100}$ +0.40* ALRD + 0.45* $ALRD_{700}$ +0.90*RES_{300} -0.12*UGR – 1.16*Y
	HOV	LUR		51.4	0.58	
NO_2[4]	Training	LUR	2399	58.8	9.38	3.30 + 22.73*MACC-NO_2 + 7.04* $ALRD_{50}$ + 3.92* $ALRD_{300}$ + 12.32* $MJRD_{100}$ + 15.73*$ALRD_{2000}$ -3.38*NAT_{400} + 4.1*POR_{700} + 5.8*RES_{300}
	HOV	LUR		57.5	9.51	
O_3 warm	Training	LUR	1730	45.5	10.07	30.00 +32.57*DEHM-O_3 -6.87* $ALRD_{200}$ –6.03* $MJRD_{100}$ -5.95*$PORT_{5000}$ - 4.79*RES_{2000} +5.70*ALT
		LUR +Kriging		69.6	7.51	
	HOV	LUR		44.5	10.15	
		LUR +Kriging		59.9	8.63	

1 Regression slope in µg/m^3, except BC (10-5m-1), multiplied by the difference between the 1th and 99th percentile of each predictor to allow comparison across predictors;

2 RMSE in µg/m^3, except BC ($10^{-5}m^{-1}$)

3 ALT = altitude, ALRD = all roads, MJRD = major roads, IND = industry, POR = ports, UGR = urban green, TBU = total build up, NAT = natural land, RES = residential, POP = sum of population, X = North-South trend, Y = East–West trend, SAT = satellite, MACC = MACC dispersion model, DEHM = DEHM CTM. Number in subscript depicts the buffer size (e.g. ALRD100 = sum of all road length within 100m)

4 No valid variograms were possible on the residuals of these models

Table M2 Validation of the ELAPSE exposure model at low pollution levels. HOV refers to the five-fold hold-out validation approach

	Validation at ELAPSE sites HOV)				**Validation at ESCAPE sites**		
PM2.5							
	R^2	RMSE	n		R^2	RMSE	n
All	0,664	2,97	543		0,648	3,41	416
<25 µg/m^3	0,649	2,70	523		0,601	2,95	390
<20 µg/m^3	0,630	2,34	439		0,612	2,37	327
<15 µg/m^3	0,532	1,92	230		0,593	1,58	191
<12 µg/m^3	0,379	1,55	130		0,405	1,32	118
<10 µg/m^3	0,379	1,26	86		0,271	1,05	74
NO_2							
	R^2	RMSE	n		R^2	RMSE	n
All	0,575	9,51	2399		0,494	11,47	1396
<40 µg/m^3	0,613	5,88	2008		0,507	6,14	1116
<30 µg/m^3	0,620	4,48	1520		0,491	4,74	844
<20 µg/m^3	0,572	3,20	841		0,384	3,31	433
BC							
					R^2	RMSE	n
All	NA	NA	NA		0,514	0,58	433
<3*10-5/m					0,424	0,45	390
<2.5*10-5/m					0,375	0,39	351
<2*10-5/m					0,311	0,33	277
<1.5*10-5/m					0,343	0,21	162
<1*10-5/m					0,119	0,14	59
<0.5*10-5/m					0,717	0,07	3
O_3 warm season							
	R^2	RMSE	n				
All	0,599	8,63	1728		NA	NA	NA
< 120 µg/m^3	0,595	8,49	1716				
< 100 µg/m^3	0,448	7,77	1366				
< 80 µg/m^3	0,073	7,78	393				
< 60 µg/m^3	0,068	6,32	40				

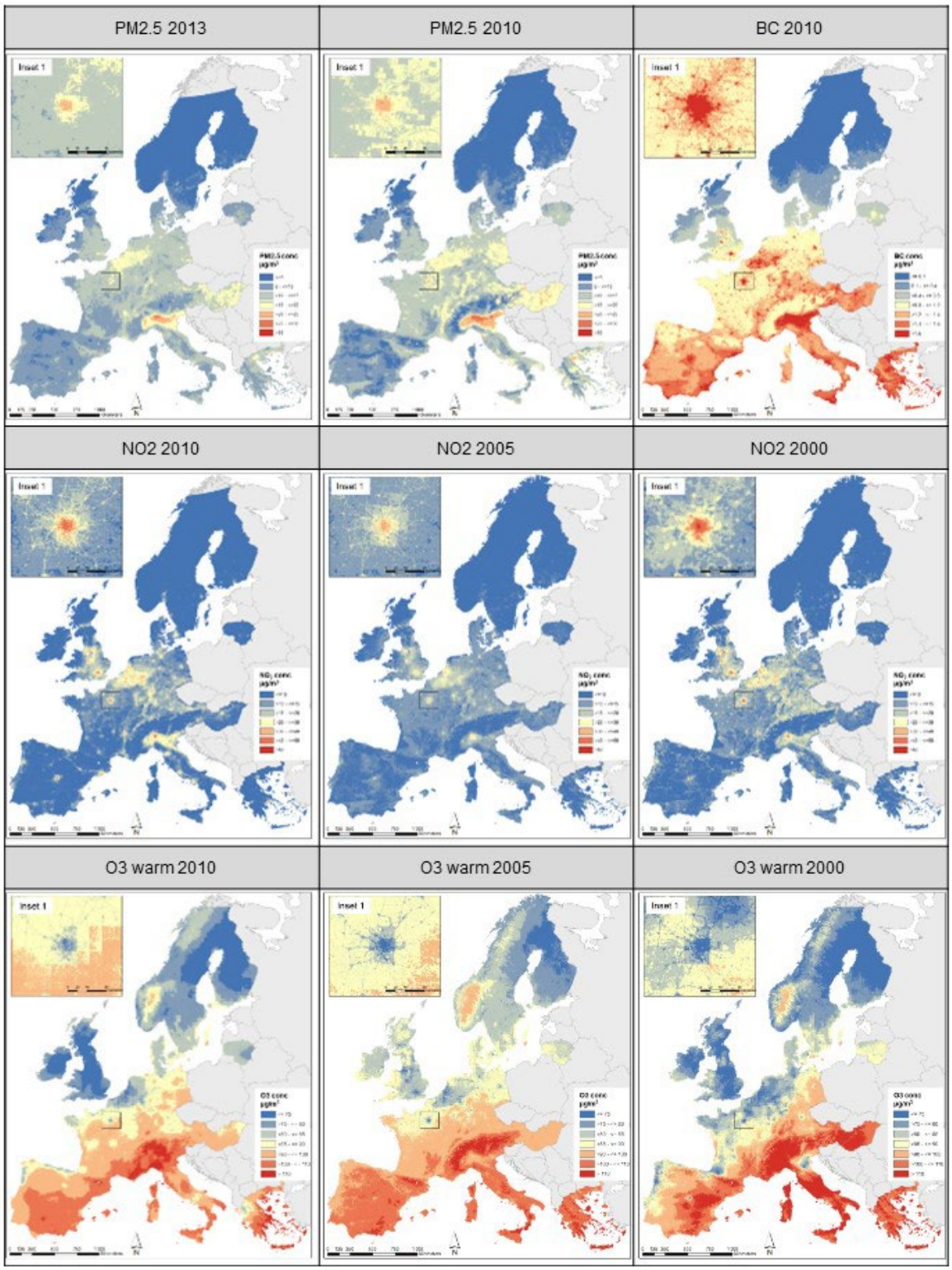


Figure M2 Mapping of hybrid West European LUR models for PM2.5, BC, NO_2 and O_3 warm season (µg/m^3, BC 10^{-5}m^{-1}). (DeHoogh et al. 2018; reproduced with permission from Elsevier)

Appendix M3: Stability of spatial air pollution exposure structure

In back extrapolation we assume that the spatial structure remains for large parts the same going back in time. To investigate the stability of the spatial structure of the 2010 PM2.5, NO_2 and O_3 LUR models, and to test this assumption, we developed LUR models for PM2.5 (2013) and NO_2 and O_3 (2000 and 2005). We then created a random point file across the study area (N = ~44,000) and intersected these points with the 2010 raster surfaces for PM2.5, NO_2 and O_3 (annual, cold season and warm season); the 2013 (PM2.5) and the 2005 and 2000 (NO_2 and O_3) surfaces. The comparison was performed at overall EU countries (relevant for the pooled cohort analysis), combined and individual ELAPSE countries, non-ELAPSE countries and NUTS1 level. We additionally evaluated the correlation of measurements annual average (plus summer and winter average for ozone) for those stations with measurements going back in time sufficiently long.

Table M3 and Figures M3 and M4 show the results of the stability tests at country and NUTS-1 level (more detailed than country). Agreement in spatial variation was generally high at the overall EU country and combined ELAPSE country level (>76%) for all comparisons, except for the O_3 cold season surface of 2000 (44%) when compared to the O_3 cold season surface of 2010. At the country level, focusing on ELAPSE countries only, we observed some heterogeneity in the associations. Both 2000 and 2005 NO_2 surfaces showed a high agreement with the 2010 NO_2 surface (all ELAPSE countries >80%). The agreement between PM2.5 surfaces developed for 2010 and 2013 showed more variability, with four ELAPSE countries >80% (UK, Sweden, Belgium and Italy), the Netherlands 70% and the rest between 48 and 60%. There was a high variability between the associations of the different O_3 surfaces. The agreement between O_3 annual surfaces of 2000 and 2005 with 2010 was reasonable, all ELAPSE countries >60% explained spatial variability, with the exception of Sweden (2000) with 45%. Except from the 2005 O_3 cold (all ELAPSE countries > 60%), the O_3 cold and warm season surfaces are less stable over time with large ranges of explained spatial variability. Italy performs poorly with 1.6%, 11.9% and 16.6% for respectively 2000 warm season, 2005 warm season and 2000 cold season (combined with the largest RMSE's). Some of the variability may be due to the use of another dispersion in the LUR models for 2010 (based on an ensemble method of multiple models) compared to 2000 and 2005 for which the ensemble model was not available.

At the NUTS1 area level (see Figure 2), like at the country level, there is a good agreement for all areas for NO_2 2000 and 2005 when compared to the 2010 surface (R2 >0.60). For the other pollutants there is more heterogeneity in the correlation coefficients across areas. When comparing the PM2.5 surfaces (2010 vs. 2013), the majority of the NUTS1 areas have a correlation coefficient > 0.40, with only a handful of areas dropping between 0.20 and 0.40. The comparison of the O_3 surfaces (2000, 2005 vs. 2013) shows a clear difference between annual and cold season versus the warm season. Both the 2000 and 2005 comparisons for warm season show a number of areas in the south of Europe with correlations of less than 0.20. This pattern is not observed in the annual and cold season comparisons.

Table M3 Stability analysis at country level

Stability analysis at country level: predictions of the 2010 LUR model versus models from other years at randomly selected points (R2 in percentages, RMSE in µg/m3)

	PM2.5 2013		NO2 2005		NO2 2000		O3 2005a		O3 2000a_L		O3 2005c		O3 2000c		O3 2005w		O3 2000w		
Country	R2 (%)	RMSE	R2 (%)	RMSE	R2 (%)	RMSE	R2 (%)	RMSE	R2 (%)	RMSE	R2 (%)	RMSE	R2 (%)	RMSE	R2 (%)	RMSE	R2 (%)	RMSE	N
All Countries	88.2	1.9	91.9	1.9	90.9	2.0	85.8	3.5	78.8	4.3	80.4	4.3	44.3	7.3	84.3	4.6	76.4	5.6	44.000
ELAPSE countries																			
Combined	89.3	1.9	92.6	2.0	91.4	2.1	82.7	3.2	82.0	3.3	87.0	3.3	45.1	6.9	81.6	4.6	78.3	5.0	34762
Austria	60.1	2.0	86.7	1.3	87.4	1.9	81.9	3.7	82.7	4.1	80.9	3.4	67.4	6.7	82.5	3.4	64.5	3.9	1050
Belgium	84.1	1.0	90.9	1.4	84.6	2.3	81.5	1.9	87.4	1.9	89.6	1.9	81.7	2.4	86.5	2.0	70.6	2.5	352
Switzerland	52.5	1.9	91.5	1.2	92.6	1.8	94.6	2.4	95.2	2.7	88.2	3.3	85.5	5.1	87.9	3.5	88.7	4.6	503
Germany	57.6	1.2	85.0	1.3	80.5	2.2	64.0	2.7	69.2	3.0	75.5	3.1	29.4	4.7	47.3	3.8	63.7	4.4	4232
Denmark	48.8	1.1	88.8	0.8	84.8	1.6	73.0	1.2	71.1	1.3	71.0	1.6	59.6	1.8	63.6	1.5	73.2	1.6	527
France	57.4	1.5	89.0	1.1	82.9	1.9	83.2	2.7	80.4	3.5	87.6	3.0	55.0	5.2	76.3	3.4	86.8	4.1	6475
Italy	82.6	1.7	81.9	1.6	82.6	2.3	59.9	4.4	64.8	4.9	90.0	4.3	16.6	9.8	11.9	5.2	1.6	12.3	3548
Netherlands	70.1	0.9	87.9	1.6	81.9	2.7	60.4	2.2	71.8	2.1	73.0	2.3	35.6	3.0	79.3	2.2	53.1	2.6	454
Norway	59.3	0.9	83.3	0.5	83.4	0.8	88.6	1.7	79.4	2.4	79.0	2.2	71.7	3.1	61.1	3.0	79.4	2.4	3449
Sweden	86.2	0.9	93.1	0.5	91.3	0.8	65.5	1.6	45.1	2.2	78.9	1.7	63.3	2.9	76.6	1.6	87.4	1.7	5353
United Kingdom	89.8	1.2	95.3	1.1	93.0	2.0	71.8	2.0	78.1	2.1	81.9	3.3	74.3	3.4	52.2	2.5	53.0	3.3	2845
Non ELAPSE countries																			
Greece	64.4	1.2	86.5	0.9	83.3	1.6	40.9	3.7	49.5	3.8	14.2	6.6	6.0	7.6	34.7	3.9	19.4	5.1	1549
Finland	44.2	1.0	92.7	0.4	89.7	0.8	52.4	1.0	46.3	1.2	25.2	2.4	67.9	1.6	70.2	1.3	69.7	1.6	4008
Hungary	53.9	0.9	84.3	0.9	84.8	1.2	50.8	1.3	38.4	1.6	21.6	3.9	59.4	2.5	54.1	1.2	38.6	2.3	1118
Ireland	73.9	0.8	92.7	0.6	90.2	1.0	52.0	1.3	49.1	1.4	79.1	2.4	68.8	2.2	61.1	1.2	61.6	2.3	841
Lithuania	56.3	0.9	89.7	0.6	85.1	1.0	52.9	1.1	40.8	1.2	65.3	1.9	74.7	1.4	54.8	0.9	24.4	1.7	780
Luxembourg	68.3	0.9	89.0	1.3	77.9	2.2	73.9	1.3	75.4	1.4	74.1	2.6	78.3	2.2	47.2	1.8	57.7	1.4	31
Portugal	63.8	1.1	85.4	1.0	87.0	1.6	71.3	1.9	67.4	2.2	62.1	3.3	51.5	3.5	33.0	2.4	37.4	3.9	1015
Spain	69.4	1.1	77.8	1.2	79.7	1.7	65.6	2.8	58.5	3.6	62.8	4.4	41.4	5.6	42.9	3.4	38.9	7.0	5974

Figure M3 Stability of spatial surface analysis at the country level. (DeHoogh et al. 2018; reproduced with permission from Elsevier)

