SUPPORTING INFORMATION FOR:

Development of Land Use Regression Models for elemental, organic carbon, PAH and hopanes/steranes in 10 ESCAPE/TRANSPHORM European study areas

Aleksandra Jedynska, Gerard Hoek, Meng Wang, Marloes Eeftens., Josef Cyrys, Menno Keuken, Christophe Ampe, Rob Beelen, Giulia Cesaroni, Francesco Forastiere, Marta Cirach, Kees de Hoogh, Audrey De Nazelle, Wenche Nystad, Christophe Declercq, Kirsten T. Eriksen, Konstantina Dimakopoulou, Timo Lanki, Kees Meliefste , Mark J. Nieuwenhuijsen, Tarja Yli-Tuomi, Ole Raaschou-Nielsen, Bert Brunekreef, Ingeborg M. Kooter

Table of Contents

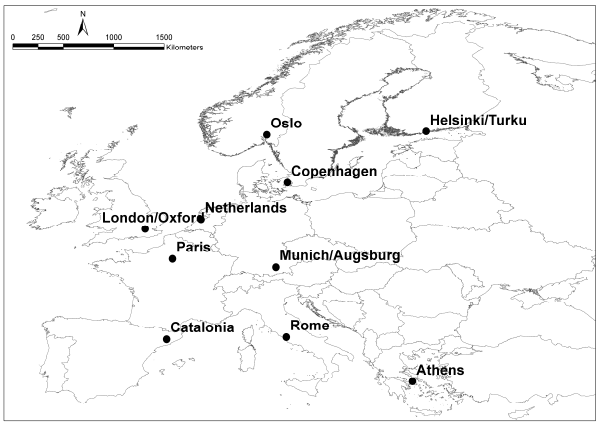


Figure S1. Map of 10 European study areas.

Table S1. Predictor variables with predefined variable names. units. defined buffer sizes. transformations of the predictor variables and directions of effect.

GIS dataset	Predictor variable	Name variable[1]	Unit	Buffer size (radius of buffer in meter)	Transformation	Direction of effect
Background						
-	Coordinate variables[2]	XCOORD, YCOORD	m	NA	Local decision	NA
CORINE	High density residential land [3]	HDRES	m^2	100, 300, 500, 1000, 5000	-	+
CORINE	Low density residential land [3]	LDRES	m^2	100, 300, 500, 1000, 5000	-	+
CORINE	Industry [3]	INDUSTRY	m^2	100, 300, 500, 1000, 5000	-	+
CORINE	Port [3]	PORT	m^2	100, 300, 500, 1000, 5000	-	+
CORINE	Urban green [3, 4]	URBGREEN	m^2	100, 300, 500, 1000, 5000	-	-
CORINE	Semi-natural and forested areas [3,5]	NATURAL	m^2	100, 300, 500, 1000, 5000	-	-
Local land use			m^2	100, 300, 500, 1000, 5000	-	Following CORINE
Population density	Number of inhabitants [3]	POP	N(umber)	100, 300, 500, 1000, 5000	-	+
Household density	Number of households	HHOLD	N(umber)	100, 300, 500, 1000, 5000	-	+
Altitude	Altitude	SQRALT	m	NA	square root	-
Traffic [6]						
Local road network	Traffic intensity [6] on nearest road	TRAFNEAR	Veh.day^{-1}	NA	-	+
Local road network	Distance to the nearest road	DISTINVNEAR1 DISTINVNEAR2	m^{-1}, m^{-2}	NA	Inverse distance and inverse distance squared	+
Local road network	Product of traffic intensity on nearest road and inverse of distance to the nearest road and distance squared	INTINVDIST INTINVDIST2	Veh.day-m^{-1} Veh.day-m^{-2}	NA	-	+
Local road network	Traffic intensity on nearest major road [7]	TRAFMAJOR	Veh.day^{-1}	NA	-	+
Local road network	Distance to the nearest major road [7]	DISTINVMAJOR1 DISTINVMAJOR2	m^{-1}, m^{-2}	NA	Inverse distance and inverse distance squared	+
Local road network	Product of traffic intensity on nearest major road and inverse of distance to the nearest major road and distance squared [7]	INTMAJORINVDIST INTMAJORINVDIST2	Veh.day-m^{-1} Veh.day-m^{-2}	NA	-	+
Local road network	Total traffic load of major roads in a buffer (sum of (traffic intensity * length of all segments)) [7]	TRAFMAJORLOAD	Veh.day-m	25, 50, 100, 300, 500, 1000	-	+
Local road network	Total traffic load of all roads in a buffer (sum of (traffic intensity * length of all segments))	TRAFLOAD	Veh.day-m	25, 50, 100, 300, 500, 1000	-	+
Local road network	Heavy-duty traffic intensity on nearest road	HEAVYTRAFNEAR*	Veh.day^{-1}	NA	-	+
Local road network	Product of Heavy-duty traffic intensity on nearest road and inverse of distance to the nearest road and distance squared	HEAVYINTINVDIST HEAVYINTINVDIST2	Veh.day-m^{-1} Veh.day-m^{-2}	NA	-	+
Local road network	Heavy-duty traffic intensity on nearest major road	HEAVYTRAFMAJOR	Veh.day^{-1}	NA	-	+
Local road network	Total heavy-duty traffic load of major roads in a buffer (sum of (heavy-duty traffic intensity * length of all segments))	HEAVYTRAFMAJORLOAD	Veh.day-m	25, 50, 100, 300, 500, 1000	-	+
Local road network	Total heavy-duty traffic load of all roads in a buffer (sum of (heavy-duty traffic intensity * length of all segments))	HEAVYTRAFLOAD	Veh.day-m	25, 50, 100, 300, 500, 1000	-	+
Central road network	Road length of all roads in a buffer	ROADLENGTH	m	25, 50, 100, 300, 500, 1000	-	+
Central road network	Road length of major roads in a buffer [8]	MAJORROADLENGTH	m	25, 50, 100, 300, 500, 1000	-	+
Central road network	Distance to the nearest road	DISTINVNEARC1 DISTINVNEARC2	m^{-1}, m^{-2}	NA	Inverse distance and inverse distance squared	+
Central road network	Distance to the nearest major road [8]	DISTINVMAJORC1 DISTINVMAJORC2	m^{-1}, m^{-2}	NA	Inverse distance and inverse distance squared	+
	Aspect ratio (sum height buildings both side of road divided by road width) [9]	CANYON [9]	m/m	NA		

