Agreement of Land Use Regression models with personal exposure measurements of particulate matter and nitrogen oxides air pollution

Denise Montagne[1)], Gerard Hoek[1,*)], Mark Nieuwenhuijsen[2)], Timo Lanki[3)], Arto Pennanen[3)], Meritxell Portella[2)], Kees Meliefste[1)], Marloes Eeftens[1)], Tarja Yli-Tuomi[3)], Marta Cirach Pradas[2)], Bert Brunekreef[1,4)]

1) Institute for Risk Assessment Sciences (IRAS), University of Utrecht, Utrecht, the Netherlands
2) Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
3) Department of Environmental Health, National Institute for Health and Welfare (THL), Kuopio, Finland
4) Julius Center for Health Sciences and Primary Care, University Medical Center, University of Utrecht, Utrecht, the Netherlands

Contents

Supporting Information 1: LUR models

Table S1 ESCAPE Land Use Regression models for PM2.5 and Soot (Eeftens et al. 2012)

Study area	**LUR model[1]**	**R^2 of model**	**R^2 validation**	**RMSE[2] (validation) (μg/m³)**	**Number of sites[3]**	**Measured concentration (μg/m³)[4]**
			PM2.5			
Helsinki/Turku, Finland	9.25 – 6.75E-6*NATURAL_500[4] + 6.34E-7*TRAFMAJORLOAD_50	67%	53%	1.0	20	8.6 [5.3 – 12.3]
Netherlands/Belgium	9.46 + 0.42*REGIONALESTIMATE + 0.01*MAJORROADLENGTH_50 + 2.28E-9*TRAFMAJORLOAD_1000	67%	61%	1.2	40	17.7 [12.7 – 21.5]
Barcelona, Spain	16.21 – 4.08E-6*GREEN_1000 + 2.04E-7*TRAFLOAD_100 + 6.82E-3*INTINVDIST2	83%	71%	2.1	20	16.3 [8.4 – 24.4]
			Soot			
Helsinki/Turku, Finland	1.15 + 2.09E-7*TRAFLOAD_50 – 1.15E-6*NATURAL_500[4]	65%	47%	0.3	20	1.1 [0.6 – 2.3]
Netherlands/Belgium	0.07 + 2.95E-9*TRAFLOAD_500 + 2.93E-3*MAJORROADLENGTH_50 + 0.85*REGIONALESTIMATE + 7.90E-9*HLDRES_5000 + 1.72E-6*HEAVYTRAFLOAD_50	92%	89%	0.2	40	1.7 [0.9 – 3.0]
Barcelona, Spain	1.01 + 7.46E-6*HDRES_300 + 2.66E-3*INTINVDIST2 + 1.11E-7**TRAFLOAD_50	86%	80%	0.4	20	2.7 [0.9 – 4.9]

1 Most variables are buffers with _xxx indicating the size of the buffer in m (e.g. HHOLD_500 is the number of households in a 500m buffer). TRAFLOAD is traffic intensity * length of road in a buffer. HDRES and LDRES are high and low density residential land use. INTINVDIST is product of traffic intensity and inverse distance to the nearest road. SQRALT is square root of altitude. Major road as road classes 0, 1 and 2 (motorways, main roads of major importance and other main roads) from the central road network or roads with more than 5,000 vehicles/day based upon local networks with linked traffic intensity. INTINVDIST is traffic intensity multiplied by inverse distance (squared).
2 RMSE is the root mean square error, which can be interpreted as the “average” residual (difference between observed and modeled concentration)
3 Number of sites that have been used for model development.
4 Mean [min - max]. Units are $10^{-5}m^{-1}$ for soot.

Table S2 ESCAPE Land Use Regression models for NO2 and NOx (Beelen et al. 2013)

Study area	LUR model [1]	R^2 of model	R^2 validation	RMSE[2] (validation) (µg/m3)	Number of sites[3]	Measured concentration (µg/m^3) [4]
			NO2			
Helsinki and Turku, Finland	7.61 + 1.18E-5*TRAFLOAD_25 + 3.43E-8*TRAFLOAD_25_1000 + 0.04*ROADLENGTH_25 + 1.24E-3*ROADLENGTH_25_300 – 9.18E-5*URBGREEN_5004	83%	75%	3.4	40	18.9 [6.1 – 40.8]
Netherlands and Belgium	-7.80 + 1.18*REGIONALESTIMATE + 2.30E-5*POP_5000 + 2.46E-6*TRAFLOAD_50 + 1.06E-4*ROADLENGTH_1000 + 9.84E-5*HEAVYTRAFLOAD_25 + 12.19*DISTINVNEARC1 + 4.47E-7*HEAVYTRAFLOAD_25_500	86%	81%	5.1	80	30.9 [12.8 – 61.5]
Barcelona, Spain	3.16 + 6.26E-3*INTINVDIST1 + 1.18E-4*HDRES_300 + 992.09*DISTINVMAJOR2 + 3.51E-4*ROADLENGTH_1000	75%	68%	11.6	40	57.7 [13.8 – 109.0]
			NOx			
Helsinki and Turku, Finland	12.56 + 3.46E-5*TRAFLOAD_25 + 4.92E-8*TRAFLOAD_25_1000 + 1.70E-2*ROADLENGTH_100 – 5.58E-5*URBGREEN_10004 + 2.54E-3*HHOLD_300	85%	74%	7.8	40	30.6 [8.6 – 94.7]
Netherlands and Belgium	3.25 + 0.74*REGIONALESTIMATE + 4.22E-6*TRAFLOAD_50 + 6.36E-4*POP_1000 + 2.39E-6*HEAVYTRAFLOAD_500 + 71.65*DISTINVMAJOR1 + 0.21*MAJORROADLENGTH_25	87%	82%	11.2	80	51.8 [17.5 – 130.8]
Barcelona, Spain	32.85 + 2.55E-4*HDRES_300 + 2815.14*DISTINVMAJOR2 + 3.87E-5*TRAFLOAD_25	73%	65%	27.7	40	101.3 [21.0 – 236.4]