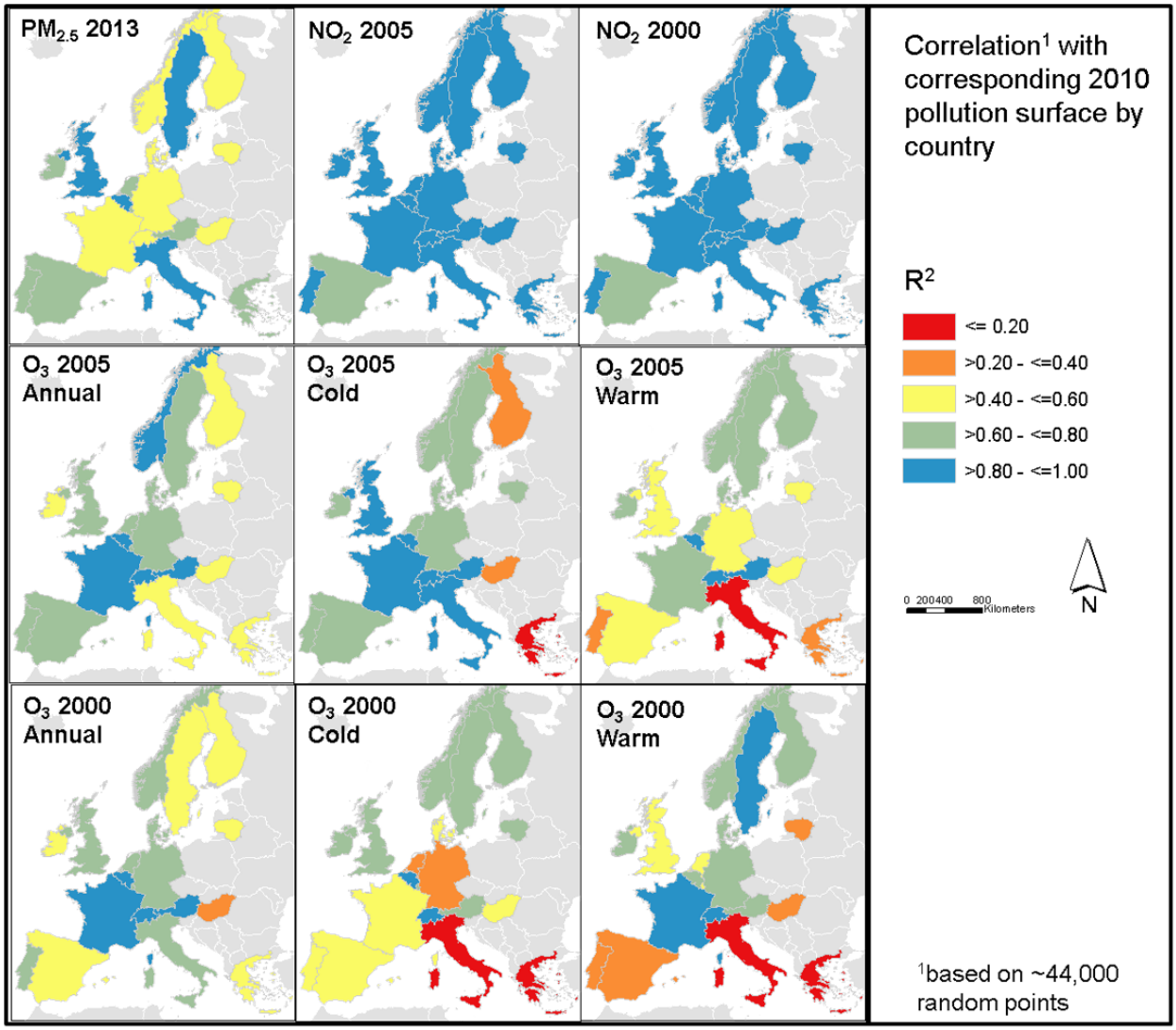
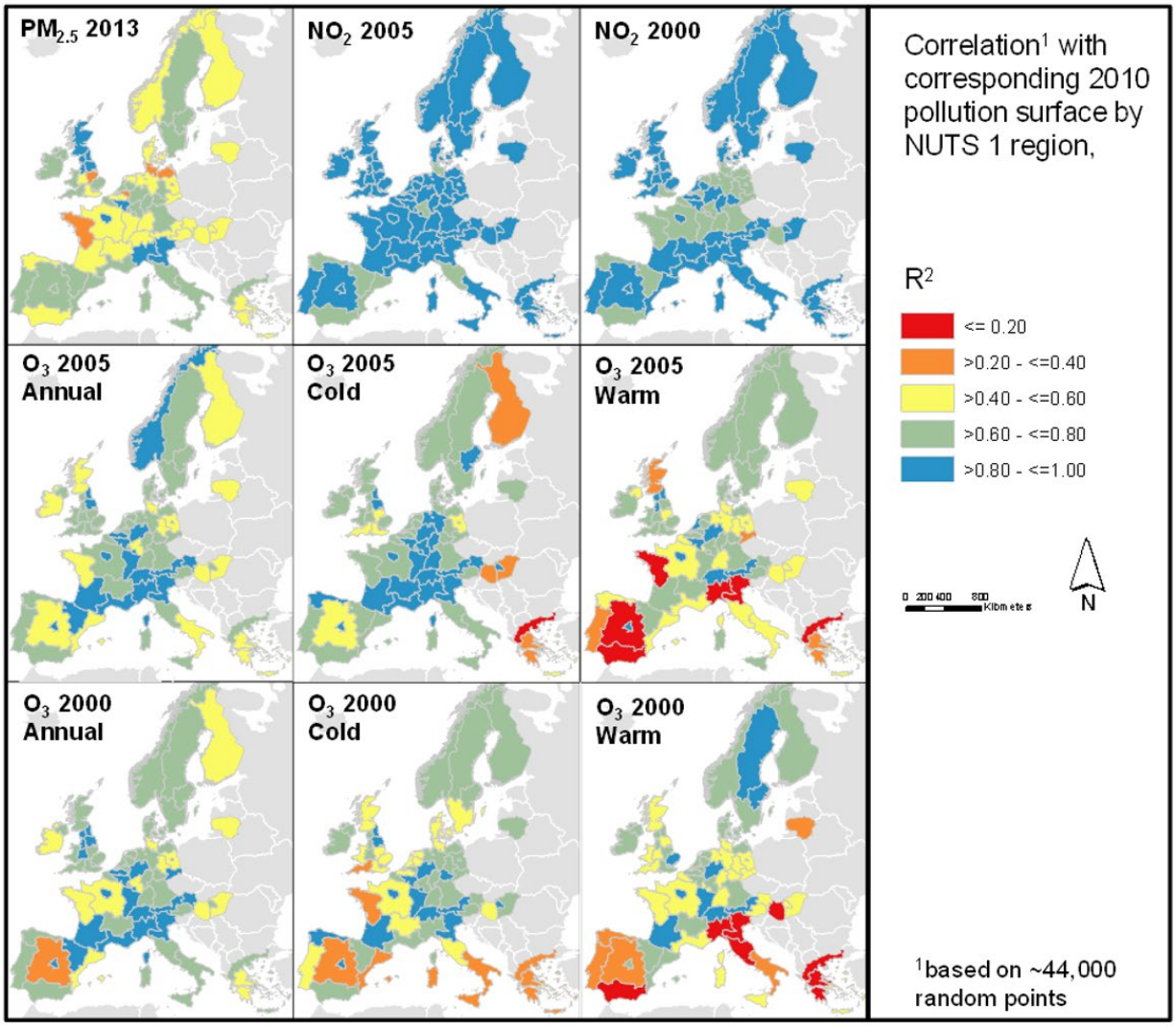


Figure M4 Stability of spatial surface analysis at the NUTS1 level (a more detailed level than country). (DeHoogh et al. 2018; reproduced with permission from Elsevier)



We evaluated the relationship between measured annual average concentrations (plus summer and winter average for ozone) for those AIRBASE stations with measurements going back in time sufficiently long between 2010 to 2005 and 2000 (see Table M3). Correlations are high at the EU level between 2010 - 2005 and 2010 - 2000 measured concentrations for NO_2 and O_3 (R2 between 0.68 – 0.87) and between measured PM2.5 2010 - 2013 (R2 0.79). Correlations for NO_2 remain high when focusing on ELAPSE countries (R2 > 0.66) for both year comparisons. The majority of ELAPSE countries observed high correlations for the other pollutants as well, with the exception of Italy and the Netherlands for both O_3 annual and O_3 warm (R2 < 0.38) for the 2 year comparisons, Norway for PM2.5 (R2 = 0.15) and O_3 annual 2000 (R = 0.03), Denmark O_3 warm 2005 (R2 = 0.19) and Sweden for O_3 annual, warm and cold season 2005 (R2 <0.30).

Table M4 Correlations between concurrent AIRBASE measurements in 2010 and 2000 and 2005 (NO_2, O_3 annual, warm and cold season) and 2013 (PM2.5) as R^2 (number of sites) for EU and separately for ELAPSE countries.

	NO_2		O_3 annual		O_3 warm		O_3 cold		PM2.5
	2000	2005	2000	2005	2000	2005	2000	2005	2013
EU	0.86 (546)	0.87 (794)	0.72 (572)	0.72 (836)	0.68 (576)	0.68 (843)	0.78 (555)	0.80 (817)	0.79 (247)
Austria	0.86 (66)	0.95 (77)	0.88 (77)	0.90 (86)	0.72 (79)	0.80 (88)	0.91 (75)	0.92 (84)	0.97 (8)
Belgium	0.95 (16)	0.93 (26)	0.88 (22)	0.88 (28)	0.77 (22)	0.76 (29)	0.92 (22)	0.95 (25)	0.86 (19)
Switzerland	0.98 (21)	0.95 (21)	0.91 (21)	0.89 (23)	0.75 (21)	0.86 (23)	0.97 (21)	0.92 (23)	n.a. (0)
Germany	0.91 (185)	0.94 (213)	0.73 (181)	0.78 (206)	0.58 (182)	0.60 (206)	0.80 (175)	0.88 (201)	0.46 (63)
Denmark	n.a. (2)	0.93 (6)	n.a. (0)	0.41 (6)*	n.a. (0)	0.19 (6)*	n.a. (0)	0.73 (6)	0.95 (3)*
France	0.86 (169)	0.90 (261)	0.71 (179)*	0.82 (301)	0.66 (184)	0.82 (307)	0.80 (173)	0.86 (294)	0.53 (57)
Great Britain	0.88 (27)	0.90 (44)	0.73 (35)	0.72 (55)	0.67 (31)	0.66 (51)	0.78 (35)	0.77 (54)	0.59 (28)
Italy	0.66 (30)	0.74 (109)	0.38 (26)	0.21 (87)	0.20 (26)	0.01 (90)*	0.75 (23)	0.68 (88)	0.84 (44)
Netherlands	0.89 (23)	0.93 (26)	0.30 (19)	0.30 (25)	0.01 (19)*	0.03 (25)*	0.59 (20)	0.70 (23)	0.68 (15)
Norway	n.a. (2)	1.00 (3)	0.03 (6)	0.50 (7)	0.46 (6)*	0.72 (7)	0.73 (6)	0.91 (7)	0.15 (5)*
Sweden	0.97 (5)	0.97 (8)	0.68 (6)	0.01 (12)*	0.41 (6)*	0.15 (11)*	0.93 (5)	0.30 (12)	0.84 (5)

*not significant (p>0.05)

Appendix M4: Back-extrapolation procedures

Aim: to estimate a pollutant concentration for each year from recruitment to end of follow-up.

In ideal circumstances we would have monitoring sites operational throughout the study period at locations representative for the cohort population for each pollutant. We would then use the temporal signal in the monitoring data and apply both the ratio and absolute difference methods to scale back the 2010 modelled PM2.5, NO_2, O_3 and BC estimates. To this end we downloaded all available AIRBASE data from 1990 up to 2014 for all four pollutants for our study areas.

We also obtained data from the Danish Eulerian Hemispheric Model (DEHM) developed by Ole Hertel and colleagues (Brandt et al. 2012). DEHM models for all four pollutants on a 26 x 26 km scale across Europe from 1979 to 2015 (downscaled from an original 50 x 50km resolution using bi-linear interpolation). This is modelled data based on emission and meteorological data.

AIRBASE sites and DEHM data were intersected with the areas shown in Figure M5. For the administrative cohorts this means that the whole country boundary was used, whereas for the pooled cohorts a circle was drawn. Due to the low density of monitoring sites around Oslo and Stockholm the circles were made larger in order to pick sufficient monitoring sites.

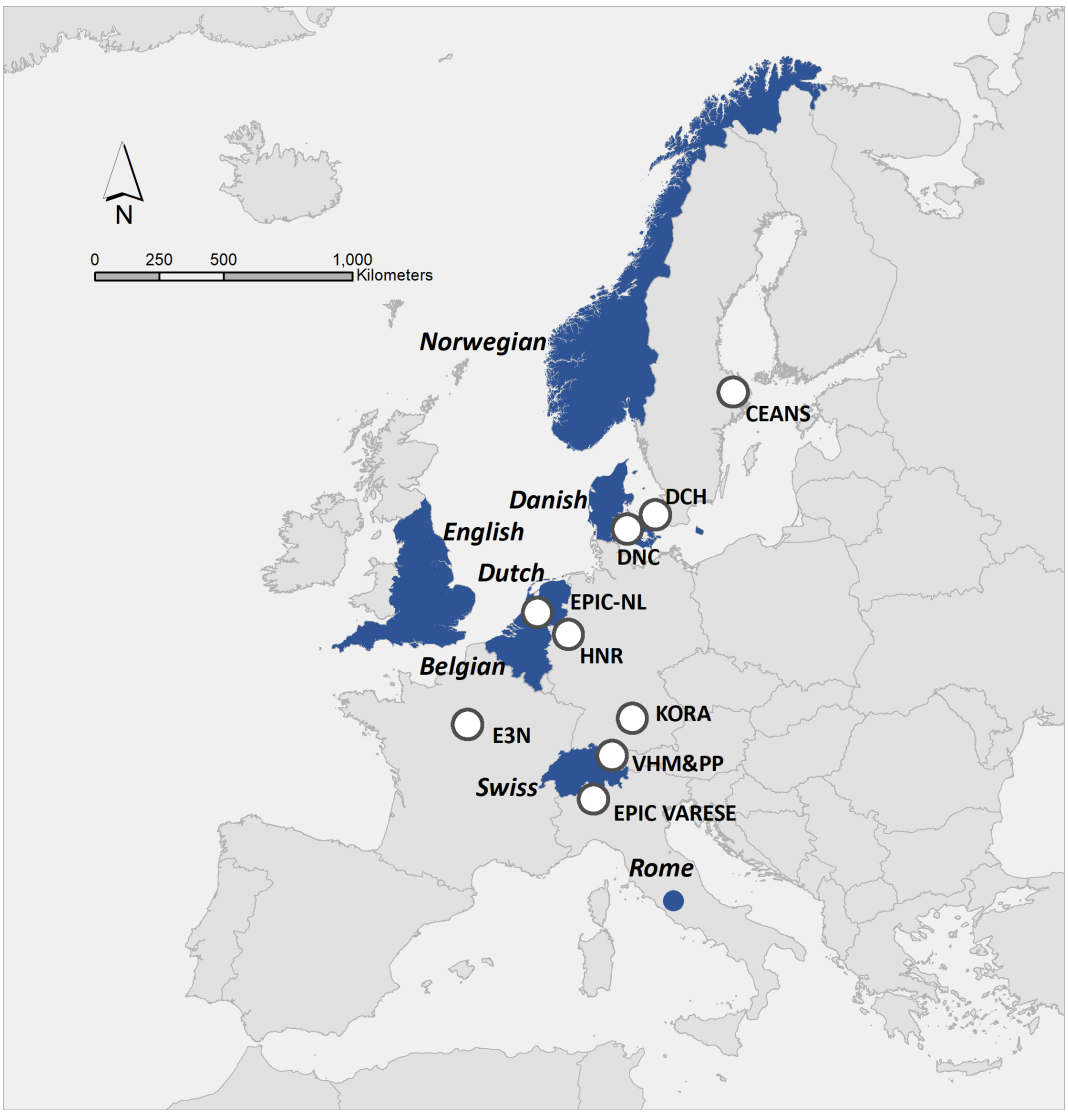


Figure M5: ELAPSE Administrative areas (blue areas) and cohort areas (circles)

AIRBASE data

We only selected those AIRBASE monitoring sites which were operational throughout the predefined time periods (i.e. 1990 - 2010, 1995 - 2010, 2000 - 2010, 2005 - 2010) and which were classified as urban or rural background. Traffic sites are too sensitive to specific changes that only apply to that particular site.

For each area the annual average pollutant concentration (NO_2, PM2.5, PM10 and O_3) and additionally the cold and warm periods for O_3 were calculated based on the AIRBASE dataset.

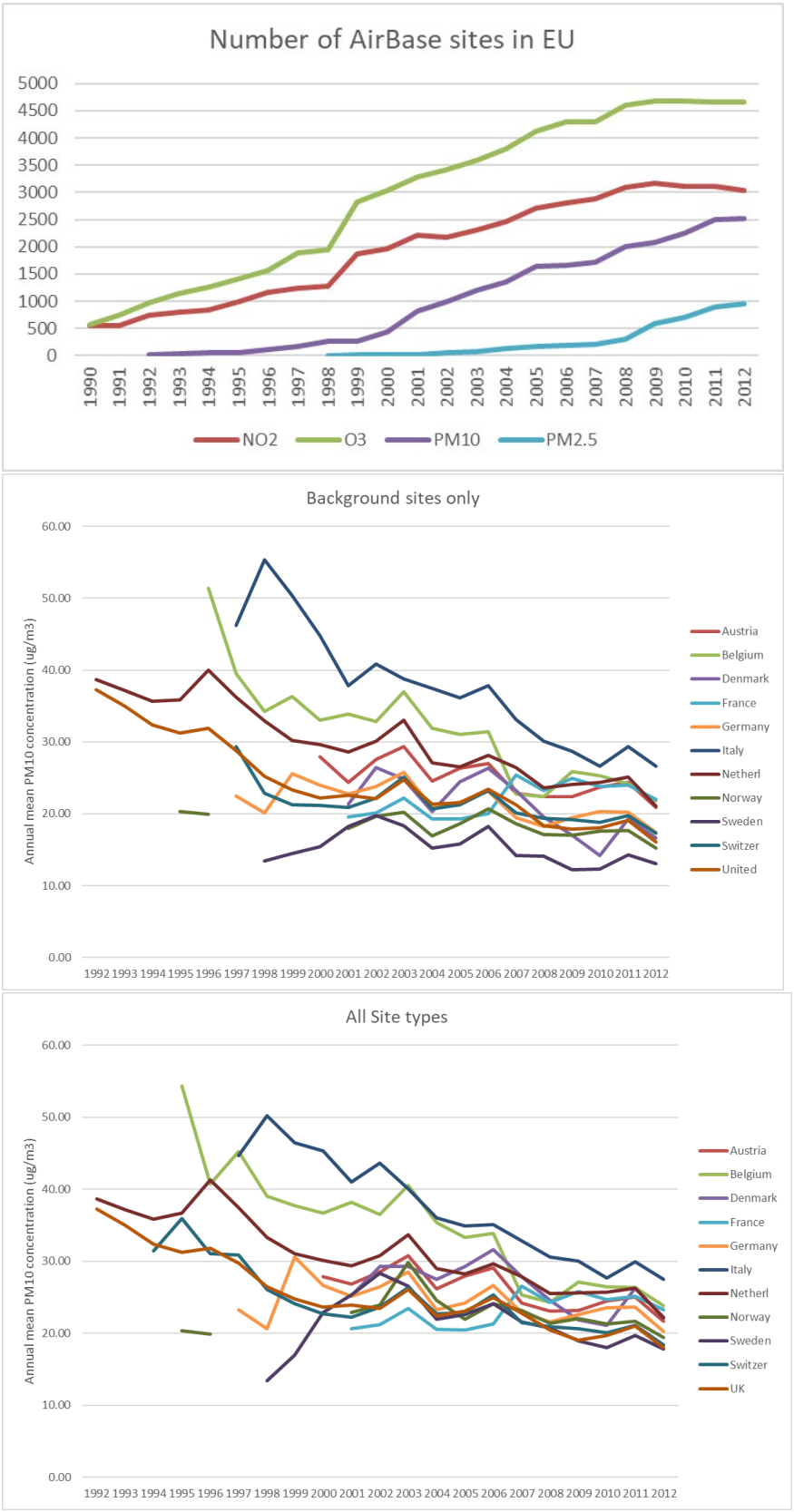
Table M5 and figure M6 shows the number of monitoring sites operational throughout the defined time periods in the different areas.

Table M5: Number of operational urban and rural background AirBase sites operating continuously during predefined time periods within the administrative and cohort areas.