Table S2. Spearman correlation between B[a]P , chrysene and traffic and wood smoke marker.

Study area	PAH	NOx	EC	levoglucosan
Oslo	chrysene	0.79	0.87	0.84
	B[a]P	0.77	0.85	0.88
Helsinki/Turku	chrysene	0.24	0.64	
	B[a]P	0.13	0.56	
Copenhagen	chrysene	0.03	0.14	
	B[a]P	0.00	0.14	
London/Oxford	chrysene	0.73	0.74	
	B[a]P	0.49	0.57	
Netherlands	chrysene	0.47	0.50	0.62
	B[a]P	0.44	0.51	0.66
Munich/Augsburg	chrysene	0.75	0.75	0.19
	B[a]P	0.46	0.50	0.51
Paris	chrysene	0.81	0.86	
	B[a]P	0.78	0.87	
Rome	chrysene	0.77	0.77	
	B[a]P	0.69	0.63	
Catalonia	chrysene	0.78	0.77	0.07
	B[a]P	0.59	0.59	0.32
Athens	chrysene	0.40	0.54	
	B[a]P	0.65	0.56	
Median	chrysene	0.74	0.75	0.41
	B[a]P	0.54	0.56	0.58

Table S3. Mean and contrast of annual averages for 10 European study areas.

Study area	n	OC			EC			ΣPAH			B[a]P			chrysene			Σhopanes/steranes		
		Mean (µg/m^3)	Range	Range/Mean (%)	Mean (µg/m^3)	Range	Range/Mean (%)	Mean (ng/m^3)	Range	Range/Mean (%)	Mean (ng/m^3)	Range	Range/Mean (%)	Mean (ng/m^3)	Range	Range/Mean (%)	Mean (ng/m^3)	Range	Range/Mean (%)
Oslo	19	1.7	3.3	199	1.0	2.0	200	1.2	3.3	280	0.14	0.41	292	0.20	0.58	295	2.3	2.8	124
Helsinki/Turku	20	1.1	2,0	189	0.8	2.1	254	1.8	3.3	183	0.17	0.39	221	0.25	0.44	178	N/A	N/A	N/A
Copenhagen	20	1.5	2.1	141	1.0	1.2	124	2.1	5.7	276	0.21	0.75	352	0.34	0.84	246	N/A	N/A	N/A
London/Oxford	20	1.4	3.1	220	1.3	5.3	399	1.0	2.7	276	0.09	0.31	354	0.17	0.52	300	N/A	N/A	N/A
Netherlands	16	1.8	1.9	107	1.2	1.4	115	1.4	2.8	200	0.14	0.27	189	0.25	0.57	226	2.8	3.7	135
Munich/Augsburg	20	2.7	3.1	117	1.5	2.2	149	1.3	2.0	157	0.13	0.24	178	0.17	0.32	188	2.2	2.0	91
Paris	20	2.2	5.5	244	1.8	7.6	420	1.5	4.0	256	0.14	0.44	307	0.27	0.91	336	N/A	N/A	N/A
Rome	20	3.7	3.1	84	2.3	2.6	115	2.0	3.7	183	0.19	0.43	222	0.31	0.80	261	N/A	N/A	N/A
Catalonia	40	2.8	5.8	210	2.2	4.8	222	1.6	2.8	179	0.17	0.35	204	0.24	0.58	243	5.0	14.9	297
Athens	20	3.5	4.3	121	1.6	2.9	177	2.0	3.9	196	0.25	0.65	266	0.24	0.54	222	N/A	N/A	N/A

Table S4. Description of land use regression model for ∑PAH (ng/m^3)

Study area	LUR model	n	R^2 (%)	LOOCV R^2 (%)	RMSE (ng/m^3)	Measured concentration (ng/m^3)
Oslo	$1.227 + 7.297 \times 10^{-12} \times$ TRAFLOAD_1000 - 0.0580 SQRALT	19	71	58	0.526	1.18 [0.06 - 3.36]
Helsinki/Turku*	-	20				1.78 [0.51 - 3.77]
Copenhagen*	-	20				2.08 [0.69 - 6.43]
London/Oxford	$0.957 + 78.388 \times$ DISTINVMAJORC2 - $4.911 \times 10^{-8} \times$ NATURAL_5000 - $1.942 \times 10^{-8} \times$ URBGREEN_5000 + $0.00000503 \times$ TRAFMAJOR + $0.000741 \times$ MAJORROADLENGTH_100	20	93	89	0.175	0.98 [0.23 - 2.93]
Netherlands	$1.621 + 0.0061 \times$ MAJORROADLENGTH_50 - $5.429 \times 10^{-8} \times$ UGNL_5000	16	58	31	0.511	1.40 [0.44 - 3.24]
Munich/Augsburg**	$1.030 + 0.00224 \times$ MAJORROADLENGTH_100	20	31	9	0.430	1.25 [0.48 - 2.43]
Paris**	$0.380 + 0.000214 \times$ MAJORROADLENGTH_500 + $2.649 \times 10^{-8} \times$ LDRES_5000 - $5.850 \times 10^{-8} \times$ URBGREEN_5000 + $0.0000123 \times$ LDRES_100	20	85	63	0.359	1.54 [0.48 - 4.43]
Rome	$1.226 + 19.120 \times$ DISTINVMAJOR1 + $0.00153 \times$ MAJORROADLENGTH_50 + $6.061 \times 10^{-8} \times$ INDUSTRY_5000	20	84	74	0.389	2.03 [0.86 - 4.58]
Catalonia	$0.761 + 0.000228 \times$ INTMAJORINVDIST1 + $9.957 \times 10^{-7} \times$ HDRES_500	40	37	27	0.597	1.56 [0.28 - 3.07]
Athens**	$1.095 + 0.00897 \times$ ROADLENGTH_25 + $3.470 \times 10^{-8} \times$ TRAFLOAD_100	20	44	8	0.657	1.99 [0.95 - 4.84]
Median			65	45		