1 Most variables are buffers with _xxx indicating the size of the buffer in m (e.g. HHOLD_500 is the number of households in a 500m buffer). TRAFLOAD is traffic intensity*length of road in a buffer. HDRES and LDRES are high and low density residential land use. INTINVDIST is product of traffic intensity and inverse distance to the nearest road. SQRALT is square root of altitude. Major road as road classes 0, 1 and 2 (motorways, main roads of major importance and other main roads) from the central road network or roads with more than 5,000 vehicles/day based upon local networks with linked traffic intensity. INTINVDIST is traffic intensity multiplied by inverse distance (squared). DISTINVMAJOR is inverse distance to a major road.
2 RMSE is the root mean square error, which can be interpreted as the "average" residual (difference between observed and modeled concentration)
3 Number of sites that have been used for model development.
4 Mean [min - max]

Supporting Information 2: Temporal adjustment

In the Netherlands, measurements were done in 32 weeks (2 extra weeks were scheduled to replace some missing data). During two of these measuring weeks, data from the reference site were missing because of technical failure. In Spain, the number of measuring weeks was 30. Out of these 30 weeks, 7 had missing data at the reference site. Four missing measurements occurred because no units were available and three because of technical failure.

For the missing measurements on the reference site in Spain and the Netherlands, imputation was applied. The measurements on the reference site were compared to measurements at fixed monitoring sites from the RIVM ('Rijksinstituut voor Volksgezondheid en Milieu', the National institute for Public Health and the Environment) in the Netherlands and the National Network in Barcelona (Gencat, Generalitat de Catalunya). The regression formula comparing the VE^3SPA reference site with the fixed site was used to calculate the concentrations for the missing data (Table S3).

In the Netherlands, the R^2 of soot on the VE^3SPA reference site compared to black smoke on site 738 of the RIVM was 0.70. NO_2 and NOx also correlated best with site 738 and had a R^2 of 0.61 and 0.73 respectively. For $PM_{2.5}$ the best correlation (R^2=0.69) was with the PM_{10} concentration on site 633 of the RIVM. In Spain, the VE^3SPA $PM_{2.5}$ concentrations at the reference site correlated with a R^2 of 0.66 with the PM_{10} (better than with $PM_{2.5}$) measured at the IES Verdaguer site (Gencat). The NO_2 concentration at Hospitalet was used for the imputation of NO_2, the R^2 was 0.64. For NO_x the concentrations of NO_2 and NO at the Hospitalet site were combined. This correlated with an R^2 of 0.75 with the VE^3SPA NOx measurements. The NO_x concentrations were also used for the imputation of soot because soot was not available in the network, the R^2 was 0.61.

Table S3 Formula's for the imputation of missing values on the reference site.

Country	**Comp**	**Independent comp, site**	**n**	**α**	**β**	**Formula**[1]	R^2
NL	Soot	BS, 738	28	0,27	0,12	Vref=0.27+0.12*F	0,70
NL	$PM_{2.5}$	$PM_{2.5}$, 633	29	-3,78	0,78	Vref=-3,78+0,78*F	0,69
NL	NO_2	NO_2, 738	30	5,82	0,95	Vref=5,82+0,95*F	0,61
NL	NO_x	NO_x, 738	30	1,73	1,51	Vref=1,73+1,51*F	0,73
Spain	$PM_{2.5}$	PM_{10}, Verdaguer	20	-10,08	0,85	Vref=-10,08+0,85*F	0,66
Spain	Soot	NO_x, Hospitalet	22	-0,06	0,03	Vref=-0,06+0,03*F	0,61
Spain	NO_2	NO_2, Hospitalet	24	-0,16	1,21	Vref=-0,16+1,21*F	0,64
Spain	NO_x	NO_x, Hospitalet	24	6,96	1,26	Vref=6,96+1,26*F	0,71

The imputed component (comp) and the independent component with the name or number of the fixed monitor site, the number of measurements in the linear regression (n), the intercept (α), the Beta (β) and the R^2 are given and entered in the imputation formula. Note: Black Smoke and soot are different methods for Black Carbon determination.
1 Vref is the VE3SPA reference site and F is the fixed reference site

The median slopes that were used for the temporal adjustment of the indoor and personal measurements per country are shown (Table S4). The slopes were consistent across the three countries and were all smaller than 1. This makes the influence of the adjustment with the reference site less, which is plausible for the indoor and personal measurements. The median correlation coefficients (R) of these regressions are shown in Table S5.