Pollutant	Admin areas	2005-2010	2000-2010	1995-2010	1990-2010	Cohort areas	2005-2010	2000-2010	1995-2010	1990-2010
PM10	Belgium	18	3	0	0	Stockholm	1	1	0	0
	Denmark	1	0	0	0	HNR	7	0	0	0
	France	116	0	0	0	KORA	1	0	0	0
	Netherlands	20	11	10	5	Oslo	1	0	0	0
	Norway	1	0	0	0	Rome	0	0	0	0
	Switzerland	20	12	0	0	Varese	1	0	0	0
	UK	6	4	1	0	Voralberg	2	0	0	0
NO_2	Belgium	21	12	5	2	Stockholm	2	2	0	0
	Denmark	5	2	0	0	HNR	8	7	4	4
	France	246	148	1	0	KORA	2	1	1	0
	Netherlands	23	19	17	7	Oslo	0	0	0	0
	Norway	3	2	0	0	Rome	5	2	0	0
	Switzerland	22	18	17	2	Varese	4	1	0	0
	UK	12	10	3	0	Voralberg	4	2	1	0
PM25	Belgium	2	0	0	0	Stockholm	0	0	0	0
	Denmark	1	0	0	0	HNR	0	0	0	0

	France	5	0	0	0	KORA	0	0	0	0
	Netherlands	0	0	0	0	Oslo	1	0	0	0
	Norway	1	0	0	0	Rome	0	0	0	0
	Switzerland	0	0	0	0	Varese	0	0	0	0
	UK	1	1	0	0	Voralberg	0	0	0	0
O_3	Belgium	27	21	3	0	Stockholm	2	2	1	0
	Denmark	6	0	0	0	HNR	5	3	2	0
	France	292	153	0	0	KORA	1	0	0	0
	Netherlands	20	15	6	0	Oslo	2	2	0	0
	Norway	5	5	0	0	Rome	4	4	0	0
	Switzerland	22	20	18	0	Varese	3	1	0	0
	UK	24	19	10	0	Voralberg	3	2	2	0

Figure M6 Trends of number of monitoring sites in Europe. (From Wolf et al., in press)



DEHM data

Similarly, mean DEHM concentrations were calculated for each area (annual for NO_2, PM2.5, PM10, O_3; cold/warm for O_3). For the administrative cohorts we also looked at the within-country variation of temporal trends; for this we used the NUTS1 boundaries (see Figure M7). As a last step population weighted NO_2, PM2.5, O_3 and BC DEHM concentrations were calculated at the NUTS1 level. 2011 European population estimates at the 1x1km level were obtained from GEOSTAT (GEOSTAT-grid-POP-1K-2011-V2-0-1.zip downloaded from http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat).

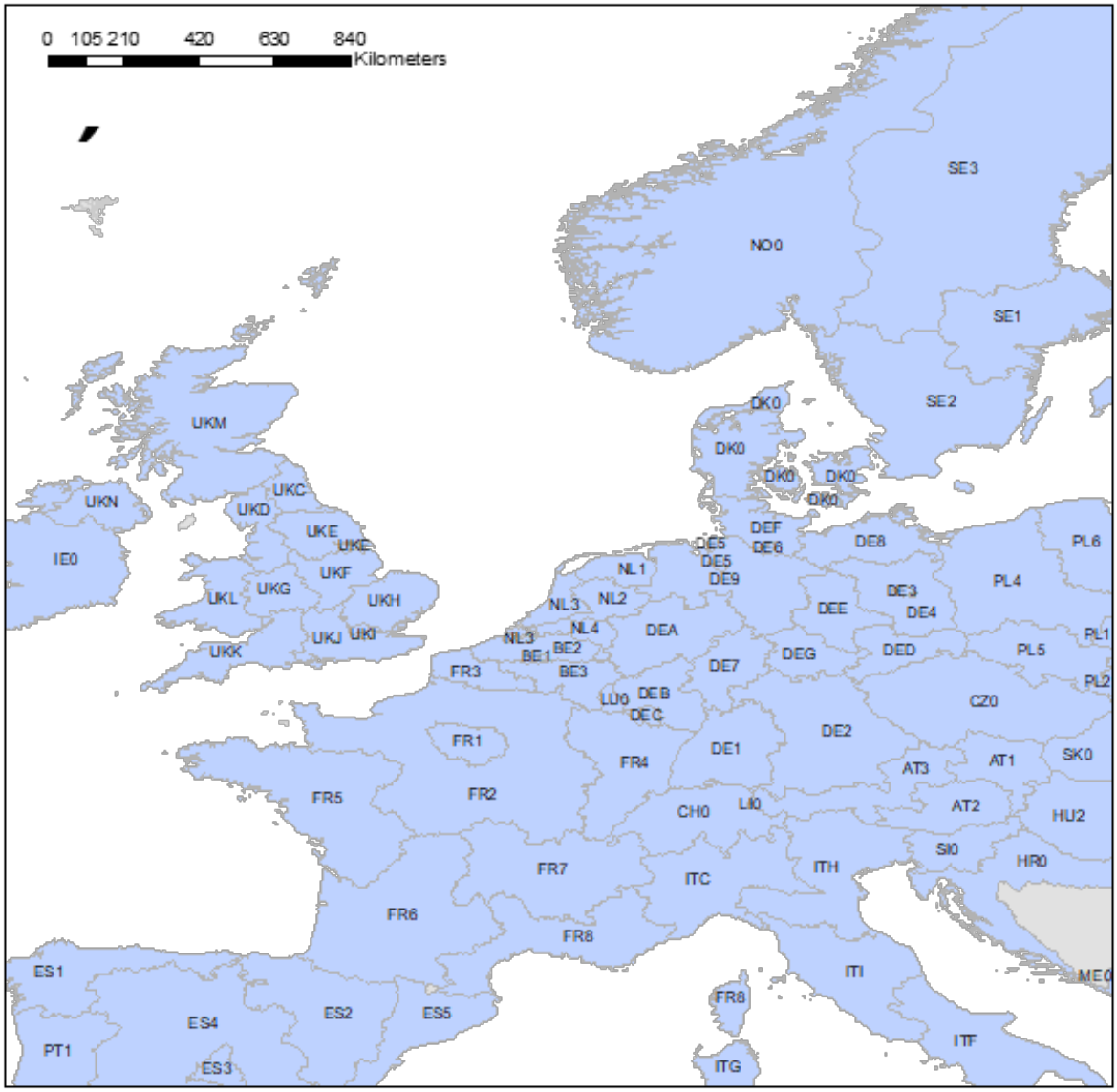


Figure M7: NUTS1 regions. (Source: Eurostat)

Methods and Results

For back-extrapolation solid measurement data from urban or rural background sites located in the study area for 2010 (ELAPSE exposures) back to the baseline period are needed. Table M5 shows the number of AIRBASE sites operational during set time periods (2005-2010, 2000-2010, 1995-2010 and 1990-2010) for the different pollutants. For PM2.5 and PM10 it is clear that no solid data exist going back in time. Already for the 2000-2010 period most administrative and cohort areas have no sites available with continuous measurements for PM2.5 and PM10. For NO_2 and O_3 there is more data available, but also for these two pollutants gaps are appearing for the 1995-2010 period. Our conclusion looking at these data is that the AIRBASE data does not give us a solid base to perform a harmonious back-extrapolation across all administrative and cohort areas.

DEHM data is complete - annual averages for all pollutants back to 1979 covering all of the study areas – and therefore has the potential to act as a solid base for back-extrapolation. We performed a couple of checks to investigate how well the temporal trend in predicted DEHM concentrations reflect trends in measured AIRBASE data. These checks could only be performed in areas where sufficient temporal AIRBASE coverage exists and we therefore concentrated on the administrative areas. We performed the check as follows:

For the administrative study areas we calculated a population weighted DEHM value per NUTS1 region and we compared these with the average values with AIRBASE sites (only urban/rural background) which were operational throughout the indicated time period (see Figure M8 showing the trend for NO_2). Trends are fairly similar for modeled and measured data. There is more variability in the measured data, reflecting either more weather-related variation or changes in monitoring configuration / site characteristics. Modelled trends for different NUTS1-areas in a country are similar but the absolute difference between the lowest and highest area are typically smaller in more recent years. Measured concentrations are typically higher than modeled concentrations, related to the 50*50 km scale of the model. Norway is the exception. For most countries the difference between the country-average measurements and the NUTS1-area with the highest concentrations (e.g. the Paris area) are modest. Differences across countries reflect differences in siting and possibly differential validity of the model. In Norway the 2 sites are regional background sites with low measured values, whereas in other countries sites are a mixture of urban and regional background sites.

Figures M8-M11 show the trends for PM2.5, NO_2 and O_3 annual and O_3 warm season average. For PM2.5, there is only limited data available. O_3 concentrations showed very small trends in time, decreasing the need for back-extrapolation. The trends in some countries (France, Belgium) did not agree between measured and modelled concentrations, but the number of sites (2 in France and 1 in Belgium) is too small to draw meaningful conclusions. The increase observed in 2009-2010 of the measured PM2.5 concentrations is unlikely to be correct. Given the regional nature of PM2.5, very different trends across neighboring countries (e.g. Netherlands and Belgium) are unlikely. To further interpret the PM2.5 trends, we added trends in measured concentrations of PM10 (figure M12). With the exception of France (and Sweden in the first years), a decreasing trend in concentrations was found, that was generally steeper prior to 2000.

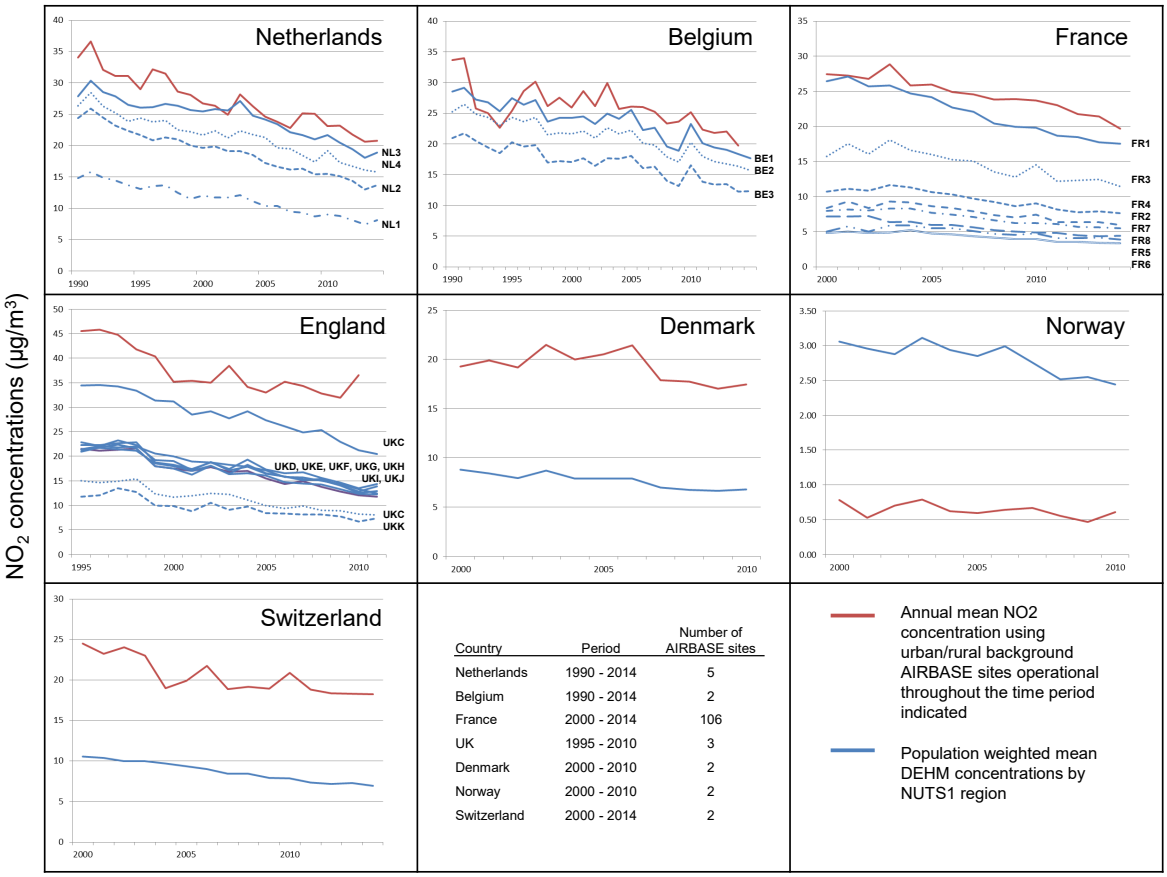


Figure M8: Comparing temporal trends between AIRBASE and DEHM NO_2 data for the administrative study areas. (From Wolf et al., in press)

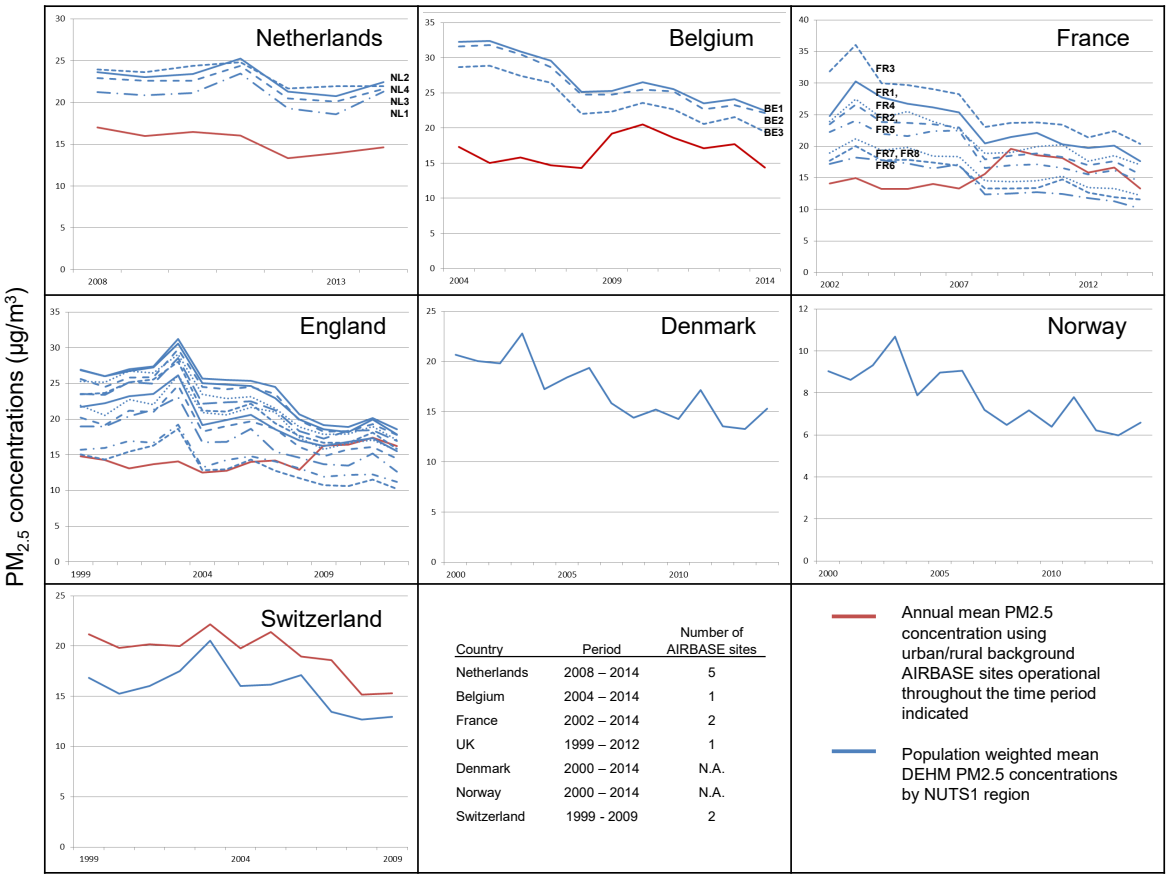


Figure M9: Comparing temporal trends between AIRBASE and DEHM PM2.5 data for the administrative study areas. (From Wolf et al., in press)

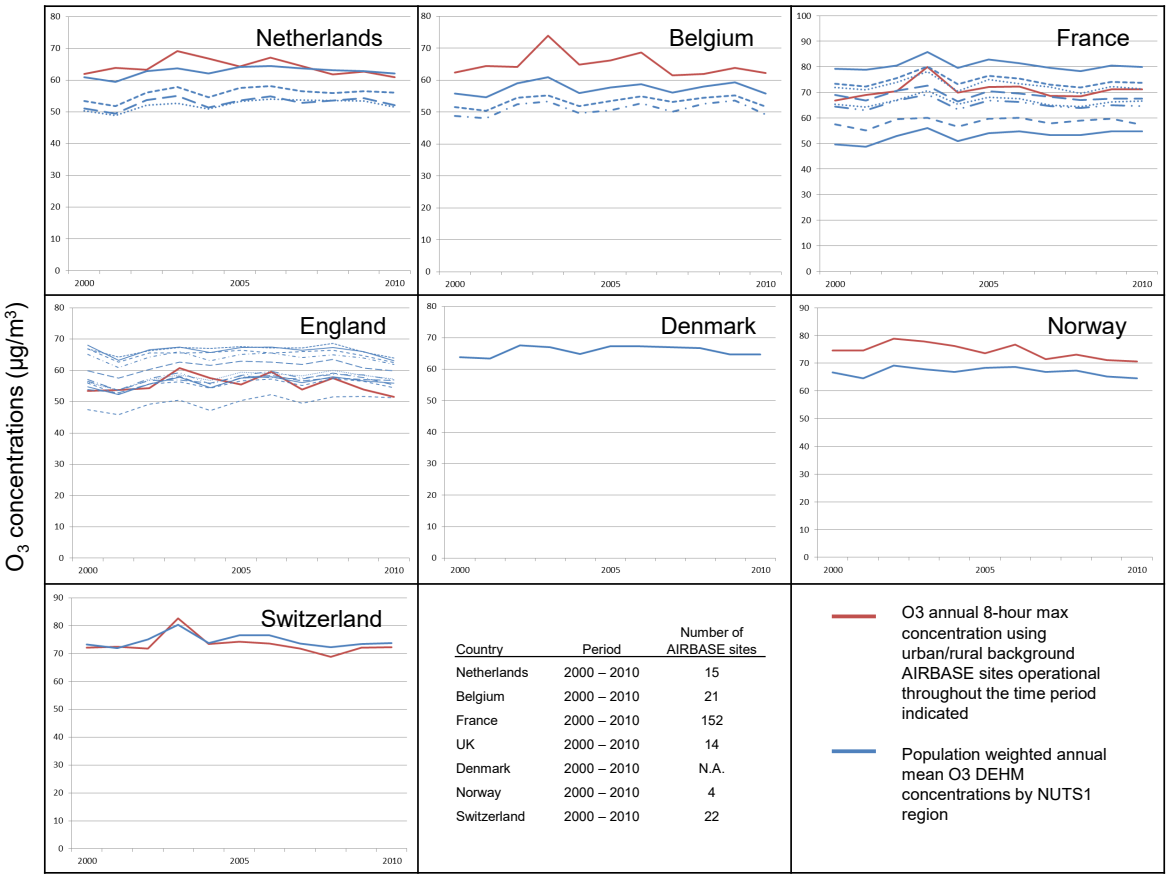


Figure M10: Comparing temporal trends between AIRBASE and DEHM O_3 annual data for the administrative study areas. (From Wolf et al., in press)