* no model possible due to an influential site

** too high Cook's D value

Table S5. Description of land use regression models for chrysene (ng/m^3)

Study area	LUR model	n	R^2 (%)	LOOCV R^2 (%)	RMSE (ng/m^3)	Measured concentration (ng/m^3)
Oslo	0.157 + 1.165 x 10^{-12} x TRAFLOAD_1000 + 0.00463 x MAJORROADLENGTH_25 + 279.4 x HEAVYINTINVDIST2 - 0.0126 x SQRALT	19	83	65	0.0810	0.20 [0.00 - 0.58]
Helsinki/Turku	0.169 +0.000760 x ROADLENGTH_50	20	23	8	0.116	0.25 [0.07 - 0.51]
Copenhagen*	-	20				0.34 [0.08 - 0.92]
London/Oxford	0.0662 + 0.0000309 x HEAVYTRAFMAJOR + 2.0342 x DISTINVMAJORC1	20	91	89	0.0377	0.17 [0.04 - 0.56]
Netherlands	0.286 + 0.00149 x MAJORROADLENGTH_50 - 1.162 x 10^{-8} x UGNL_5000	16	71	50	0.0927	0.25 [0.07 - 0.64]
Munich/Augsburg	0.0907 + 2.826 x 10^{-9} x TRAFMAJORLOAD_300 + 0.000582 x MAJORROADLENGTH_50_v2	20	46	33	0.0739	0.17 [0.06 - 0.38]
Paris	0.106 + 7.963 x 10^{-9} x INDUSTRY_5000 + 1.346 x 10^{-7} x POP_5000 - 1.439 x 10^{-8} x URBGREEN_5000 + 4.808 x 10^{-8} x LDRES_1000	19	84	73	0.0439	0.27 [0.05 - 0.97]
Rome	0.0110 + 0.000410 x POP_100 + 6.556 x 10^{-7} x LDRES_300 + 0.000977 x MAJORROADLENGTH_50	20	90	83	0.0648	0.31 [0.10 - 0.90]
Catalonia	0.0764 + 0.0000526 x INTMAJORINVDIST1 + 1.915 x 10^{-7} x HDRES_500	40	55	47	0.0893	0.24 [0.04 - 0.62]
Athens	0.205 + 1.686 x 10^{-8} x TRAFMAJORLOAD_25	19	41	30	0.0470	0.24 [0.10 - 0.64]
Median			71	50		

* no model possible due to an influential site

Table S6. Description of land use regression model for Σhopanes/steranes (ng/m^3)

Study area	LUR model	n	R^2 (%)	LOOCV R^2 (%)	RMSE (ng/m^3)	Measured concentration (ng/m^3)
Oslo	1.938 + 5.456 x 10^{-7} x PORT_5000	19	18	2	0.823	2.29 [1.30 - 4.13]
Netherlands	5.750 + 28.658 x DISTINVMAJOR2 + 0.0000862 x POP_500 - 0.00000759 x YCOORD	16	78	63	0.519	2.75 [1.46 - 5.16]
Munich/Augsburg	52.895 + 0.000849 x ROADLENGTH_100 + 0.000165 x INTMAJORINVDIST - 0.0000115 x XCOORD	20	58	41	0.403	2.18 [1.00 - 2.98]
Catalonia	1.081 + 0.00000190 x HDRES_1000	40	27	21	3.004	5.00 [1.56 - 16.4]
Median			43	31		

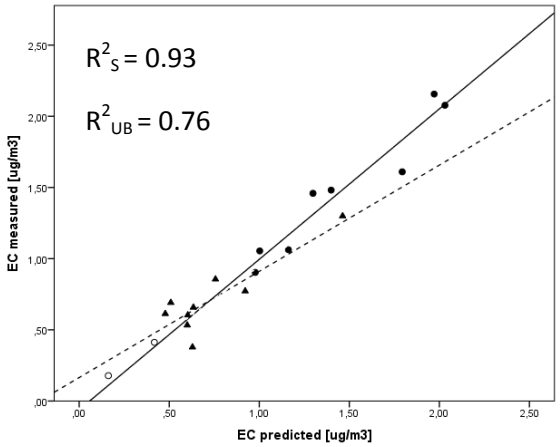
Table S7. Average of the model R2 and LOOCV R2 for EC. OC. B[a]P. ΣPAH and chrysene per study area

	R2 (%)	LOOCV R2 (%)
Oslo	76	61
Helsinki/Turku	29	22
Copenhagen	24	21
London/Oxford	87	81
Netherlands	74	56
Munich/Augsburg	51	35
Paris	86	72
Rome	71	60
Catalonia	50	40
Athens	59	43