Table S4 The median slopes (β) per country of the within person (or temporal) relation between the outdoor measurements and the personal/indoor measurements.

Correlation β	**Finland**		**Netherlands**		**Spain**	
Component	**indoor**	**pers**	**indoor**	**pers**	**indoor**	**pers**
Soot (10-5/m)	0,49	0,42	0,68	0,66	0,62	0,40
PM2.5 (µg/m3)	0,38	0,60	0,41	0,44	0,59	0,64
NO2 (µg/m3)	0,11	0,35	0,29	0,37	0,20	0,25
NOx (µg/m3)	0,65	0,80	0,44	0,75	0,88	0,60

Table S5 The median correlation coefficients (R) of the association between the home outdoor measurements and the personal/indoor measurements per participant, per city

	Helsinki		**Utrecht**		**Barcelona**	
Component	**indoor**	**pers**	**indoor**	**pers**	**indoor**	**pers**
$PM_{2.5}$	0.51	0.67	0.82	0.71	0.79	0.41
Soot	0.79	0.72	0.95	0.89	0.81	0.67
NO_2	0.25	0.77	0.57	0.77	0.48	0.43
NOx	0.85	0.91	0.84	0.85	0.95	0.86

*All correlation coefficients were statistically significant ($p<0.05$)

Table S6 Median correlation coefficients (R) of the outdoor measurements versus the measurements at the reference site (including the imputed values) per ID

Median Correlation	**Finland**	**Netherlands**	**Spain**
Component	**ref**	**ref**	**ref**
Soot (10-5/m)	0,90	0,88	0,72
PM2.5 (µg/m3)	0,88	0,92	0,79
NO2 (µg/m3)	0,97	0,92	0,89
NOx (µg/m3)	0,95	0,93	0,86

After imputing the missing measurements at the reference site, the measurements at the reference site were plotted against the outdoor measurements per participant to show the temporal association of the measurements at the reference site and the outdoor concentrations at the participant address. The median correlation coefficients are shown in Table S6. For most of the volunteers in the three countries, the reference site was well able to predict the temporal fluctuation of the outdoor measurements at the home addresses, the median correlation coefficients in Finland and the Netherlands were very good. This supports the use of the reference site to correct for temporal variability. In Spain, the median correlation coefficients were lower for especially soot and $PM_{2.5}$, but still good. If only the non-imputed measurements at the reference site are plotted, the median correlation coefficients for Spain are a bit higher for soot (0.78) and lower (0.73) for $PM_{2.5}$. This indicates that the reference site in Spain was less predictive for the temporal fluctuations at the sites, which was not explained by the larger number of imputed values.

The corrected mean concentrations correlated well with the uncorrected mean of the measurements (Table S7). The correlation of the corrected and uncorrected mean was less for $PM_{2.5}$ than for the other components, as observed in the ESCAPE project.[1, 2] This was explained by the authors by the observation that the measured $PM_{2.5}$ concentration had lower seasonal variability and were more dependent on the weather, resulting in larger within season variability. Therefore the uncorrected mean could deviate more from the corrected mean.[1]

The outdoor corrected measurements had lower correlations with the uncorrected measurements than the indoor and personal measurements. This was a result of the procedure in which for the outdoor measurements we directly subtracted the differences at the reference site, while we used the difference multiplied by an indoor/outdoor slope (<1) for the indoor and personal measurements.

Table S7 The correlation of the corrected mean versus the uncorrected mean concentrations.

PM2.5	**Outdoor R^2**	**Indoor R^2**	**Personal R^2**
Netherlands	0,40	0,96	0,85
Finland	0,26	0,99	0,86
Spain	0,41	0,77	0,85
Soot	**Outdoor R^2**	**Indoor R^2**	**Personal R^2**
Netherlands	0,93	0,89	0,81
Finland	0,81	0,99	0,98
Spain	0,69	0,81	0,87
NO2	**Outdoor R^2**	**Indoor R^2**	**Personal R^2**
Netherlands	0,95	0,99	0,97
Finland	0,81	1,00	0,98
Spain	0,88	0,99	0,97
NOx	**Outdoor R^2**	**Indoor R^2**	**Personal R^2**
Netherlands	0,96	1,00	0,97
Finland	0,82	0,99	0,98
Spain	0,91	0,85	0,87

Supporting Information 3: Modeled PM2.5, NO2, NO_x and measured personal soot

In Table S8 the associations of the measured soot concentrations with the modeled personal $PM_{2.5}$, NO_2 and NO_x concentrations are shown. Associations are usually stronger than with the personal exposure of the component itself.

Table S8 The coefficient of determination (R^2), the regression coefficient (β), the intercept (α) and the p-value (p) of the regression analysis of the ESCAPE measured soot concentrations versus the mean modeled PM2.5, NO2 and NOx concentrations.