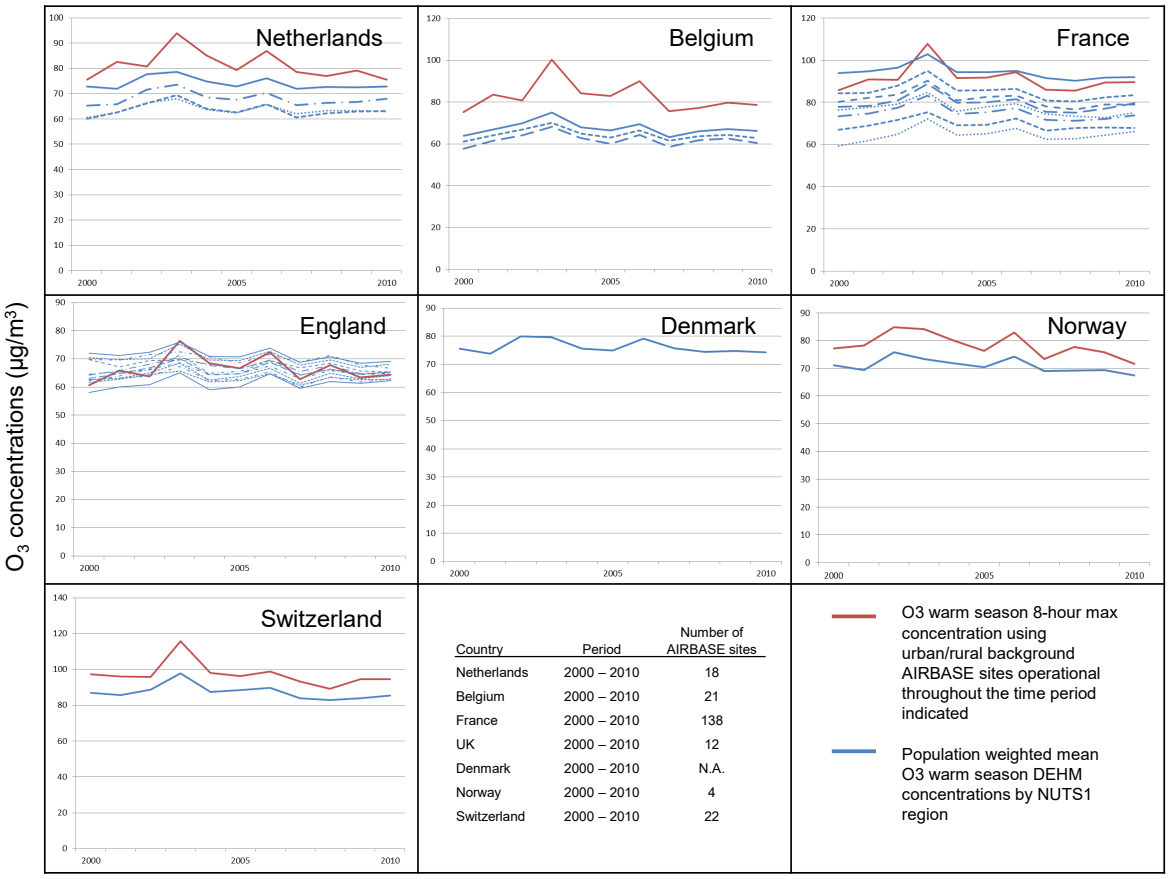
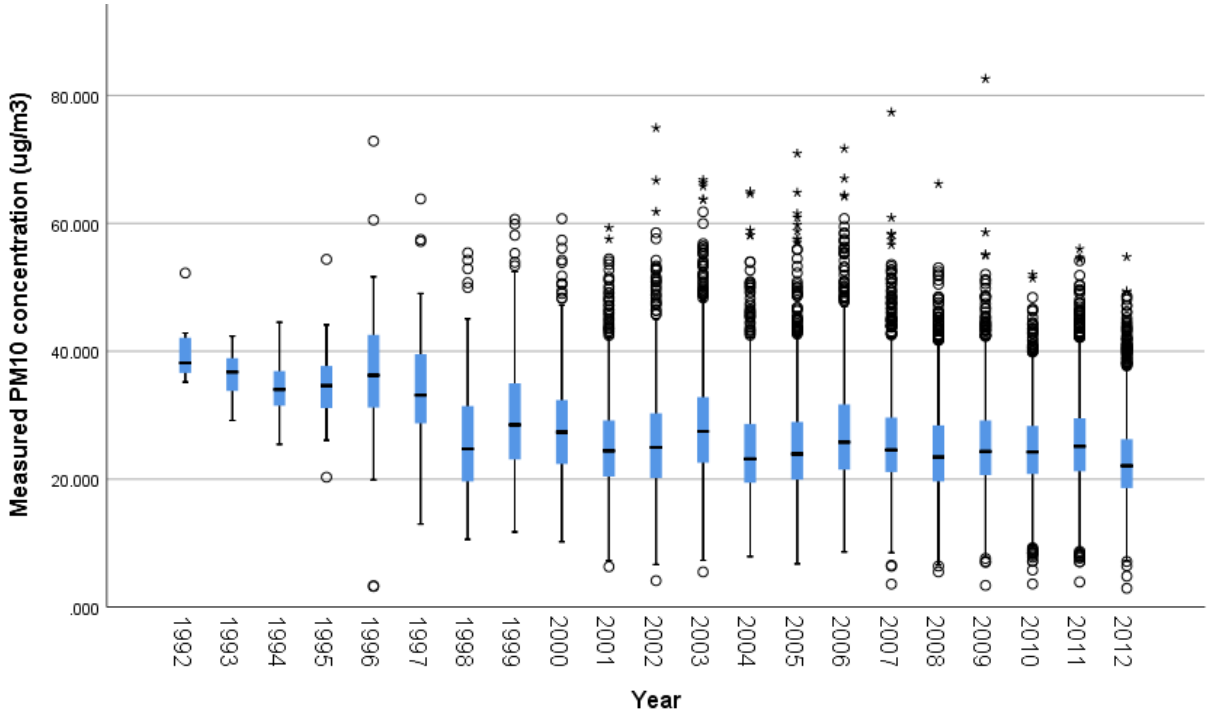


Figure M11: Comparing temporal trends between AIRBASE and DEHM O_3 warm season data for the administrative study areas

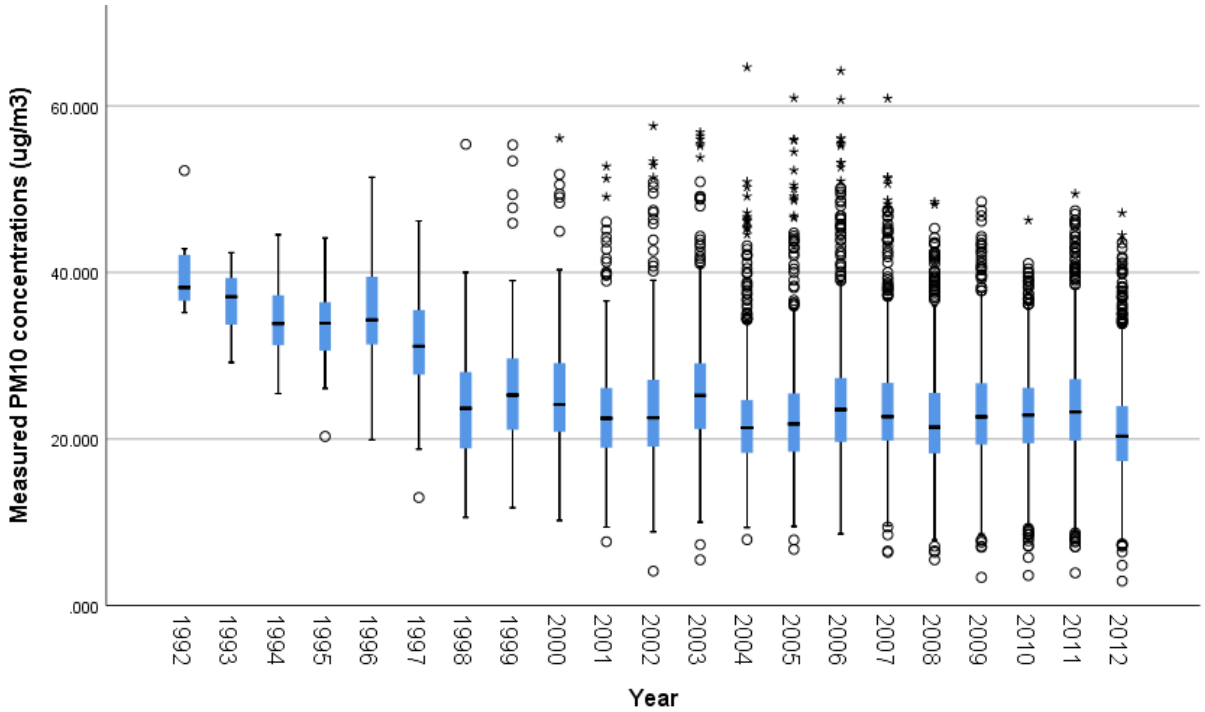
For BC, measurements were available for one year only across the EU. Measurements of Black Smoke in the UK and the Netherlands showed strongly decreasing trends in time, similar to trends in modelled data of BC. BC and Black Smoke cannot be quantitatively compared. Advantages of using the modelled data included completeness and consistency of method. For PM2.5 – a key pollutant- there is virtually no useful monitoring data. Disadvantages include the spatial scale and possibly lower validity compared to measured data.

Figure M12 Trends of measured PM10 concentrations over Europe (only including ELAPSE countries) and specific countries

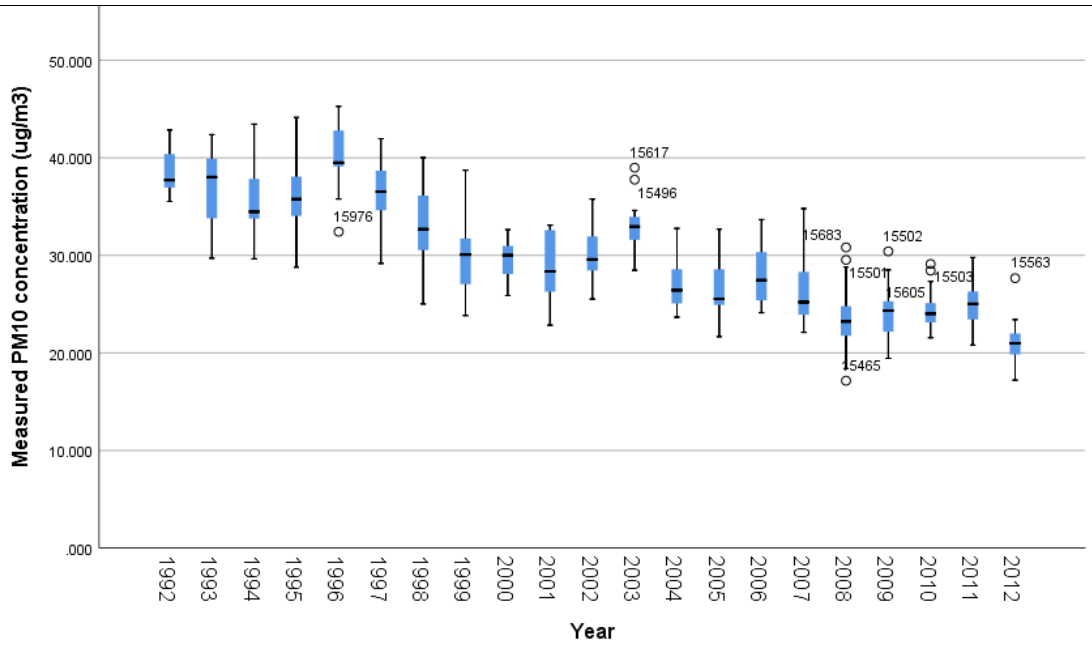
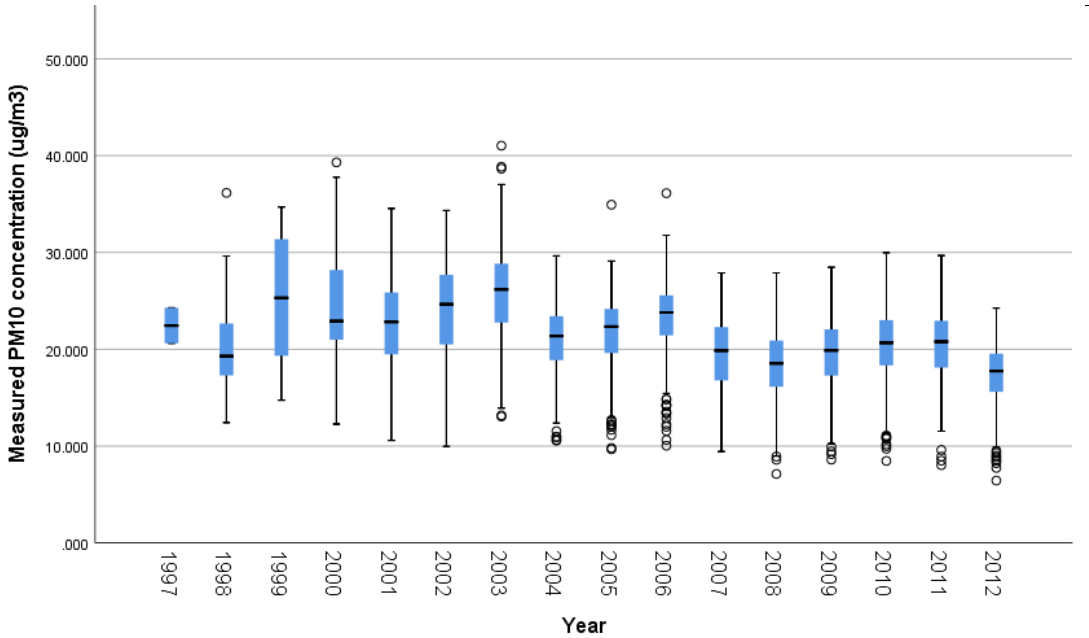
(a) Trend in measured PM10 concentrations at Airbase sites using all sites

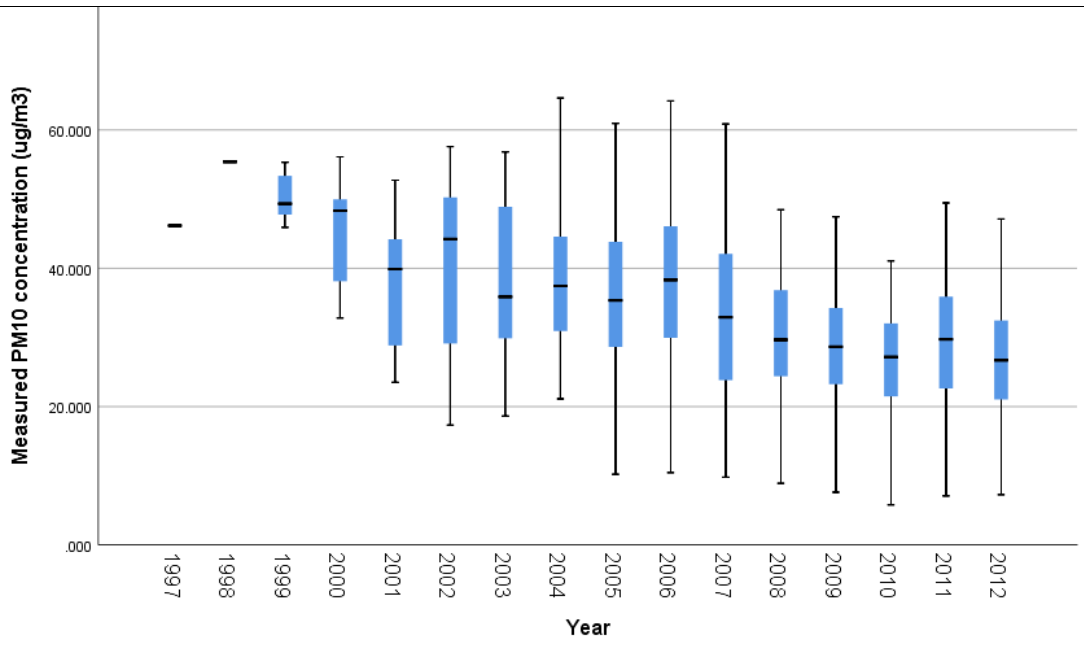
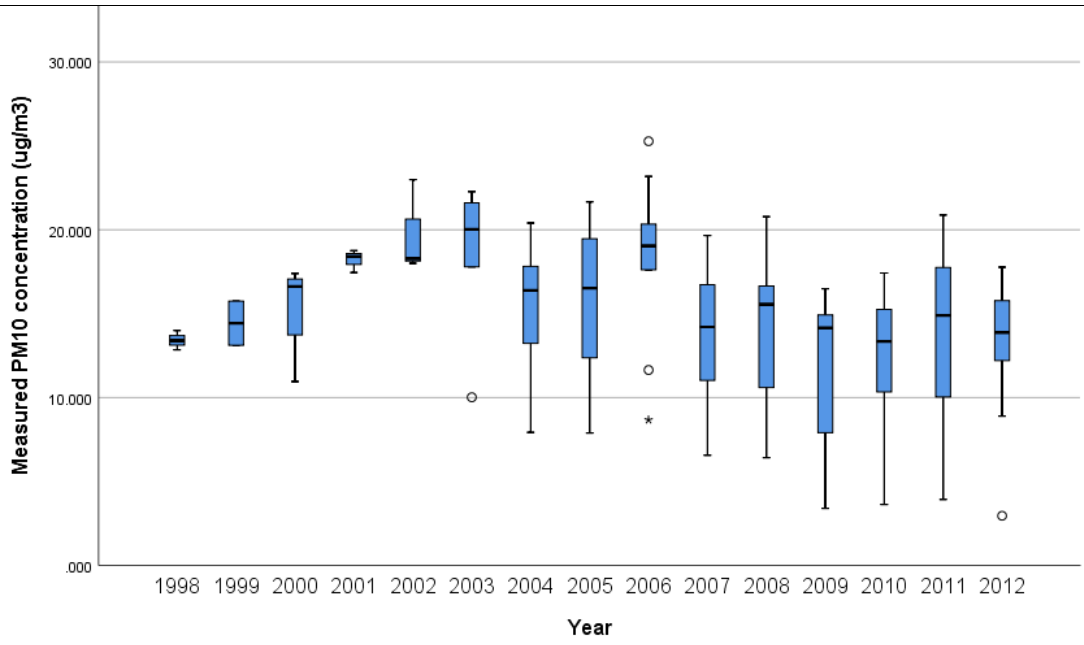