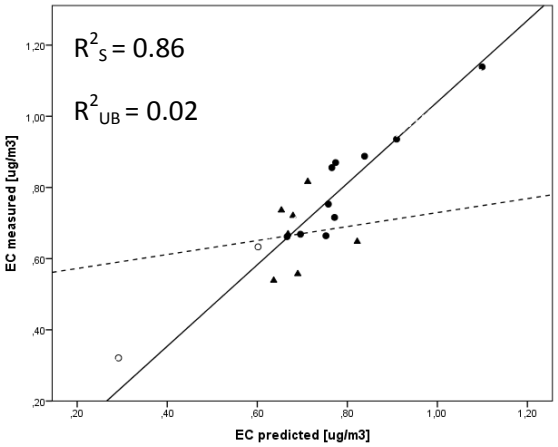
Table S8. Liniear regression between EC and PM2.5absorbance

Study area		Regression	R^2
Oslo	with intercept	EC = -0.407 + 1.061*PM2.5abs	0.894
	w/o intercept	EC = 0.789*PM2.5abs	0.960
Helsinki/Turku	with intercept	EC = -0.327 + 1.017*PM2.5abs	0.842
	w/o intercept	EC = 0.752*PM2.5abs	0.957
Copenhagen	with intercept	EC = -0.136 + 0.909*PM2.5abs	0.636
	w/o intercept	EC = 0.807*PM2.5abs	0.949
LondonOxford	with intercept	EC = -0.672 + 1.223*PM2.5	0.890
	w/o intercept	EC = 0.896*PM2.5abs	0.924
Netherlands	with intercept	EC = -0.090 + 0.819*PM2.5	0.951
	w/o intercept	EC = 0.767*PM2.5abs	0.995
Munich/Augsburg	with intercept	EC = -1.352 + 1.480*PM2.5	0.955
	w/o intercept	EC = 0799*PM2.5	0.958
Paris	with intercept	EC = -1.342 + 1.575*PM2.5	0.908
	w/o intercept	EC = 1.039*PM2.5	0.898
Rome	with intercept	EC = -0.443 + 0.944*PM2.5	0.695
	w/o intercept	EC = 0.798*PM2.5	0.965
Catalonia	with intercept	EC = -0.668 + 1.144*PM2.5	0.832
	w/o intercept	EC = 0.904*PM2.5	0.956
Athens	with intercept	EC = -0.596 + 0.941*PM2.5	0.619
	w/o intercept	EC = 0.700*PM2.5abs	0.943