Model	R2	β	α	p
		Helsinki		
PM2.5	0.29	0.29	-1.68	0.04
NO2	0.47	0.04	0.01	0.00
NOx	0.49	0.02	0.16	0.00
		Utrecht		
PM2.5	0.32	0.09	-0.44	0.03
NO2	0.60	0.02	0.46	0.00
NOx	0.55	0.01	0.61	0.00
		Barcelona		
PM2.5	0.00	0.01	2.00	0.81
NO2	0.16	0.01	1.42	0.15
NOx	0.21	0.00	1.64	0.08

Supporting Information 4: Site description

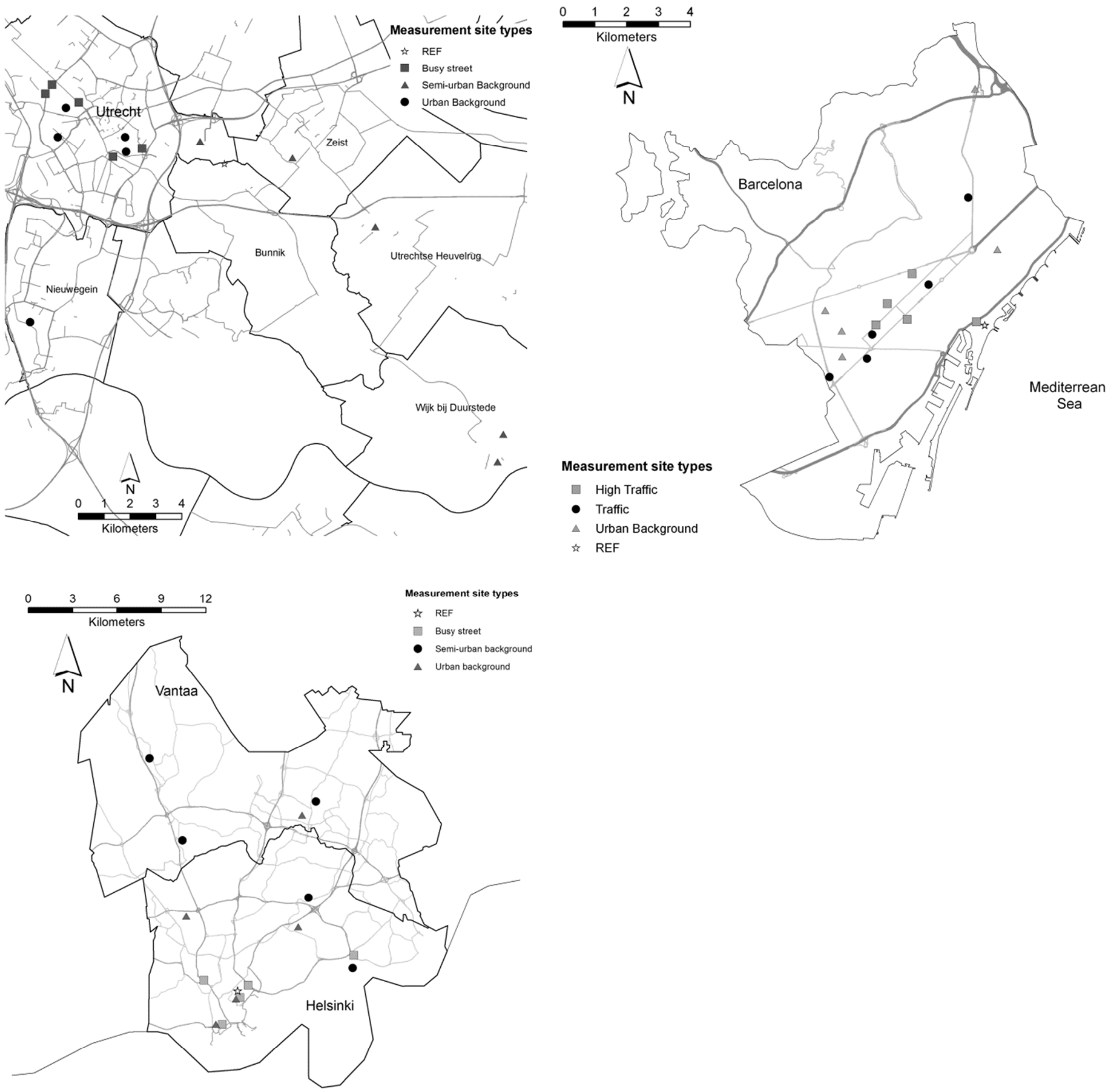
Table S9 Distribution of home characteristics

		Utrecht	Barcelona	Helsinki
Traffic intensity ($veh.day^{-1}$)*	Traffic site	16805 (5066)	21722 (9840)	18675(13886)
	Urban background	1511 (1660)	10254 (7386)	1053 (747)
	Semi-urban background	542 (453)	1040 (721)	867 (1770)
Sampling height (m)*	Traffic site	3.3 (7.2)	5.2 (4.1)	15.4 (9.2)
	Urban background	4.3 (0.8)	6.4 (3.4)	10.8 (6.6)
	Semi-urban background	4.4 (3.3)	9.4 (4.9)	2.4 (1.4)
House type	Detached family home	0	0	2
	Attached family home	11	0	2
	Flat/apartment	4	15	11
Home volume	<100 m^3	0	3	3
	100-200 m^3	6	8	6
	200-300 m^3	5	2	4
	>300 m^3	4	2	2
Built	1800-1944	8	3	2
	1945-1979	5	10	8
	>1980	2	2	5
Type kitchen is open (n (%))		9 (60)	0 (0)	6 (40)
Air conditioning (n)		1 (6.7)	8 (53)	0 (0)
Heating	Central in home	11	2	1
	District	1	0	11
	Separate gas/oil heaters	3	6	0
	Electric	0	4	3