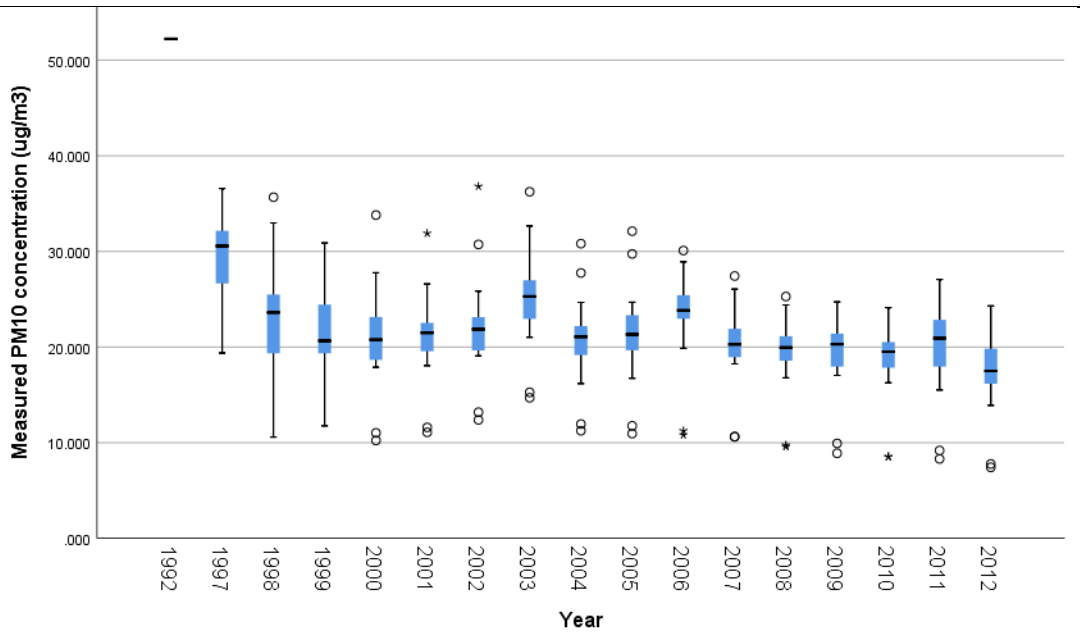
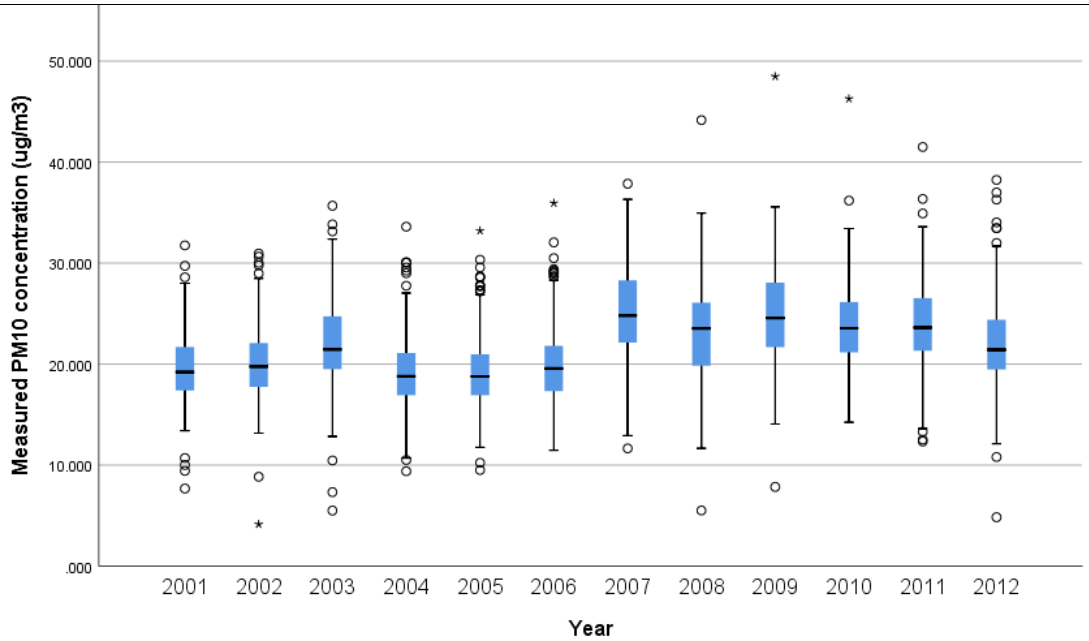
(b) Trend in measured PM10 concentrations at Airbase sites using background sites only

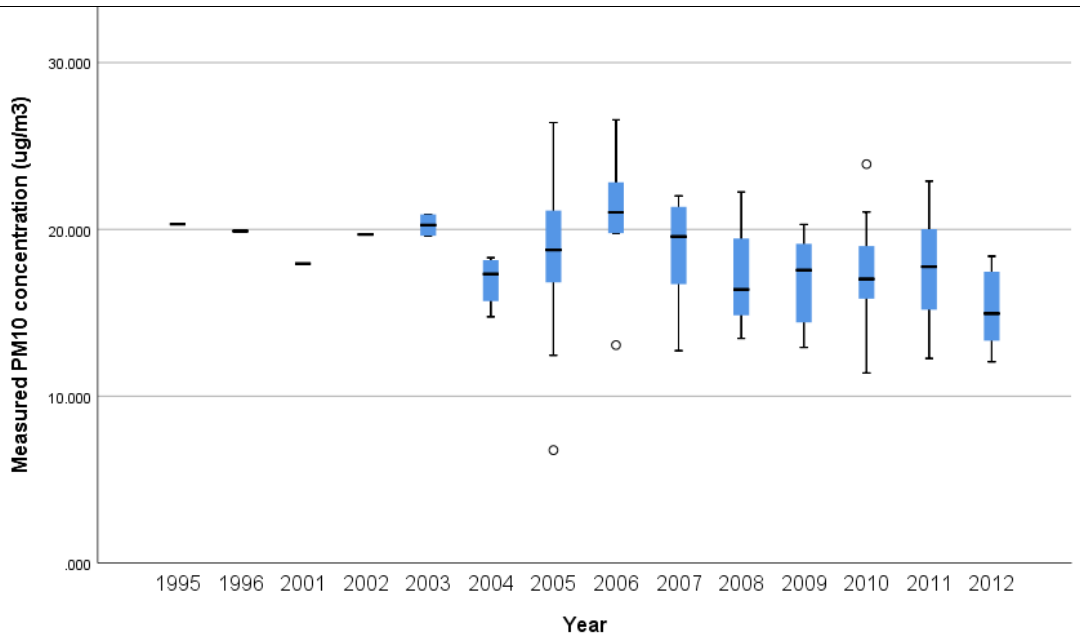
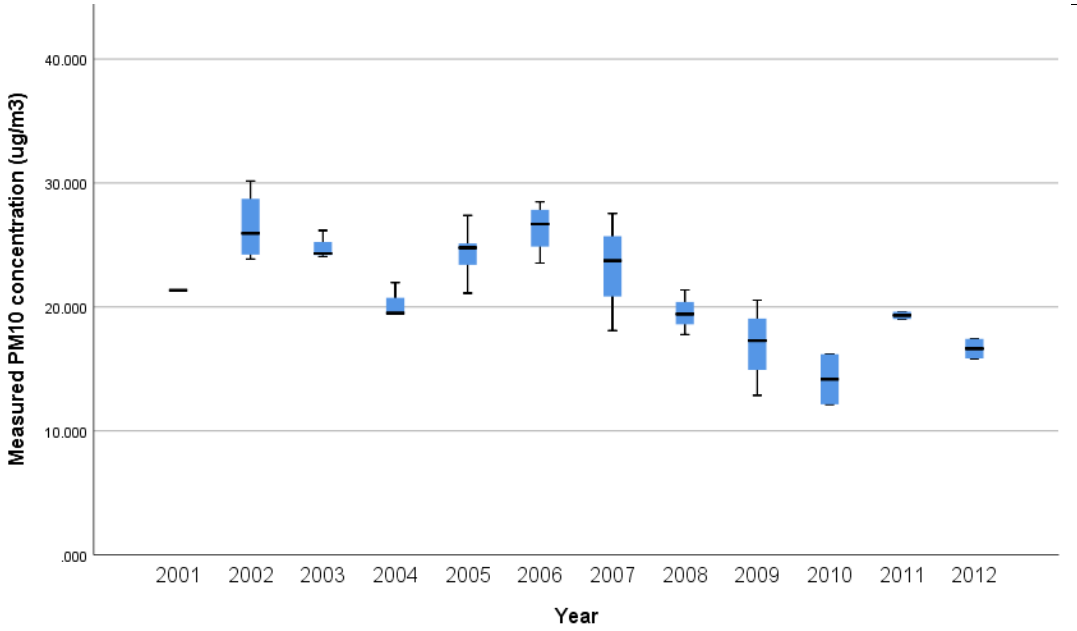


(c) Trends in measured PM10 per ELAPSE country, only using Background sites

NL	
DE	

IT	 Measured PM10 concentration (ug/m3) 60.000 40.000 20.000 .000 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 Year
SE	 Measured PM10 concentration (ug/m3) 30.000 20.000 10.000 .000 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 Year

CH	 Measured PM10 concentration (ug/m3) 50.000 40.000 30.000 20.000 10.000 .000 1992 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 Year
FR	 Measured PM10 concentration (ug/m3) 50.000 40.000 30.000 20.000 10.000 .000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 Year

DK	
NO	

Conclusion

We used the trends predicted by the DEHM for all four pollutants to obtain annual average concentrations for all years from recruitment up to end of follow-up. To allow for varying trends per country, we used the population weighted average of all 26 x 26 km cells within NUTS-1 region for national cohorts. For smaller study areas, we used the population weighted average of all 26 x 26 km cells in the (approximated) study area.

Appendix M5: Area-level SES variables ELAPSE

In the ELAPSE study manual (version 2, 31-10-2016) we identified that area-level SES variables were needed as potential confounders in the epidemiological analysis of the pooled cohort and the administrative cohorts. The Area-level SES Workgroup identified that harmonized SES data were not available from European databases for the neighborhood scale. This implies that all local partners were asked to obtain the area-level SES data. We specified what data should be obtained and linked to the cohort data. In the section "**Definition of data"** we define the data that we wanted the teams to obtain. In the section **Background**, the rationale for this selection is given, including quotes of the relevant sections of the study manual and text sections from recent main cohort studies including the other cohorts used in the HEI low level air pollution call.

Definition of data

Tables M6 and M7 define the data that we aimed to obtain, realizing that not all data are available (such as mean income in the UK) or necessary (e.g. regional scale for the Rome Longitudinal cohort).

All variables were collected for a small area (neighborhood) and large area (region) to allow for confounding at commonly used spatial scales in previous cohort studies.

A **neighborhood** is a part of a city, with about 1,000 – 10,000 people. Ideally, we use a standard definition, referring to externally defined areas. If this is not available, postal codes were used, if they refer to the approximate number of people defined above. Examples include "buurt" and "wijk" in the Netherlands including on average 1,400 and 6,000 subjects; parish or census district (~4,300 subjects) in Denmark; Lower Super output area in the UK. In Switzerland (SNC) postcode boundaries (n=4,156) were used for neighborhood. Quite a few of the ESCAPE cohorts have used municipality (or local administrative unit 2 (LAU2, former NUTS5)) in the ESCAPE project. We now aimed at a finer spatial scale, given that many of the included ESCAPE cohorts include a large metropolitan area with surrounding smaller towns. For smaller towns (e.g. below 10,000 subjects), the town (community) level was deemed to be fine. In some countries, data are available for multiple scales within the specified range. Availability of type of data (in multiple years) and comparability with other cohorts were criteria to select the scale. A very fine scale e.g. below 1,000 subjects is problematic for computational reasons (random effect models), particularly if the outcome is relatively rare. Table M8 lists as an example the available scales in the UK. The smallest unit is the Census Output Area (COA) which is used as the building blocks for all other geographical units. Not all information is available for all levels. Census data (including information on ethnicity, education, unemployment) is available for all levels; information from administrative sources such as % population on income support for Lower-layer Super Output Areas (LSOA) upwards. LSOA or ward were the scales to select.

The **region** is important for national cohorts. In the Dutch Duels cohort, COROP area was used, classifying the Netherlands in ~30 regions. In Switzerland (SNC) canton (n=26) was used for region.

Each cohort defined this locally and judged whether a region scale is needed. For the Rome cohort, a region scale is not needed. See also the examples in North American studies. When both a neighborhood and region scale is used, neighborhood should be nested within region.

SES has multiple dimensions, including income, education, occupation and employment. We use national composite scores that combine the different dimensions and in addition the main individual components as the association with air pollution and health may differ between dimensions. SES scores at regional scale were calculated by aggregating the raw variables to region level and then calculate the SES score.

The years have been picked as guideline reflecting years close to recruitment and during follow-up, before and after the 2008 economic crisis. The specific years have been selected because in at least the UK, Switzerland and Italy most data are from the censuses which were conducted in these years.

Table M6 Specification of area-level socio-economic status variables

Characteristic	Specification	Comment
Spatial scale	Neighborhood Region	Data on both scales required
Variables	Composite score	Combining different dimensions in one overall score, such as Carstairs index in the UK. Available in many countries, different in each country.
	Mean income	Mean household income
	Low household income	% of low household income (low nationally defined)
	Income support	% of working age population on income support
	Unemployment rate	% of working age population unemployed
	Education-low	% with low education to be defined nationally
	Education-high	% with high education to be defined nationally
	Ethnicity	% non-western immigrants*
Period	Year 1991 Year 2001 Year 2011	Time-varying variable. Close to recruitment and just before and after the economic crisis.
Spatial identification	Simple ID for both scales Coordinate of centroid of region**	ID to allow random effects Cox models (number from 1 to n). Coordinate to allow sub-national back-extrapolation of exposure. Coordinate not for neighborhood in anticipation of privacy issues

* Using the national definition which may differ in definition (e.g. how eastern Europeans are classified)

** Given the purpose, a simple geographic coordinate is sufficient

Table M7 Variable naming

Variable name	Explanation
Score_Neighbor_1995	Composite SES score at neighborhood level, year 1995
Score_ Region_1995	Composite SES score at regional scale, year 1995
Score_Neighbor_2005	Composite SES score at neighborhood level, year 2005
Score_ Region_2005	Composite SES score at regional scale, year 2005
Score_Neighbor_2010	Composite SES score at neighborhood level, year 2010
Score_ Region_2010	Composite SES score at regional scale, year 2010
Meanincome_Neighbor_1995	Mean household income at neighborhood level, year 1995
Meanincome_ Region_1995	Mean household income at regional scale, year 1995
Meanincome_Neighbor_2005	Mean household income at neighborhood level, year 2005
Meanincome_ Region_2005	Mean household income at regional scale, year 2005
Meanincome_Neighbor_2010	Mean household income at neighborhood level, year 2010
Meanincome_ Region_2010	Mean household income at regional scale, year 2010
Etc.	

Table M8 Geographical units in England and Wales

Level	Name	Average population	N
Local	Census Output Area (COA)	300	181,408
Neighborhood	Lower layer Super Output Area (LSOA)	1,500	34,753
	Ward	6,500	8,570
	Middle layer Super Output Area (MSOA)	8,000	7,201
Municipality	District/Local Authority	150,000	348
Region	Government Office Region	5,600,000	10

Background

After the brief rationale, we show how area-level SES was treated in recent major cohort studies.

Rationale for the selected indicators

We evaluated the possibility to obtain European data. We have concluded that obtaining neighborhood level SES data from European databases is not possible. We have not been able to trace adequate data on the neighborhood scale. Contacts with research groups working on socio-economic inequality research in Europe confirmed this. European data for very large spatial scale do exist (e.g. from Urban Audit or Eurostat), but we prefer to use the national data to be consistent with the neighborhood data. The conclusion is that the data were to be obtained by each cohort / country.

We have further reviewed administrative and individual-level cohorts outside the ELAPSE consortium and noted that these often North American studies have typically characterized SES more comprehensively than we have done previously in ESCAPE.

SES has multiple dimensions, including income, education, occupation and employment status. Various countries have developed indicators that combine the different dimensions. Examples include the Townsend and Carstairs index in the UK, the SES score in the Netherlands (Fischer et al. 2015), the Danish Deprivation Index at the parish level (Meijer et al. 2013), socio-economic position index at the census block level in Rome (Cesaroni et al. 2013), the Swiss neighborhood index of SEP for areas with 50 households (Panczak et al. 2012) and a composite deprivation score

at the Small Area Market Statistics level (SAMS, N~9000 in Sweden (Hamano et al. 2013). These scores probably better reflect the impact on health than each individual component. Where available, we used them.

Disadvantages of the composite scores are that they are different between countries and not available everywhere. Creating an index from the basic data is beyond the scope of ELAPSE. Furthermore, the correlation between air pollution and the various dimensions may be different. An illustration of different correlations of air pollution with income and education is found in a recent Canadian study: r=0.46 with education and 0.19 with low income (Pinault et al. 2016). We therefore use the individual SES components as well. The aim is to use multiple dimensions for each cohort, moving away from using (as examples) income at municipality level in Denmark and unemployment rate in Varese as the only indicators of SES.

Due to economic and political developments, changes in area-level SES may have occurred during follow-up. We therefore aimed at collecting SES data for multiple periods: at recruitment (~1991), during follow-up before the 2008 economic crisis (2001) and after the economic crisis (2011). North American cohorts have taken this approach as well. We realize this was not feasible everywhere. The specific years have been picked as these are years of the censuses (major source of information) in the UK, Switzerland and Italy.

When we collect information from different years it is likely that the boundaries of the geographical units change (they certainly did in the UK). If we assign these units to cohort participants based on their x-y-coordinate this should not be a big issue but we should be aware of the fact that the information (e.g. SES index) collected for one year might not necessarily represent the same neighborhood in another year.

Despite efforts to harmonize data, differences exist in the final data, both in terms of availability and definitions. An example of the problem of obtaining comparable data can be found in a recent paper comparing air pollution equity in the UK and the Netherlands (Fecht et al. 2015). Data on mean income were not available in the UK, data on education not at the neighborhood level in the Netherlands.

Because of the inevitable differences between countries, we analyze the data using categorical variables (low, medium, high), to be defined with national definitions of e.g. income. This step was made after data has been collected, consistent with the approach for individual level confounders.

The identification of the neighborhoods and regions by simple ID numbers allows random effect analyses. For the administrative national cohorts, random effects by neighborhood may not be possible for computational reasons. Inclusion of neighborhood SES is still possible, without random effect.