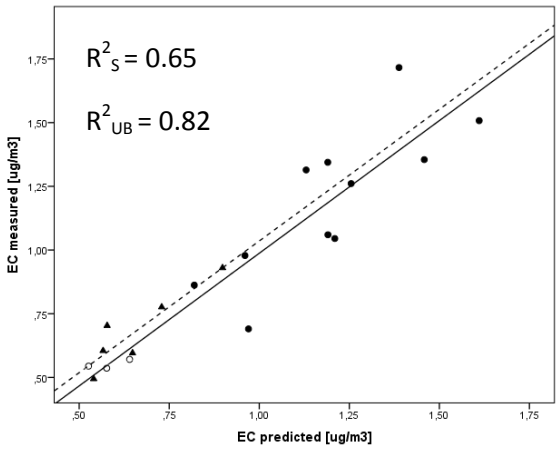
Oslo



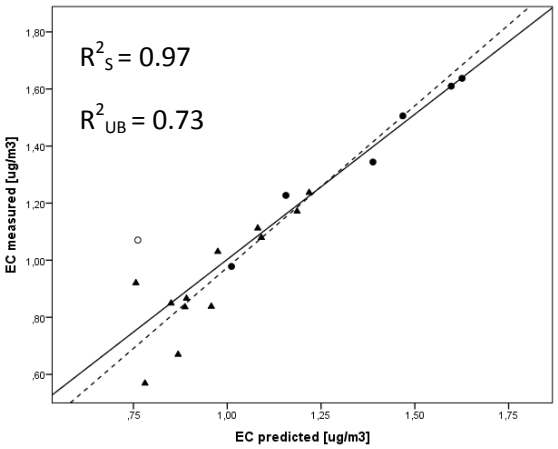
Helsinki/Turku



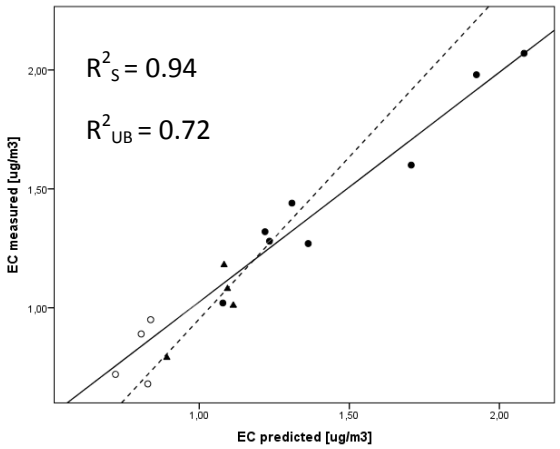
Copenhagen



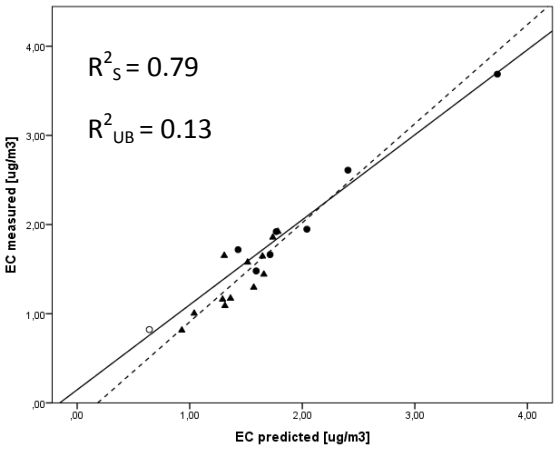
London/Oxford



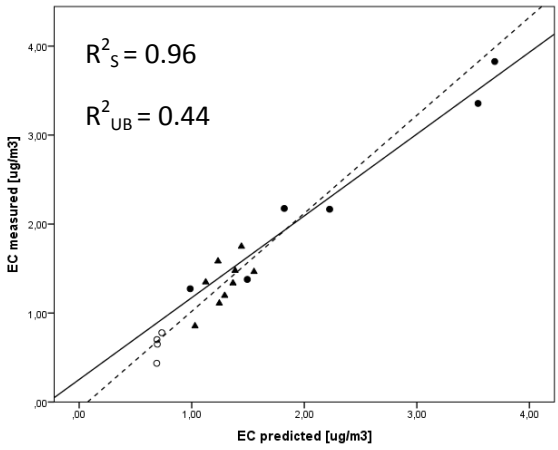
Netherlands



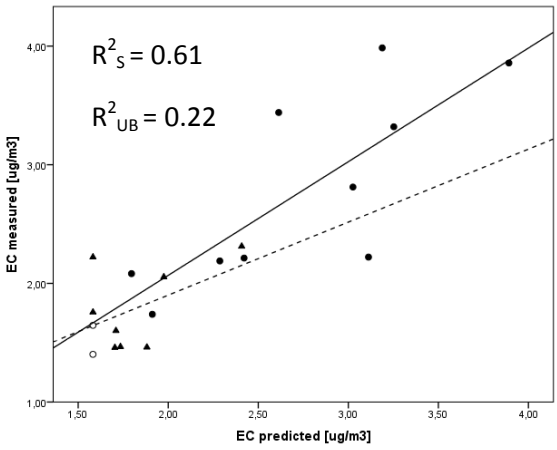
Munich/Augsburg