	No heating	0	3	0
Cooking on gas (n %)		12 (80)	14 (93)	1 (6.7)
Living room on	Ground level	6	0	6
	First floor	6	8	3
	>2nd floor	3	7	6
Fume hood	No fume hood	5	5	3
	Exhaust air recirculated	2	0	3
	With external vent (*n* (%))	8	10	9
Pet in home	Cat	6	0	3
	Dog	4	1	3
	Other	5	1	1
Floor cover	Smooth	9	10	6
	Carpet	2	4	0
	Rug (larger than 1 m^3)	4	1	9

* Mean (standard deviation). Traffic counted manually (van Roosbroeck, 2007)

Figure S1 Maps of the VE3SPA sites in Helsinki (top left), Barcelona (top right) and Utrecht (bottom left).



Supporting Information 5: Quality assurance, quality control

Table S10 describes the results of the field blanks that were collected. Overall, the field blanks were low, although there were a few blank $PM_{2.5}$ measurements that had quite high negative values. It could be that while transporting the filters to the filter cassette, a small part of the support filter was shipped of. This was not noted by the field worker and thus the blank was still included in the analyses.
The soot measurements were all above the detection limit (DL), most $PM_{2.5}$ and NO_2 as well (Table S11).
The measurements that were below the detection limit are nonetheless included in the analysis. Finland had the most measurements under the detection limit, because of the overall low concentrations in Finland (section 6.2.6).
The results of the duplicate measurements are described in Table S12. Most coefficients of variation are about 10% which is acceptable. The CV for $PM_{2.5}$ is larger for the Netherlands than for Finland and Spain. For some of the duplicate measurements a pump unit was used that had given some technical problems in the past. The best units were used for the actual measurements. The coefficient of variance (CV) for $PM_{2.5}$ of 21.63% in the Netherlands thus likely gives an upper estimate of the uncertainty. A CV value of 10% in individual measurements translates into a CV of the average (based upon six measurements) of about 4% (10% / √6). A CV value of 20% in individual measurements translates into a CV of the average (based upon six measurements) of about 8% (20% / √6). Because of the small number of personal duplicates for $PM_{2.5}$, soot and NO_2, these CV values are less interpretable.

Table S10 Field blanks and detection limit for PM2.5 (μg/m3), soot reflectance , NO2 and NO (μg/m3)

	N field blanks		**Average field blank**				**Detection limit**			
Study area	**PM2.5**	**NOx**	**PM2.5 (μg)**	**Soot***	**NO2 (μg)**	**NOx (μg)**	**PM2.5 (μg/ m^3)**	**Soot (10^{-5} m^{-1})**	**NO2 (μg/ m^3)**	**NOx (μg/ m^3)**
Utrecht	15	13	0.00	102.7	-.01	-0,99	1.22	0.12	2,1	12,0
Helsinki	16	17	0.00	101.9	0.51	1,24	2.35**	0.11	5,8	5,8
Barcelona	15	13	0,01	102.3	0.54	1,75	5.17†	0.13	1,9	4,3

*this is the R0 in the formula for the absorption coefficient that can be found on page 11.
** When the sample of S13 8-10-2010 (filter 10756) was excluded, the detection limit was 1.15 μg/m^3

† When the sample of F113 at 6-12-2010 (filter 10534) was excluded, the detection limit was 0.73 μg/m^3

Table S11 Number of included samples below the detection limit.

	measurements (n)			
Study area	**$PM_{2.5}$**	**Soot**	**NO_2**	**NO_x**
Utrecht	1	0	0	7
Helsinki	16*	0	5	8
Barcelona	8*	0	1	0

* When the Blank filters 10756 and 10534 were excluded from the detection limit calculations, none of the filters were below detection limit.

Table S12 Duplicate measurements

	n duplicates		CV (%)			
Outdoor/Indoor Duplicates						
Study area	**$PM_{2.5}$/ soot**	**NO_x**	**$PM_{2.5}$**	**Soot**	**NO_2**	**NO_x**
Utrecht	13	17	21,6%	11,1%	9,0%	10,2%
Helsinki	15	14	11,5%	12,6%	9,0%	8,3%
Barcelona	15	5	9,2%	6,8%	20,3%	5,7%
Personal Duplicates						
Utrecht	4	5	7,7%*	21,8%*	4,7%	3,0%
Helsinki	5	5	9,2%	7,5%	6,6%	26,7%†
Barcelona	5	3	40,8%**	27,6%	4,4%	7,4%

*if we don't include volunteer N16, the Dutch CV would be 2.31% for PM2.5 and 7.10% for Soot.
** If we don't include the personal duplicate measurement of volunteer S19 than the CV would be 23,65% for $PM_{2.5}$
†9.1% if not including volunteer F994

Supporting Information 6: Scatterplots modeled versus measured concentrations

Figure S2 Regression plots of the corrected measured average versus modeled concentration for NO2 and NOx in Utrecht, Helsinki and Barcelona