Area-level SES adjustment in selected other cohort studies (quotes from publications)

Canadian cohort (Crouse et al. 2015)

“”We adjusted our models for aboriginal ancestry, visible minority status, immigrant status, marital status, highest level of education, employment status, occupational classification, and quintiles of household income (see Supplemental Material of the Crouse paper, Table S1, for coding). We also calculated **time-varying contextual variables** from the closest census year (i.e., either 1991, 1996, 2001, or 2006) adjusted for regional variations across Canada (i.e., census division means subtracted from census tract means) describing the proportion of unemployed adults, the proportion of adults who had not completed high school, and the proportion of individuals in the lowest income quintile. Census tracts correspond roughly to the size of a neighborhood, and census divisions correspond roughly to the size of a city or county.””

ACS study (Turner et al. 2016)

“”Models were adjusted a priori for the following covariates assessed at enrollment as in previous work (6, 20, 25, 28–30): education; marital status; body mass index (BMI) and BMI squared; cigarette smoking status; cigarettes per day and cigarettes per day squared; years smoked and years smoked squared; started smoking at younger than 18 years of age; passive smoking (hours); vegetable, fruit, fiber, and fat intake; beer, wine, and liquor consumption; occupational exposures; an occupational dirtiness index; and **six sociodemographic ecological covariates at both the postal code and postal code minus county-level** mean derived from the 1990 U.S. Census (median household income and percentage of African American residents, Hispanic residents, adults with postsecondary education, unemployment, and poverty) (Table E5).””

Medicare cohort (Wang et al. 2017)

Census Variables and Behavioral Variables

“”We obtained the percentage of people below the poverty level, percentage of less educated people, median income, and median home value for each ZCTA (=zip code tabulation area) from the US Census Bureau 2000 Census Summary File 333 and the American Community Survey (AC S) 5-year estimates of 2009–2013.34 To account for the time-varying nature of these variables, we assigned the Census 2000 variables to observations from 2000 to 2006 and assigned the AC S estimates to observations from 2007 to 2013. The rural areas were defined as areas with a population density below the first tertile of the population density in the seven states (51 per square mile) according to the Census 2000 data. We also obtained age-adjusted yearly prevalence estimates of percentage of smokers and percentage of obesity from CDC Behavioral Risk Factor Surveillance System (BRFSS).””

NIH-AARP cohort (Thurston et al. 2016)

“”We also included two contextual characteristics of the participants’ residential census tracts found to modify the PM2.5–mortality HR estimates and have statistical significance in our analyses (data not shown): a) median census tract household income; and b) percent of census tract population with less than a high school education, based on the 2000 decennial census for the residence at study entry, as included in the cohort data set (NIH-AARP 2006).””

Table M9 Availability of indiviudal covariate variables in the pooled cohort .

Cohort	Age	Sex	BMI	Smoking status	Smoking intensity and duration	Marital status	Employment status	Occupational status	Education	Diet	Alcohol
CEANS-SDPP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CEANS-SIXTY	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CEANS-SALT	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CEANS-SNACK	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DCH	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
DNC-1993	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DNC-1999	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
E3N	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
EPIC-NL-MORGEN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
EPIC-NL-PROSPECT	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HNR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
KORA-S3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
KORA-S4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VHM&PP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
EPIC VARESE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes

Table M10 Availability of area-level socio-economic status and ethnicity covariate variables for 2001 in the pooled cohort

Cohort	Definition neighborhood	Definition region	Neighborhood					Region				
			Income	Unemployment rate	Income support	Education	Ethnicity	Income	Unemployment rate	Income support	Education	Ethnicity
CEANS-SDPP	Neighborhood	NA	Yes[1]	No	No	No	No	Yes	No	Yes	Yes	No
CEANS-SIXTY	Neighborhood	NA	Yes[1]	No	No	No	No	Yes	No	Yes	Yes	No
CEANS-SALT	Neighborhood	NA	Yes[1]	No	No	No	No	Yes	No	Yes	Yes	No
CEANS-SNACK	Neighborhood	NA	Yes[1]	No	No	No	No	Yes	No	Yes	Yes	No
DCH	Municipality	Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DNC-1993	Municipality	Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DNC-1999	Municipality	Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
E3N	IRIS/commune[2]	NA	Yes	Yes	No	Yes	Yes	No	No	No	No	No
EPIC-NL-MORGEN	Neighborhood	COROP area	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes
EPIC-NL-PROSPECT	Neighborhood	COROP area	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes
HNR	City district	Municipality	Yes	Yes	No	No	No	No	No	No	No	No
KORA-S3	Boroughs of Augsburg + municipality	NA	Yes	No	No	No	No	Yes	Yes	No	No	Yes
KORA-S4	Boroughs of Augsburg + municipality	NA	Yes	No	No	No	No	Yes	Yes	No	No	Yes
VHM&PP	Municipality	NA	Yes	Yes	No	Yes	Yes	No	No	No	No	No
EPIC VARESE		NA	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes

NA= not applicable because of size of study area

[1] estimated from neighborhood income in 2011 and the ratio of regional level income in 2001 and 2011

[2] IRIS is a small administrative unit of about 2000 subjects in cities; for smaller towns commune = municipality was used

Appendix M6: Harmonization of different definitions of diet and alcohol consumption

In ELAPSE we pooled data from the participating individual cohorts. Differences between covariate definitions were present in several variables as, *e.g.*, dietary variables, which have been defined with actual quantitative consumption (*e.g., grams/day*) or with frequency of consumption (*e.g., daily, weekly, seldom*). Different variable definitions were handled by harmonization between cohorts and transformation of the variable where needed.

For most dietary factors we analyzed the impact of adding categorical, continuous and or a harmonized variable defined a low/medium/ high intake combined in three separate analyses. The analysis was thus performed in three different sets of cohorts. In each case, models with and without the dietary variable were compared between analyses using the same number of observations. All data was analyzed as low / medium / high to avoid outliers and relax the linearity assumption. Below a detailed analysis is presented,

Fruit

Five of the eight cohorts in the pooled dataset report fruit based on consumption frequency (**fruit_cat**, Figure M13). Three cohorts report fruit consumption in quantity (**fruit**, Table M11).

Figure M13 Distribution of available **fruit (categorical)** variables per cohort (1='Daily', 2='Weekly', 3='Seldom', 4='Never')

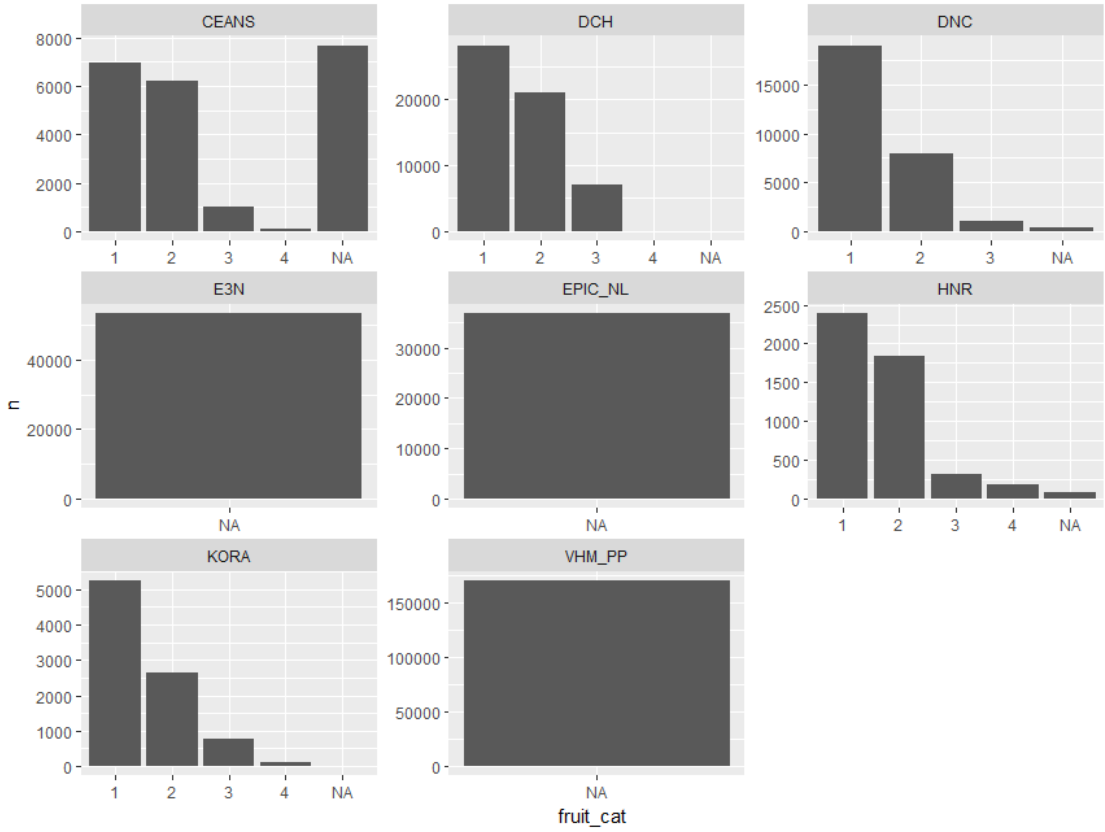


Table M11 Descriptives of available **fruit (continuous)** variables per cohort

##	study	n	mean_fruits_g_day
## 1	DCH	56254	181
## 2	E3N	53517	257
## 3	EPIC_NL	36740	261

In order to harmonize the fruit variables, we created a new **fruit_elapse** variable with *low-medium-high* categories, harmonizing the frequency-based **fruit_cat** and quantity-based **fruit** variables.

For cohorts where only **fruit_cat** variable is available, recoding is fairly straightforward: daily is high; weekly is medium and seldom/never is low. We interpreted 'weekly' as 'less than daily', then *medium* seems right. We combined *seldom* and *never* categories into *low*.

For cohorts where quantity-based **fruit** data is available, we decided to calculate tertiles. These showed substantial differences between the cohorts (Table M12). To make use of these differences, we then calculated the tertiles after pooling data from the respective cohorts. These tertiles were used as *low-medium-high* cut-off values for the harmonized **fruit_elapse** variable (Figure M14).

The DCH cohort is (currently) the only one with both fruit consumption definitions available. However, the comparison of **fruit_cat** and **fruit_elapse** (based on pooled **fruit** tertiles) showed substantially different pictures (Figures M13 and M14). Due to lack of available data, we are unable to make similar comparison for other pooled cohorts. Because of the poor comparison within DCH of the two fruit definitions, we performed three separate analyses to test the sensitivity of the air pollution effect estimates to inclusion of fruit: fruit as a continuous variable, fruit as a categorical variables and fruit as a combined variable. These analyses were performed in three different sets of cohorts (all with and without fruit to assess the effect of fruit on air pollution effect estimates).

We thus analyzed:

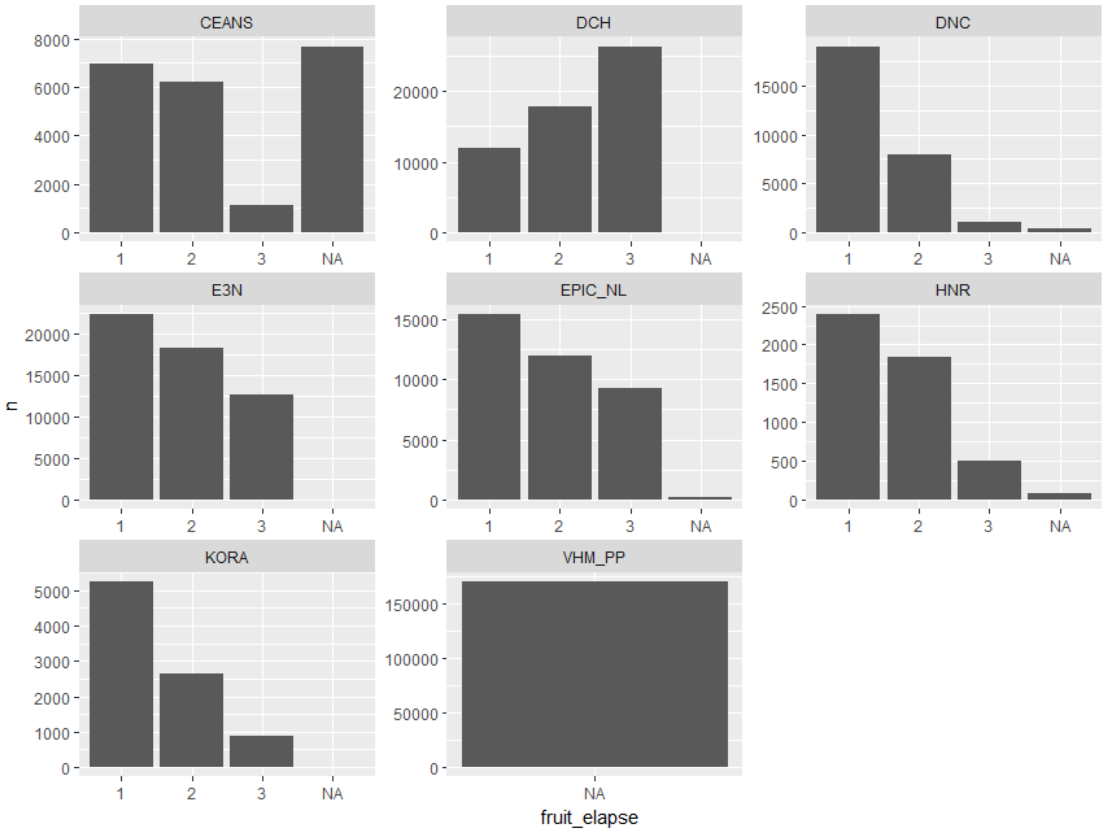
1. **fruit_cat**, (n=108,527)
2. **fruit_elapse (**n=191,518)
3. **Fruit (**n=138,991)

Quantity-based **fruit** was the most likely candidate to continue with in further analyses, given the high number of observations and its best accuracy in describing the actual intake. The continuous data was analyzed as low / medium / high to avoid outliers and relax the linearity assumption.

Table M12 Harmonization of fruit variables (**bold** is applied)

Variable	Definition	Low	Medium	High	Comment
Fruit	**Categorical**	**seldom/never**	**weekly**	**daily**	***Weekly* interpreted as 'less than daily'**
	POOLED	**< 137**	**137 - 264**	**> 264**	**Based on tertiles**
	DCH	< 97	97 - 197	> 197	Based on tertiles
	E3N	< 164	164 - 300	> 300	Based on tertiles
	EPIC-NL	< 165	165 - 303	> 303	Based on tertiles

Figure M14 Distribution of the harmonized **fruit_elapse** variable across cohorts (1='High', 2='Medium', 3='Low'). Note: for cohorts with only categorical fruit, distribution in this figure is identical to figure M13.



Total vegetable

All cohorts that have vegetable consumption data available, have continuous data except HNR. Therefore, there is little need to harmonize categorical and quantitative consumption data. We

analyzed only the continuous data (step 3 listed in the fruit analyses). The data was analyzed as low / medium / high to avoid outliers and relax the linearity assumption, using the same procedure as outlined for fruit.

Figure M15 Distribution of available **vegetable (total, categorical)** variables per cohort (1='Daily', 2='Weekly', 3='Seldom', 4='Never')

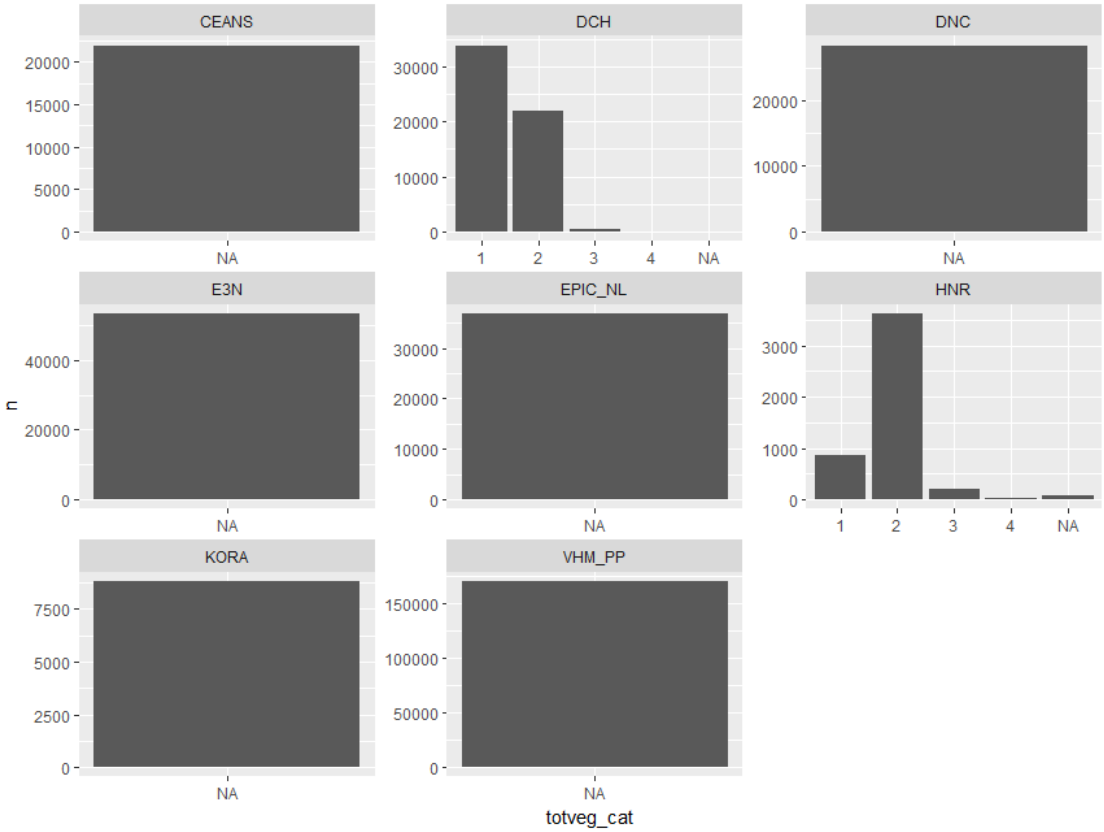


Table M13 Descriptives of available **vegetable (total, continuous)** variables per cohort

##	study	n	mean_totveg_g_day
## 1	DCH	56254	174
## 2	E3N	53517	274
## 3	EPIC_NL	36740	138

Table M14 Harmonization of total vegetable variables (**bold** is applied)

Variable	Definition	Low	Medium	High	Comment
Vegetable (total)	**Categorical**	**seldom/never**	**weekly**	**daily**	
	POOLED	**< 134**	**134 - 222**	**> 222**	**Based on tertiles**
	DCH	< 121	121 - 197	> 197	Based on tertiles
	E3N	< 204	204 – 307	> 307	Based on tertiles
	EPIC-NL	< 111	111 - 152	> 152	Based on tertiles

Raw vegetable

Two cohorts have continuous and three cohorts have categorical raw vegetable consumption data available. We analyzed categorical, continuous and combined, as with fruit. The data was analyzed as low / medium / high to avoid outliers and relax the linearity assumption, using the same procedure as outlined for fruit.