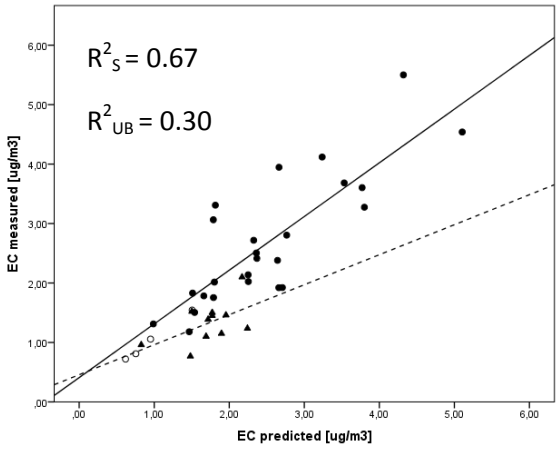
Paris



Rome



Catalonia



Athens

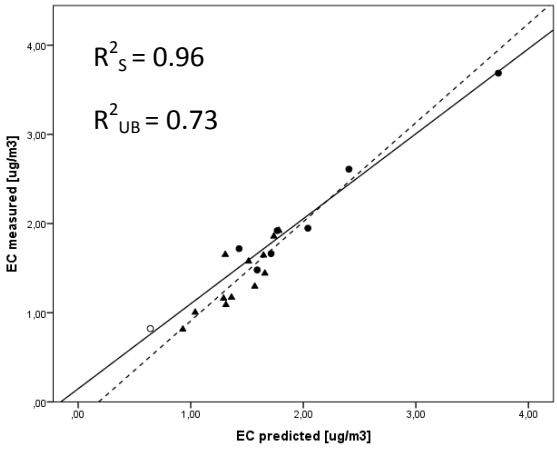
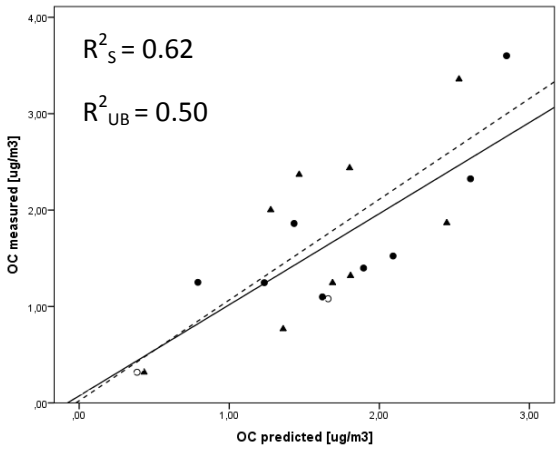
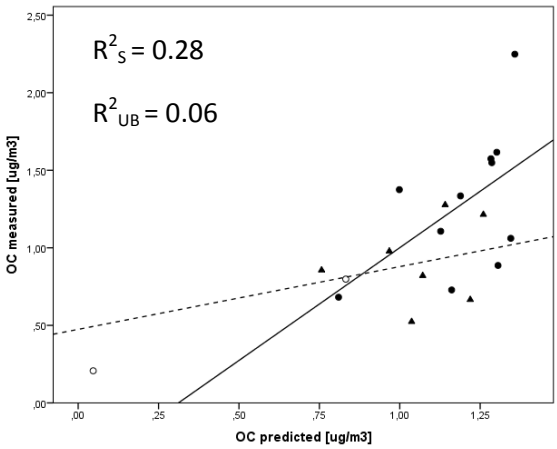


Figure S2. Correlation between measured and predicted EC annual averages for S - ●,UB - ▲. RB - ○. Solid line – S. dotted line - UB