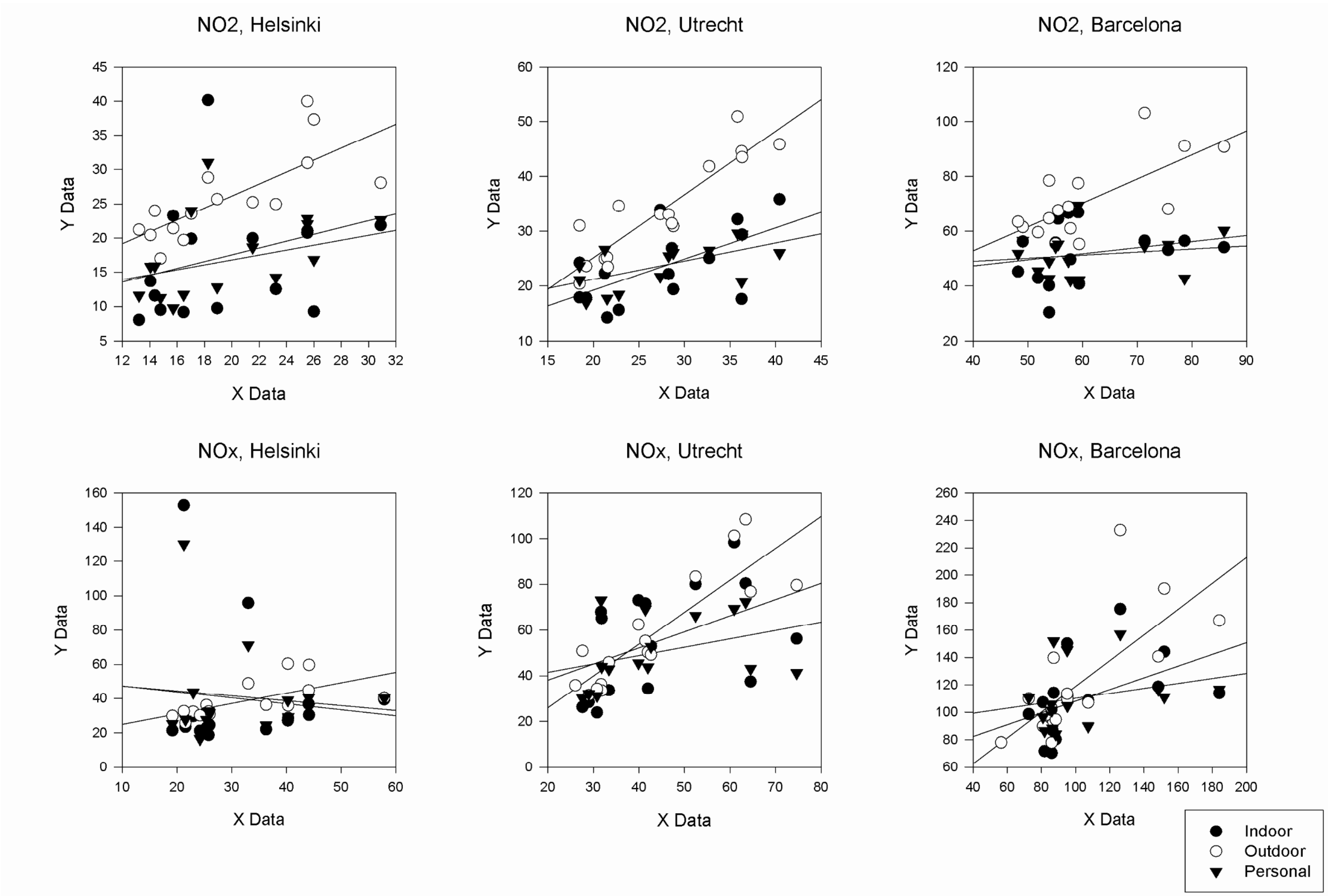
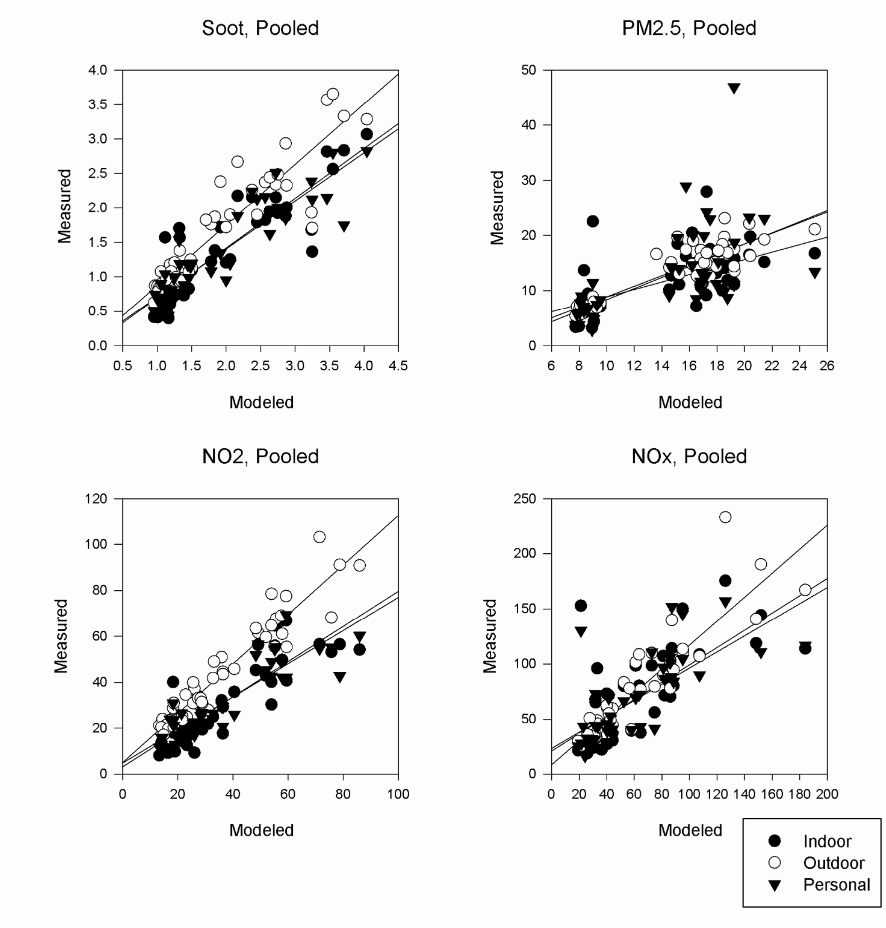