Figure M16 Distribution of available **vegetable (raw, categorical)** variables per cohort (1='Daily', 2='Weekly', 3='Seldom', 4='Never')

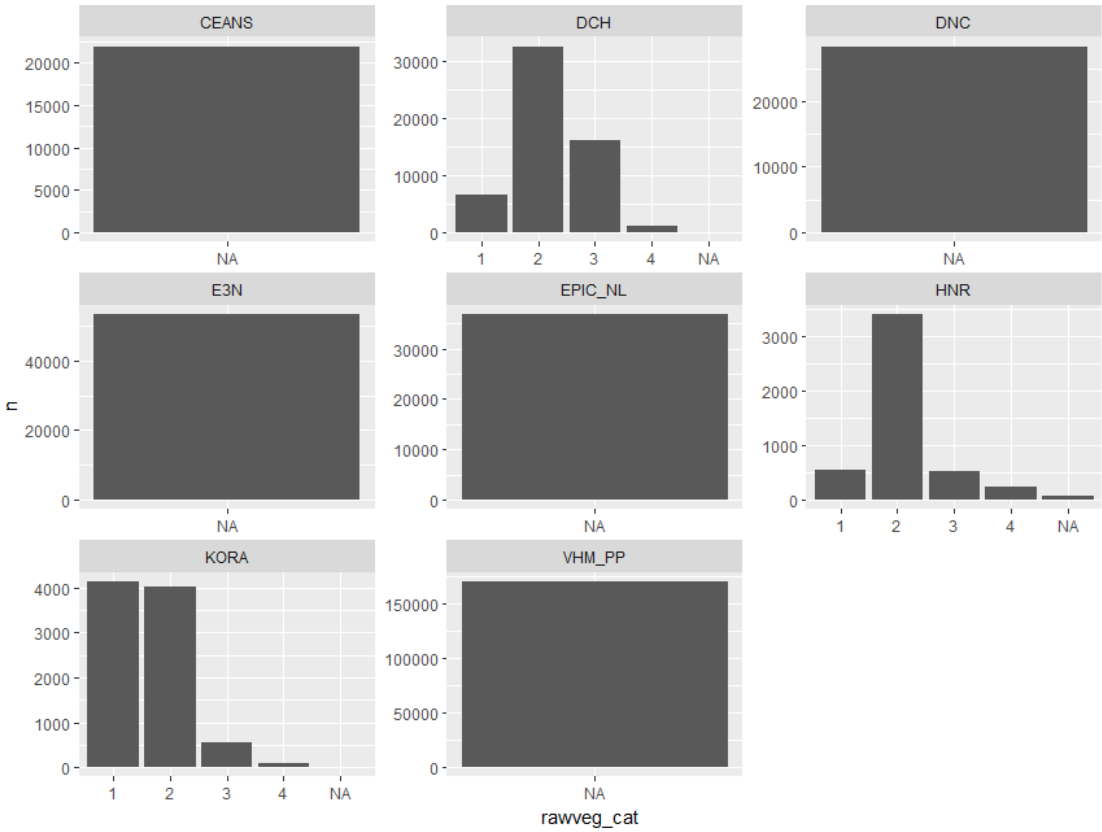


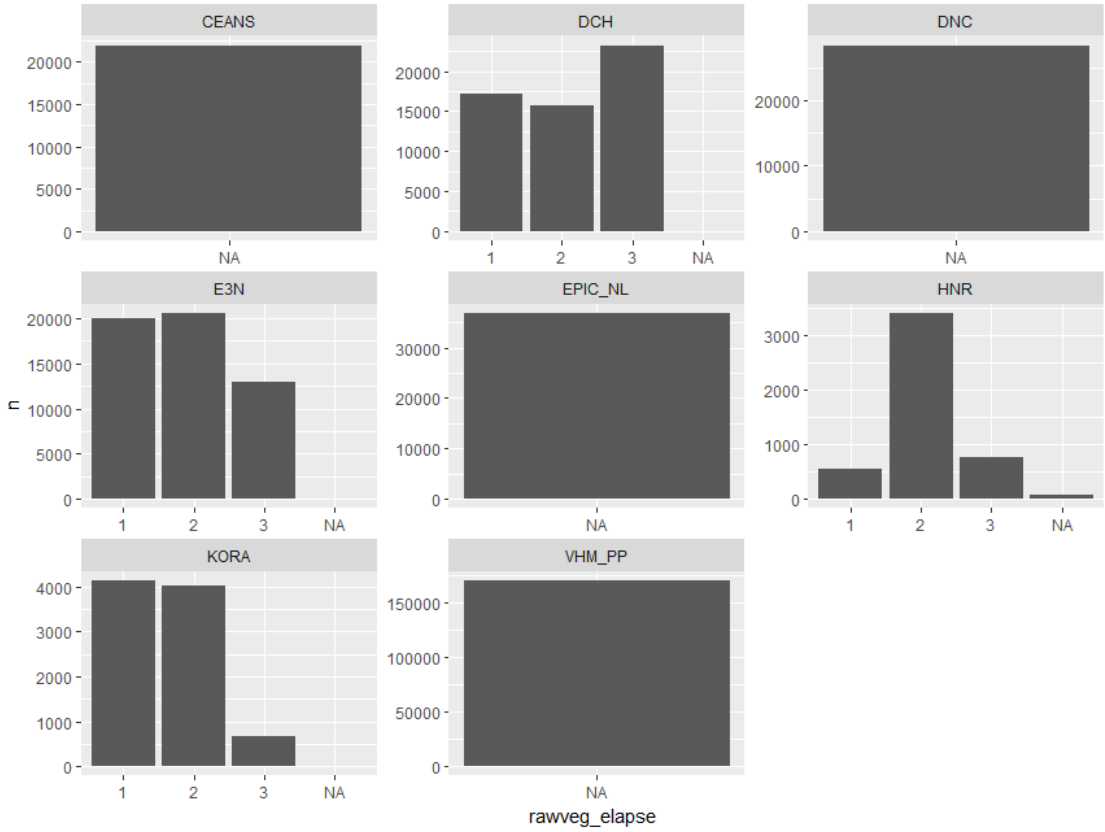
Table M15 Descriptives of available **vegetable (raw, continuous)** variables per cohort

##	study	n	mean_rawveg_g_day
## 1	DCH	56260	64
## 2	E3N	53517	69

Table M16 Harmonization of raw vegetable variables (applied in **bold**)

Variable	Definition	Low	Medium	High	Comment
Vegetable (raw)	**Categorical**	**seldom/never**	**weekly**	**daily**	
	POOLED	**< 32**	**32 - 78**	**> 78**	**Based on tertiles**
	DCH	< 22	22 - 71	> 71	Based on tertiles
	E3N	< 42	42 - 84	> 84	Based on tertiles

Figure M17 Distribution of the harmonized **rawveg_elapse** variable across cohorts (1='High', 2='Medium', 3='Low')



Meat

Three cohorts have continuous and three cohorts have categorical meat consumption data available. We analyzed categorical, continuous and combined, as with fruit. The data were analyzed as low / medium / high to avoid outliers and relax the linearity assumption, using the same procedure as outlined for fruit.

Table M17 Descriptives of available **meat (continuous)** variables per cohort

##	study	n	mean_meat_g_day
## 1	DCH	56254	182
## 2	E3N	53517	82
## 3	EPIC_NL	36740	89

Figure M18 Distribution of available **meat (categorical)** variables per cohort (1='Daily', 2='Weekly', 3='Seldom', 4='Never')

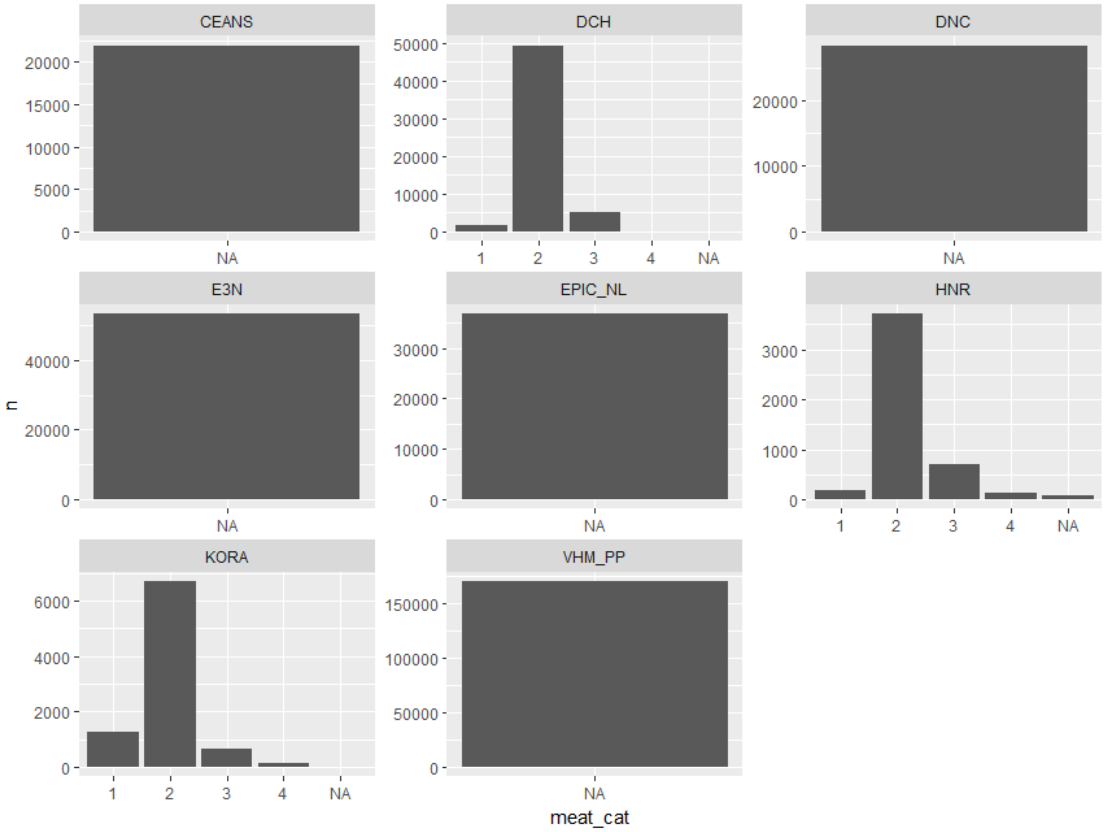
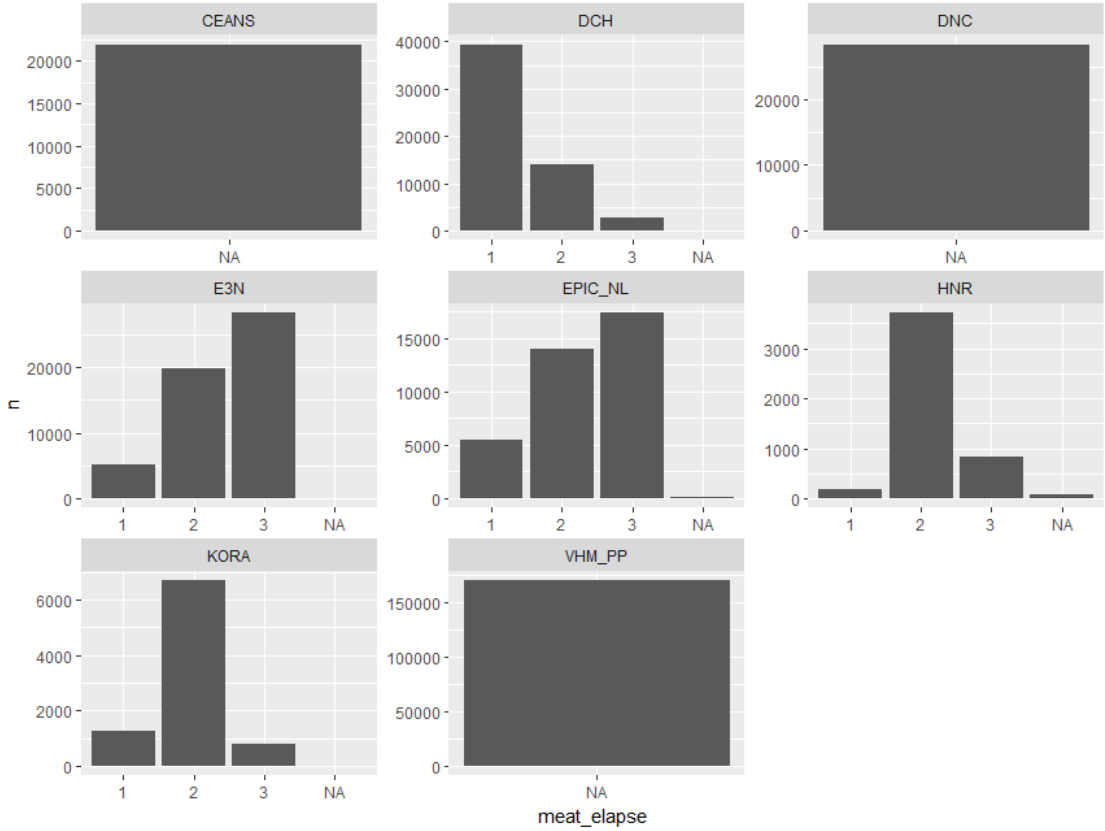


Table M18 Harmonization of meat variables (**bold** applied)

Variable	Definition	Low	Medium	High	Comment
Meat	**Categorical**	**seldom/never**	**weekly**	**daily**	
	POOLED	**< 83**	**83 - 138**	**> 138**	**Based on tertiles**
	DCH	< 143	143 - 200	> 200	Based on tertiles
	E3N	< 61	61 - 97	> 97	Based on tertiles
	EPIC-NL	< 63	63 - 106	> 106	Based on tertiles

Figure M19 Distribution of the harmonized **meat_elapse** variable across cohorts (1='High', 2='Medium', 3='Low')



Alcohol

All cohorts that have alcohol consumption data available, have continuous data except CEANS. Therefore there is little need to harmonize categorical and quantitative consumption data. We l analyzed only the continuous data (step 3 listed in the fruit analyses). The data was analyzed as low / medium / high to avoid outliers and relax the linearity assumption, using the same procedure as outlined for fruit.

Table M19 Descriptives of available **alcohol (continuous)** variables per cohort

##	study	n	mean_alc
## 1	DCH	56259	21
## 2	DNC	27579	16
## 3	E3N	53517	11
## 4	EPIC_NL	36740	11
## 5	HNR	4686	8
## 6	KORA	8804	16

Figure M20 Distribution of available **alcohol (categorical)** variables per cohort (1='Daily', 2='Weekly', 3='Seldom', 4='Never')

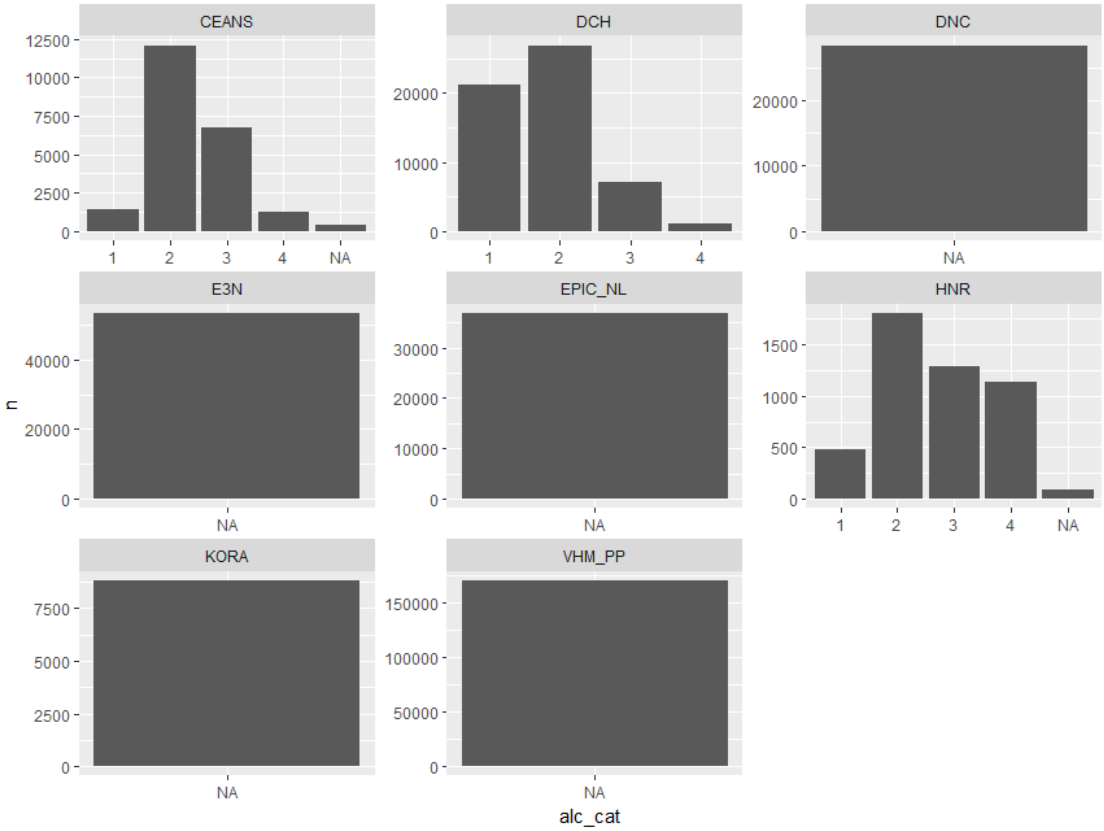


Table M20 Harmonization of alcohol variables (**bold** applied)

Variable	Definition	Low	Medium	High	Comment
Alcohol	**Categorical**	**seldom/never**	**weekly**	**daily**	
	POOLED	**< 4**	**4 - 15**	**> 15**	**Based on tertiles**
	DCH	< 8	8 - 20	> 20	Based on tertiles
	DNC	< 5	5 - 19	> 19	Based on tertiles
	E3N	< 3	3 - 12	> 12	Based on tertiles
	EPIC-NL	< 1	1 - 11	> 11	Based on tertiles
	HNR	< 0	0 - 5	> 5	Based on tertiles
	KORA	< 2	2 - 17	> 17	Based on tertiles

Appendix M7: Missing covariate value assessment

Multiple imputation (MI) is an attractive and effective approach for statistical analysis of incomplete data. The main idea is to create multiple data sets that reflect the potential values of the missing data. More precisely, random draws are made from the posterior distribution of the missing values given the observed data, usually under the missing at random (MAR) assumption. Estimates are combined across imputed data sets using Rubin's rules (Rubin 1987). Although MI techniques have been proposed for imputing a covariate that may be complete missing from one study from the rest of the studies that are pooled under a multi-cohort approach, we decided against this considering the differences in the underlying populations between participating cohorts and instead prompted for the extensive sensitivity analysis on the choice of covariates in Model 3. We nevertheless tested robustness of the effect estimates for the association with total mortality to the missing data in the covariates that were available across cohorts and included in the main Model 3. When missing values occur in multiple variables, and in particular when these are a mixture of continuous and discrete variables, the method of multiple imputation by chained equations (MICE) is particularly attractive (van Buuren et al. 1999). This involves specifying a separate imputation model for each incomplete variable given all the other variables and repeatedly imputing the variables in an iterated sequence. As with MI in general, it is crucial that the imputation model is consistent (or congenial) with the model of interest, which will subsequently be fitted to the imputed data sets. Hence, MICE was applied for each cohort with availability of covariate data but with missing values to produce 5 complete datasets per cohort. In this way the imputation of the missing values was based on the cohort-specific data and did not use information from the rest contributing cohorts in the pooled data set. Consequently, the cohort-specific corresponding complete datasets were pooled and Model 3 was applied for each of the 5 complete pooled datasets. The effect estimates from these models were pooled using the Rubin's rules. We applied MICE by filling in missing data for all covariates in main Model 3. We used the R library *mice*.