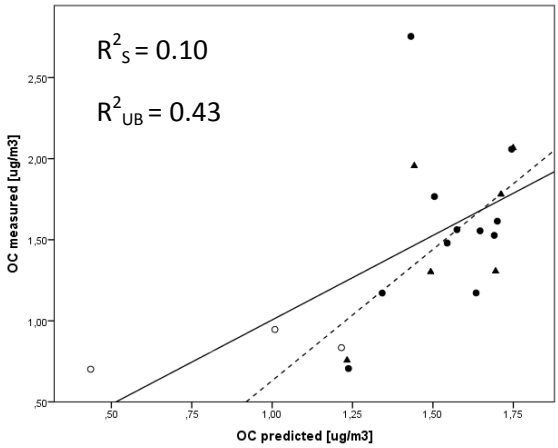
Oslo



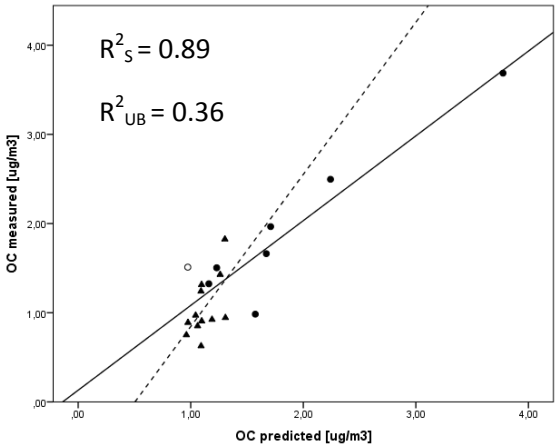
Helsinki/Turku



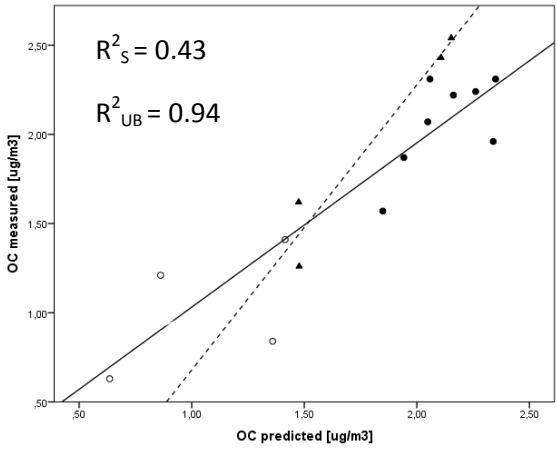
Copenhagen



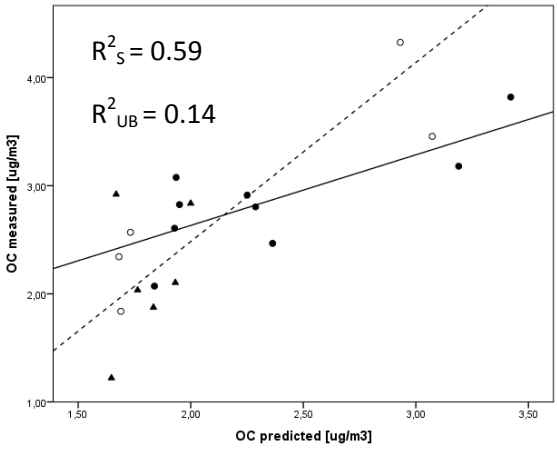
London/Oxford



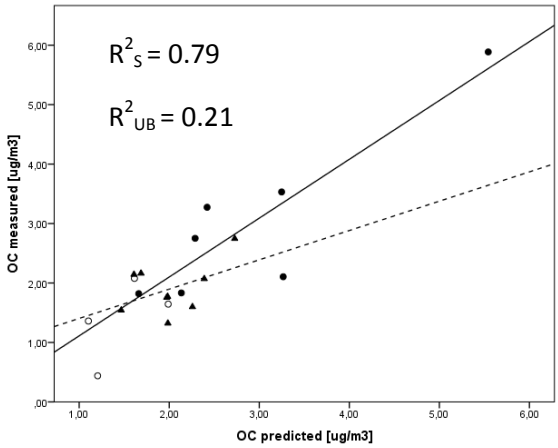
Netherlands



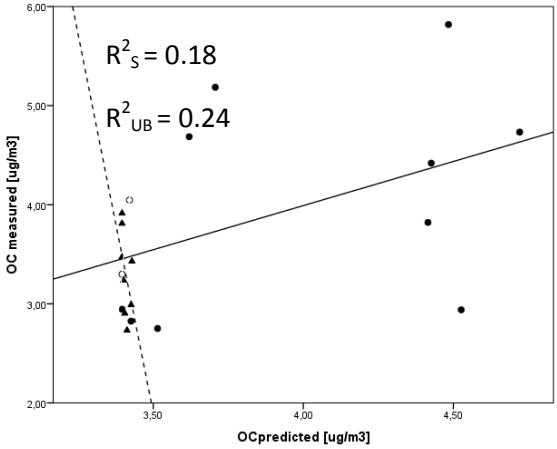
Munich/Augsburg



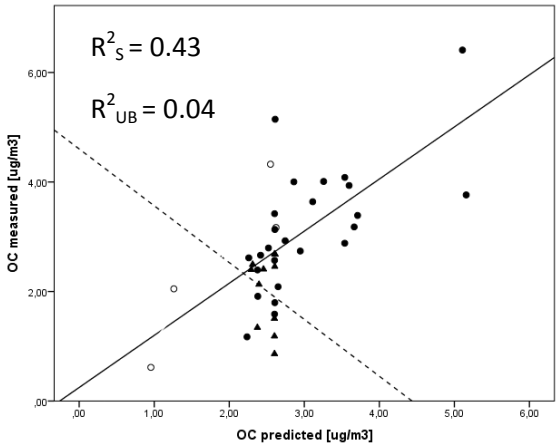
Paris



Rome



Catalonia



Athens

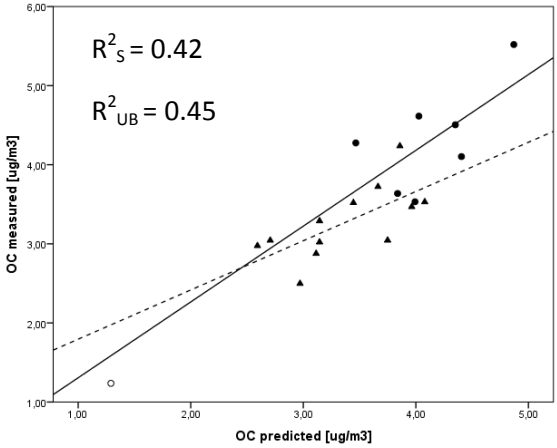
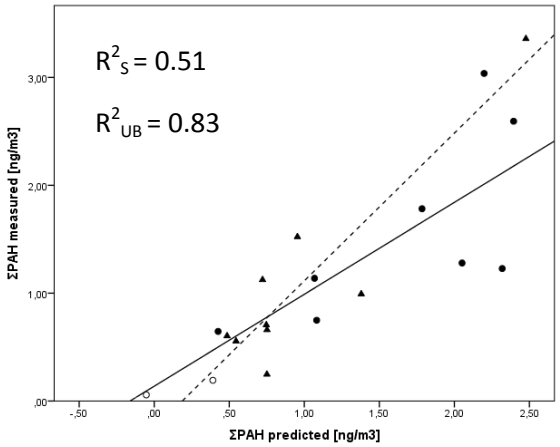
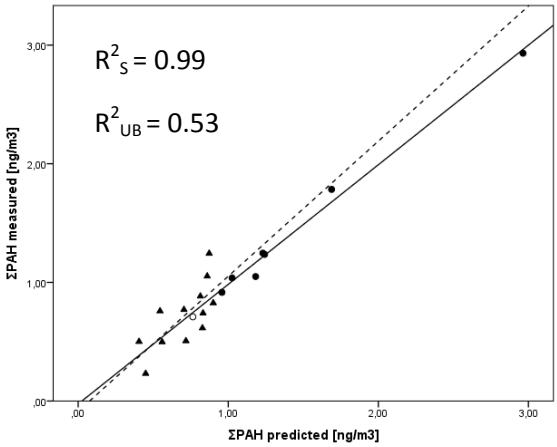


Figure S3. Correlation between measured and predicted OC annual averages for S - ●,UB - ▲, RB - ○. Solid line – S, dotted line - UB