Figure S3 Regression of the pooled data set of the three countries.



Supporting Information 7: Associations of modeled and measured concentrations using an alternative adjustment method for temporal variation

When the concentrations were corrected by including the measurements at the reference site in the model, the results were comparable with the absolute differences correction method (Table S13). The p-values for the modeled concentrations (pm) show the probability that adding the modeled concentrations in the regression model is significant. With this method, the outdoor NO_x in Finland is not significant anymore. Also with the reference site included in the model, we see that the models significantly predict the outdoor concentrations of soot, NO_2 and NO_x in the three countries. Table S13 shows that the indoor NO_x for the Netherlands and Spain was significant when including the reference site in the model. With the absolute differences method, these associations were borderline significant. No associations were found with personal exposure.

Table S13 The coefficient of determination (R^2), the regression coefficient of the reference site (βr) the regression coefficient of the modeled concentrations (βm) and the intercept (α) of the regression analysis of the ESCAPE modeled concentrations versus the mean measured VE3SPA concentrations, with the measurements at the reference site included in the model

PM2.5	**Outdoor**				**Indoor**				**Personal**			
	R2	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**
Netherlands	0,66	0,86*	1,08*	-13,88	0,57	1,37*	0,32	-10,94	0,35	0,62*	-0,39	10,12
Finland	0,78	0,93*	0,61	-4,88	0,23	1,44	1,34	-14,08	0,21	0,46	1,14	-6,61
Spain	0,57	0,84*	0,23	-0,79	0,06	0,27	0,00	11,86	0,06	0,64	-0,34	18,43
Soot	**Outdoor**				**Indoor**				**Personal**			
	R2	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**
Netherlands	0,83	1,60*	1,16*	-1,92	0,81	1,39*	0,65*	-1,24	0,68	1,03*	0,37*	-0,48
Finland	0,60	1,26*	0,94*	-1,22	0,41	1,84	1,93*	-2,99	0,47	1,13	1,45*	-1,86
Spain	0,67	-0,02	0,90*	0,01	0,61	-0,07	0,72*	0,16	0,49	0,53†	0,28	0,38
NO2	**Outdoor**				**Indoor**				**Personal**			
	R2	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**
Netherlands	0,89	1,84*	1,08*	-39,08	0,57	1,38*	0,52*	-23,28	0,58	1,01*	0,30*	-8,45
Finland	0,51	0,87	0,80*	-14,60	0,09	0,59	0,31	-5,78	0,28	0,58	0,47	-7,81
Spain	0,67	-0,23	1,05*	20,27	0,11	0,45	0,19	16,85	0,10	-0,36	0,20	58,37
NOx	**Outdoor**				**Indoor**				**Personal**			
	R2	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**	**R2**	**βr**	**βm**	**α**
Netherlands	0,87	2,49*	1,36*	-89,87	0,76	3,67*	0,61*	-104,27	0,65	2,63*	0,30	-57,34
Finland	0,41	1,28†	0,49†	-28,28	0,11	2,54	-0,31	-48,15	0,13	2,09	-0,28	-30,44
Spain	0,58	0,12	1,01*	9,67	0,31	0,22	0,48†	43,59	0,14	-0,22	0,25	104,39

Concentrations are corrected for weather differences using the inclusion of the reference site in the regression analysis method. βr reflects temporal variation. Bm the spatial component. The R^2 cannot be compared with Table 4 from the main paper.
* Significant at the p<0.05 level
† Significant at the p <0.10 level