Appendix M8: Incorporation of cohort, neighborhood and large area effects in analysis of the pooled and administrative cohorts

In the process of finalizing the methodology for the analysis of the pooled **cohort** data we assessed several options to account for the level of grouping in the data corresponding to the underlying participating cohorts in the dataset. Since the association under investigation in the ELAPSE project is the health effects of long-term residential exposure to specific air pollutants, a second level is inherent in the data that corresponds to the neighborhood (small area) level of the participants' residence.

We assessed the sensitivity of our health effect estimates depending on the control for the underlying cohort with five approaches (Debray et al. 2015): 1) not accounting for the cohort, 2) dummies for cohorts, 3) strata per cohort, 4) a frailty term for cohort identification and 5) a random intercept per cohort under a mixed Cox modeling approach. For the composite cohorts (CEANS and EPIC-NL) indicators of or control for the individual cohorts were applied. We assessed the proportional hazards (PH) assumption in terms of the log-log plots and compared models in terms of the Akaike Information Criteria (AIC). We further accounted for the second level of clustering at the neighborhood level by either a 1) random intercept per neighborhood or 2) applying variance correction, recognizing that these allow for different interpretations following either a conditional or marginal approach correspondingly.

Not accounting for the original study in the model assumes that baseline hazards and hazard ratios (HRs) are the same for each cohort hence this be labeled as a naive approach, although it has the advantage of exploiting the whole exposure contrast across cohorts. On terms of estimation it is an unconditional model applying a maximum partial likelihood (ML), generally using Newton–Raphson (NR). When we include dummy variables to characterize the cohorts the model follows a fixed effects approach. It is a conditional model where cohorts act proportionally on the baseline risk, while estimation again follows from NR. Using a stratified model for the contributing cohort allows baseline hazards to vary by cohort and relaxes the proportional hazards (PH) assumption between cohorts. Under stratification the baseline hazards play no role in the estimation (again by the ML approach) and the model does not estimate between cohorts variance. The lack of structure makes this choice the most general of the conditional models, although being conditional there are no between-cohort comparisons and all information comes from the within-cohort comparisons, although this is partly overcome by large sample sizes. The remaining two approaches (frailty and mixed Cox; (O'Quigley and Stare 2002) follow a random effects approach with random intercepts for participating studies. They estimate between cohorts' heterogeneity but assume different underlying distributions for the baseline hazards. Hence the frailty model treats cohort effects as a sample from a (usually) gamma distribution, where the frailties represent unmeasured factors affecting cohort-specific baseline risks and are assumed to act multiplicatively on the average baseline risk. Estimation is typically through penalized partial likelihood algorithm. A mixed Cox model with random intercept for the underlying cohorts basically is the same with the frailty but assumes a Gaussian distribution of underlying hazards.

There is some discussion in the literature for the optimal choice. (Glidden and Vittinghoff 2004) support the use of random effects Cox models in multi-center studies, while others indicate that for large sample sizes the two approaches provide identical estimates (O'Quigley and Stare 2002). (Giganti et al. 2015) similarly reviewed seven approaches (including the ones in the present report) to account for heterogeneity in HIV treatment cohorts and concluded that hazard ratio estimates varied slightly between approaches, and differences were not clinically meaningful. Nevertheless, most previous publications focus on discrete exposures. Basically the two approaches have different assumptions and advantages, as the random effects models (frailty or mixed Cox) estimate between cohorts' heterogeneity and assume a structure for baseline hazards, while it further allows to incorporate random slopes per cohort under a mixed Cox approach. Heterogeneity in exposure effect can also be investigated by specifying a random effects distribution.

Similarly, in the analysis of the **administrative** cohorts we initially considered two levels of spatial correlation: one in large administrative regions across the country and one in the smaller administrative unit at the neighborhood level. We assessed the sensitivity of our health effect estimates with the following approaches: 1) not accounting for the big administrative regions, 2) dummies for regions, 3) strata per region, 4) a frailty term for region identification and 5) dummies for regions with variance correction of the effect estimates at the neighborhood level and 6) only variance correction of the effect estimates at the neighborhood level. Especially for the Rome administrative cohort only the model not accounting for area level adjustment and option 6 were assessed, as the cohort was restricted to the large administrative unit of Rome and the definition of region, under the rationale described above, was not relevant.

All analyses were done in R version 3.4.0 with packages: *survival, coxme, Matrix, foreach, glmnet, multcomp, survey, splines, Hmisc, mfp, VIM, ggplot2, frailtySurv, survsim, eha, stamod*. R provides several options for frailty and Cox mixed models. Specifically, the function *frailty* of *survival* package allows for either: 1) a Gamma distribution and a choice between estimation-maximization algorithm (that is the default) or restricted maximum –likelihood (REML); or 2) a Gaussian frailty under REML estimation. Library *coxme* applies a Gaussian Cox mixed effects models under EM algorithm.

Table M21 Overview of area-level SES confounder data including year and spatial scale

Cohort	Definition neighborhood (number in cohort)	Definition region (number in cohort)	Area-level SES data including scale	Year of data
Dutch	Neighborhood ("Wijk" in Dutch, on average 6000 subjects (n~2700)	Main Public Health authority COROP (n=40)	SES score 4-position postal code (~neighborhood). Mean income, %low income, unemployment rate, ethnicity at neighborhood (wijk) and COROP scale.	2006 and 2012
English	N/A	Strategic Health Authority	Index of Multiple Deprivation Lower Super Output Area	2010
Rome	Neighborhood (n =94)	Not applicable	SES score census block and mean income, education %low and % high) per neighborhood	2001
Danish	Municipality and possibly parish	Region (n=5), possibly province	Parish and province avg. income unemployment rate, benefits, education	2001
Norwegian	Large neighborhood (N=1543)	County (N=19)	Income, low income, low education, unemployed, renting, non-western ethnicity, living in dwelling, single parent	2001 and 2011
Swiss	Postal code (N=3175)	Canton (n=26)	Neighborhood (postcode) and canton SES score 2001, plus education and employment	2001 and 2011
Belgian	Sections (n=6344), being a level between municipalities LAU 2 (n=589) and census tracts (n=19781)	Arrondissements (n=43) NUTS 3 level	mean income, unemployment rate, low education level rate, and ethnicity rate for both neighborhood and region	2001 and 2011

Table M22 Administrative cohorts and available survey data used for indirect adjustment approach

Cohort	Survey	Smoking	BMI	Alcohol use	Physical activity
Dutch	National Health Monitor of 2012, ~400,000 subjects, elderly overrepresented by design	Status and amount	YES	Status and amount	YES
English	Not applicable: individual data available in the full cohort	Status	YES	NO	NO
Rome	SIDRIA: ~7,000 Rome young parents ~1,000 subjects, population based, 65+	Status and amount	NA	NA	NA
Danish	Danish National Survey, 2010, ~160,000 subjects	YES	YES	YES	YES
Norwegian	CONOR, cohort with participants from all regions in Norway, approx. 117,000	Status, amount and duration	YES	Status	YES
Swiss	Swiss Health Survey of 1992 (n~15,000); 2002 survey not available for researchers [a]	Status and amount	YES	Status and amount	YES
Belgian	Population representative Health Interview Survey 2001 (n~10,000)	Status in 3 categories (current, former and never)	YES in 4 categories defined by WHO	NA	NA

NA = not applicable

[a] No geocodes in the survey, so only option to use this data is if it's linked directly to SNC

Table M23 Overview of spatial scales in the analysis of administrative cohorts

Cohort	Definition neighborhood for area-level SES (number in cohort)	Definition region for area-level SES (number in cohort)	Definition of area for large area adjustment (frailty, strata)
Dutch	Neighborhood ("Wijk" in Dutch, on average 6000 subjects (n~2700); Postal code for score (N=4019)	Main Public Health authority COROP (n=40)	Province (n=12)
English	Lower Super Output Area	N/A	Strategic Health Authority
Rome	Neighborhood (n =94)	Not needed	Not needed
Danish	Municipality (n = 98) and parish (n = 2,200)		Region (N=5)
Norwegian	Large neighborhood (N=1543).	County (N=19)	
Swiss	Postal code (N=3175)	Canton (n=26)	Region (n=7)
Belgian	Sections (n=6344)	Arrondisements NUTS 3 (n=43)	Region (n=3)

Appendix M9:Measurement error correction using regression calibration

We applied regression calibration (Carroll et al. 1995; Keogh and White 2013) to assess measurement error. We estimated the attenuation factor from a simple linear regression of the measured concentrations at the ESCAPE monitoring sites vs. the estimated concentrations from the ELAPSE model for PM2.5 and NO_2. ESCAPE monitoring data were not used in model development for PM2.5 and NO_2. For BC and O_3 we needed another procedure as BC models were developed based upon ESCAPE data and O_3 was not measured in ESCAPE. For BC and O_3 we regressed ELAPSE model predictions from the hold-out validation procedure at ESCAPE and Airbase sites respectively on measured concentrations. Appendix M9 provides further details. As the measurement error model was based on fixed site ambient concentrations, personal covariates such as smoking in our main model, could not incorporated in the measurement error model. We therefore applied regression calibration in the minimally adjusted model 1, as it is important that the health model and the measurement error model contain the same covariates. We used model 1, because a model without any adjustments (no age, sex and cohort), is not interpretable. Specifically, the approach proceeded in three steps:

Step 1. We run a simple Cox model using the estimated exposure as the only independent variable (age as time axis and strata for sub-cohort and sex are specified in this model). We may define this as our naive model

Step 2. We run a linear regression model with the concentrations measured at the ESCAPE/AIRBASE monitoring sites as dependent variable and the estimated exposure with the ELAPSE model as independent variable and extracted the regression slope and variance-covariance matrix from this model.

Step 3. The corrected HR and its variance was estimated using the attenuation factor (lambda) from the RC model of step 2. The formulas below document the calculations:

bcorr = bnaive/lambda

varcorr = (varnaive/(lambdaˆ2))+((bnaive/(lambdaˆ2))ˆ2)*lambdavar

with:

bcorr	is the corrected slope of the mortality -pollution association
bnaive	is slope of the mortality pollution association from step 1, not corrected for measurement error
lambda	is the attenuation factor, the regression slope from step 2 model
lambdavar	is the attenuation factor variance-covariance matrix from step 2 model
varcorr	the variance of the corrected slope and
varnaive	the variance of the uncorrected slope

References

Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. 2013. Development of no2 and nox land use regression models for estimating air pollution exposure in 36 study areas in europe - the escape project. Atmospheric Environment 72:10-23.

Brandt J, Silver JD, Frohn LM, Geels C, Gross A, Hansen AB, et al. 2012. An integrated model study for europe and north america using the danish eulerian hemispheric model with focus on intercontinental transport of air pollution. Atmos Environ 53:156-176.

Carey IM, Atkinson RW, Kent AJ, van Staa T, Cook DG, Anderson HR. 2013. Mortality associations with long-term exposure to outdoor air pollution in a national english cohort. American journal of respiratory and critical care medicine 187:1226-1233.

Carroll RJ, Ruppert D, Stefanski LA. 1995. Measurement error in non-linear models. London:Chapman and Hall.

Cesaroni G, Porta D, Badaloni C, Stafoggia M, Eeftens M, Meliefste K, et al. 2012. Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. Environmental health : a global access science source 11:48.

Cesaroni G, Badaloni C, Gariazzo C, Stafoggia M, Sozzi R, Davoli M, et al. 2013. Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in rome. Environmental health perspectives 121:324-331.

Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV, et al. 2015. Ambient pm2.5, o(3), and no(2) exposures and associations with mortality over 16 years of follow-up in the canadian census health and environment cohort (canchec). Environmental health perspectives 123:1180-1186.

de Hoogh K, Chen J, Gulliver J, Hoffmann B, Hertel O, Ketzel M, et al. 2018. Spatial pm2.5, no2, o3 and bc models for western europe - evaluation of spatiotemporal stability. Environment international 120:81-92.

Debray TP, Moons KG, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RH, et al. 2015. Get real in individual participant data (ipd) meta-analysis: A review of the methodology. Research synthesis methods 6:293-309.

Eeftens M, Tsai MY, Ampe C, Anwander B, Beelen R, Bellander T, et al. 2012. Spatial variation of pm2.5, pm10, pm2.5 absorbance and pmcoarse concentrations between and within 20 european study areas and the relationship with no2 - results of the escape project. Atmospheric Environment 62:303-317.

Fecht D, Fischer P, Fortunato L, Hoek G, de Hoogh K, Marra M, et al. 2015. Associations between air pollution and socioeconomic characteristics, ethnicity and age profile of neighbourhoods in england and the netherlands. Environmental pollution 198:201-210.

Fischer P, Marra M, Ameling C, Hoek G, Beelen R, De Hoogh K, et al. 2015. Air pollution and mortality - the dutch environmental longitudinal study (duels). Environmental health perspectives.

Giganti MJ, Luz PM, Caro-Vega Y, Cesar C, Padgett D, Koenig S, et al. 2015. A comparison of seven cox regression-based models to account for heterogeneity across multiple hiv treatment cohorts in latin america and the caribbean. AIDS research and human retroviruses 31:496-503.

Glidden DV, Vittinghoff E. 2004. Modelling clustered survival data from multicentre clinical trials. Statistics in medicine 23:369-388.

Hamano T, Kawakami N, Li X, Sundquist K. 2013. Neighbourhood environment and stroke: A follow-up study in sweden. PloS one 8:e56680.

Hertel O, Jensen SS, Andersen HV, Palmgren F, Wåhlin P, Skov H, et al. 2001. Human exposure to traffic pollution. Experience from danish studies. Pure and Applied Chemistry 73:137-145.

Hooyberghs J, Mensink C, Dumont G, Fierens F. 2006. Spatial interpolation of ambient ozone concentrations from sparse monitoring points in belgium. Journal of environmental monitoring : JEM 8:1129-1135.

Keogh RH, White IR. 2013. Using full-cohort data in nested case-control and case-cohort studies by multiple imputation. Statistics in medicine 32:4021-4043.

Ketzel M, Berkowicz R, Hvidberg M, Jensen SS, Raaschou-Nielsen O. 2011. Evaluation of airgis - a gis-based air pollution and human exposure modelling system. International Journal for Environment and Pollution 47:226-238.

Ketzel M., Wåhlin P, Berkowicz R, Palmgren F. 2003. Particle and trace gas emission factors under urban driving conditions in copenhagen based on street and roof level observations. Atmos Environ 37:2735-2749.
Lefebrve W, Vrankxs S. 2013. Validation of the ifdm-model for use in urban applications.VITO.
Meijer M, Engholm G, Grittner U, Bloomfield K. 2013. A socioeconomic deprivation index for small areas in denmark. Scandinavian journal of public health 41:560-569.
O'Quigley J, Stare J. 2002. Proportional hazards models with frailties and random effects. Statistics in medicine 21:3219-3233.
Panczak R, Galobardes B, Voorpostel M, Spoerri A, Zwahlen M, Egger M, et al. 2012. A swiss neighbourhood index of socioeconomic position: Development and association with mortality. Journal of epidemiology and community health 66:1129-1136.
Pinault L, Tjepkema M, Crouse DL, Weichenthal S, van Donkelaar A, Martin RV, et al. 2016. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the canadian community health survey cohort. Environmental health : a global access science source 15:18.
Rubin DB. 1987. Multiple imputation for nonresponse in surveys. New York:Wiley and sons.
Thurston GD, Ahn J, Cromar KR, Shao Y, Reynolds HR, Jerrett M, et al. 2016. Ambient particulate matter air pollution exposure and mortality in the nih-aarp diet and health cohort. Environmental health perspectives 124:484-490.
Turner MC, Jerrett M, Pope CA, Krewski D, Gapstur SM, Diver WR, et al. 2016. Long-term ozone exposure and mortality in a large prospective study. American journal of respiratory and critical care medicine 193:1134-1142.
van Buuren S, Boshuizen HC, Knook DL. 1999. Multiple imputation of missing blood pressure covariates in survival analysis. Statistics in medicine 18:681-694.
Velders GJM. 2016. Grootschalige concentratie- en depositiekaarten nederland, rapportage 2016. RIVM National Institute for Public Health and the Environment report 2016-0068, 2016. Bilthoven:RIVM.
Wang Y, Shi L, Lee M, Liu P, Di Q, Zanobetti A, et al. 2017. Long-term exposure to pm2.5 and mortality among older adults in the southeastern us. Epidemiology 28:207-214.
Wolf K, Hoffmann B, Andersen Z, Atkinson R, Bauwelinck M, Bellander T, et al. In press. Long-term exposure to low-level ambient air pollution and incidence of stroke and acute coronary heart disease – Results from six European cohorts within the ELAPSE Project. Lancet Planetary Health.