P values Table S13

PM2.5	Outdoor		Indoor		Personal	
	pr	pm	pr	pm	pr	pm
Netherlands	0,002	0,011	0,002	0,589	0,037	0,394
Finland	0,000	0,182	0,131	0,584	0,274	0,315
Spain	0,002	0,380	0,393	0,996	0,435	0,707
Soot	**Outdoor**		**Indoor**		**Personal**	
	pr	**pm**	**pr**	**pm**	**pr**	**pm**
Netherlands	0,004	0,000	0,001	0,000	0,004	0,008
Finland	0,011	0,007	0,120	0,026	0,136	0,012
Spain	0,946	0,001	0,799	0,002	0,060	0,136
NO2	**Outdoor**		**Indoor**		**Personal**	
	pr	**pm**	**pr**	**pm**	**pr**	**pm**
Netherlands	0,001	0,000	0,030	0,017	0,015	0,026
Finland	0,153	0,018	0,510	0,500	0,333	0,132
Spain	0,646	0,000	0,474	0,501	0,424	0,310
NOx	**Outdoor**		**Indoor**		**Personal**	
	pr	**pm**	**pr**	**pm**	**pr**	**pm**
Netherlands	0,000	0,000	0,000	0,027	0,001	0,179
Finland	0,055	0,071	0,263	0,734	0,233	0,691
Spain	0,858	0,002	0,690	0,052	0,637	0,201

Supporting Information 8: Environmental Tobacco Smoke

Table S14 Regression analyses of the modeled outdoor versus the measured indoor/personal concentrations, the measurements with environmental tobacco smoke (ETS) are excluded from analysis. Only participants with more than one measurement are included.

ETS excluded from analyses										
PM2.5	**Indoor**					**Personal**				
Country	**N**	**R^2**	**β**	**α**	**p**	**N**	**R^2**	**β**	**α**	**p**
Netherlands	15	0,01	0,31	6,74	0,67	15	0,07	-0,39	18,18	0,35
Finland	15	0,04	1,90	-8,36	0,45	15	0,03	0,85	-0,58	0,51
Spain	13	0,00	0,08	14,96	0,86	13	0,01	-0,31	28,26	0,76
Soot	**Indoor**					**Personal**				
Country	**N**	**R^2**	**β**	**α**	**p**	**N**	**R^2**	**β**	**α**	**p**
Netherlands	15	0,64	0,63	0,10	0,00	15	0,42	0,35	0,50	0,01
Finland	15	0,30	1,83	-1,29	0,04	15	0,38	1,34	-0,75	0,01
Spain	13	0,35	0,49	0,69	0,03	13	0,14	0,25	1,42	0,21
NO2	**Indoor**					**Personal**				
Country	**N**	**R^2**	**β**	**α**	**p**	**N**	**R^2**	**β**	**α**	**p**
Netherlands	15	0,34	0,56	8,22	0,02	15	0,27	0,27	16,09	0,05
Finland	15	0,02	0,23	13,16	0,65	15	0,18	0,49	7,60	0,12
Spain	15	0,11	0,32	31,27	0,23	15	0,03	0,14	42,81	0,54
NOx	**Indoor**					**Personal**				
Country	**N**	**R^2**	**β**	**α**	**p**	**N**	**R^2**	**β**	**α**	**p**
Netherlands	15	0,12	0,53	30,80	0,21	15	0,05	0,23	39,32	0,42
Finland	15	0,01	-0,33	50,54	0,72	15	0,01	-0,25	48,74	0,72
Spain	15	0,17	0,40	67,87	0,13	15	0,01	0,10	101,20	0,67

The number of participants included in the analysis (N), the coefficient of determination (R^2), the regression coefficient (β), the intercept (α) and the p-value (p) of the model.

Supporting Information 9: Leave one out cross-validation

Table S15 The minimum (MIN), maximum(MAX) and standard deviation (STD) of the R squares derived from the leave one out cross validation. N=15, for the 15 models where N-1=14 participants.

		Outdoor Soot	Outdoor $PM_{2.5}$	Outdoor NO_2	Outdoor NOx
Helsinki	MIN	0.50	0.12	0.50	0.36
	MAX	0.68	0.28	0.71	0.62
	MEAN	0.57	0.21	0.56	0.42
	STD	0.04	0.05	0.05	0.06
Utrecht	MIN	0.72	0.37	0.79	0.72
	MAX	0.80	0.54	0.87	0.83
	MEAN	0.75	0.43	0.82	0.75
	STD	0.02	0.05	0.02	0.03
Barcelona	MIN	0.26	0.02	0.41	0.42
	MAX	0.49	0.17	0.60	0.68
	MEAN	0.33	0.10	0.49	0.51
	STD	0.06	0.04	0.05	0.06
		Personal Soot	**Personal $PM_{2.5}$**	**Personal NO_2**	**Personal NOx**
Utrecht	MIN	0.37	0.01	0.28	0.07
	MAX	0.61	0.28	0.50	0.30
	MEAN	0.44	0.07	0.35	0.14
	STD	0.06	0.06	0.06	0.07
Barcelona	MIN	0.05	0.00	0.00	0.04
	MAX	0.42	0.06	0.12	0.14
	MEAN	0.20	0.01	0.03	0.07
	STD	0.08	0.02	0.03	0.03
Helsinki	MIN	0.28	0.03	0.15	0.01
	MAX	0.45	0.17	0.40	0.08
	MEAN	0.39	0.08	0.21	0.02
	STD	0.05	0.04	0.06	0.02

References

1. Eeftens, M.; Tsai, M.; Ampe, C.; Anwander, B.; Beelen, R. 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* **2012,** *62,* 303-317.

2. Eeftens, M.; Beelen, R.; de Hoogh, K., et al Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. *Environ. Sci. Technol.* **2012,** *46,* 11195-11205.