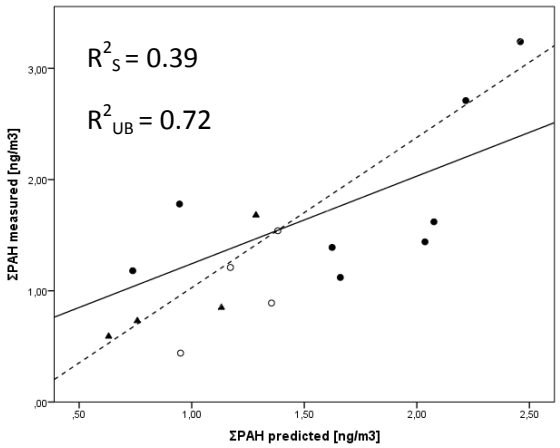
Oslo



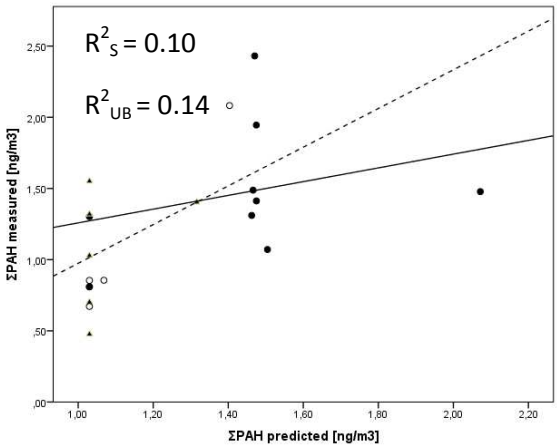
London/Oxford



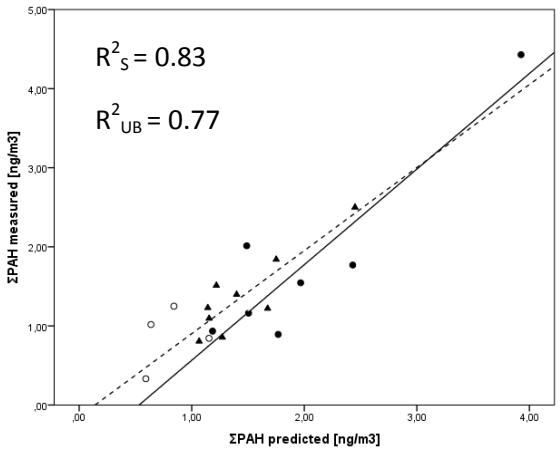
Netherlands



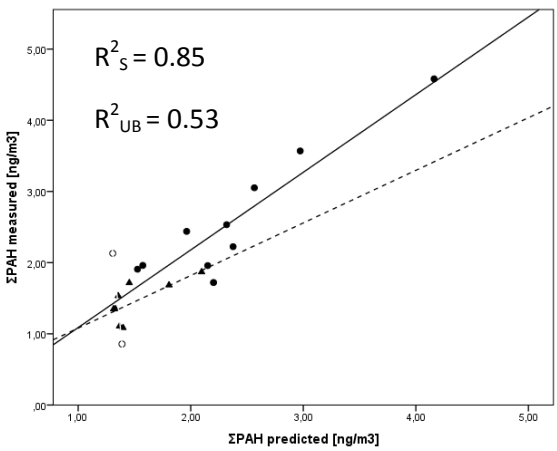
Munich/Augsburg



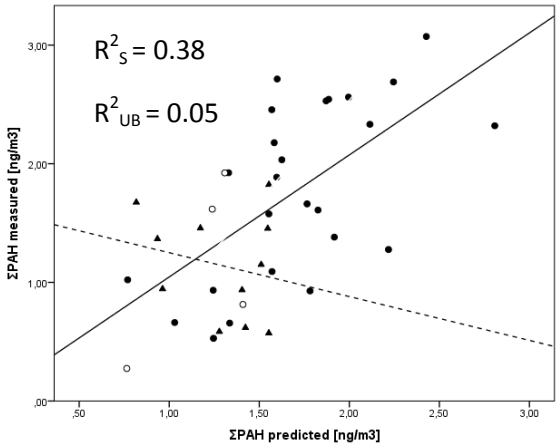
Paris



Rome



Catalonia



Athens

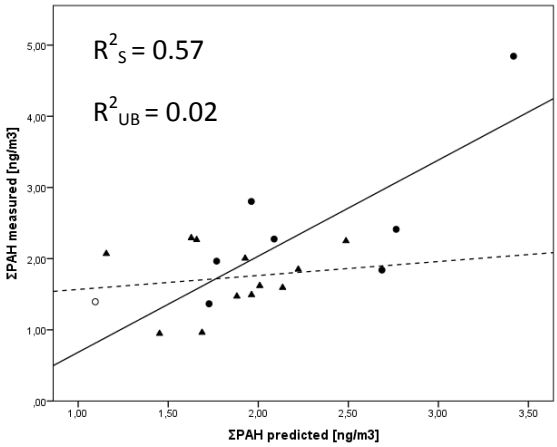
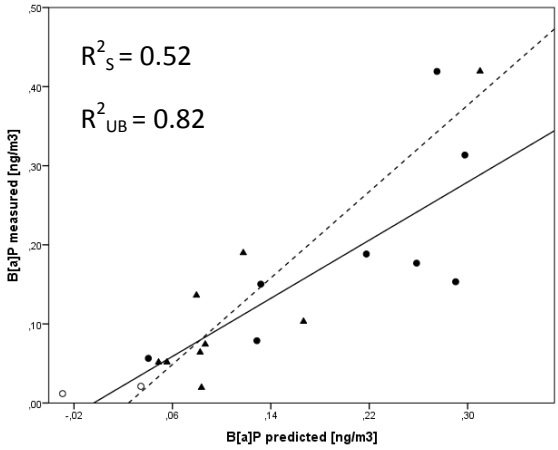
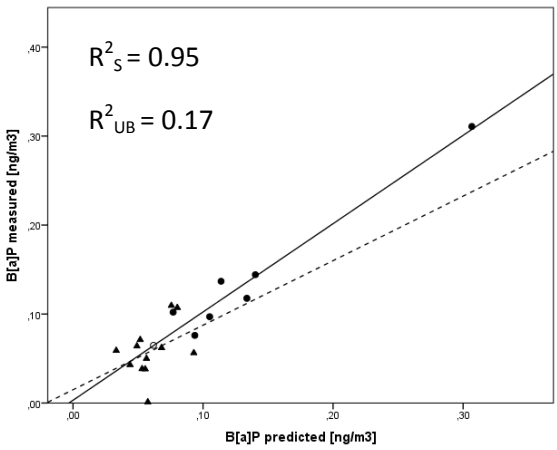


Figure S4. Correlation between measured and predicted ΣPAH annual averages for S - ●, UB - ▲, RB - ○. Solid line – S, dotted line - UB

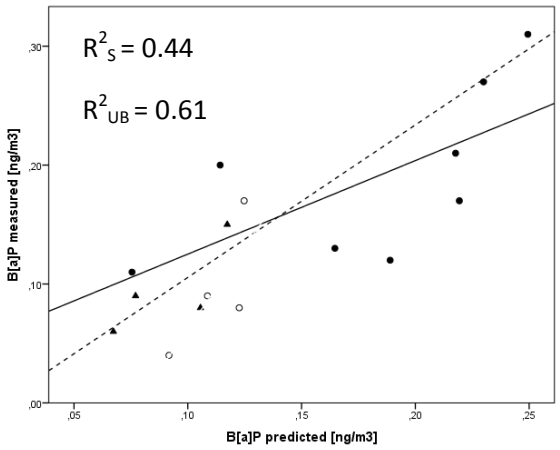
Oslo



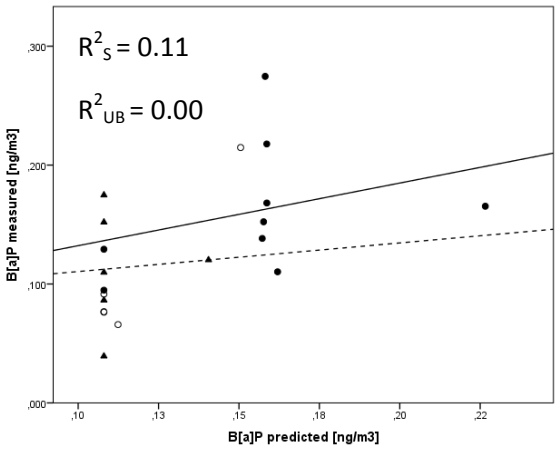
London/Oxford



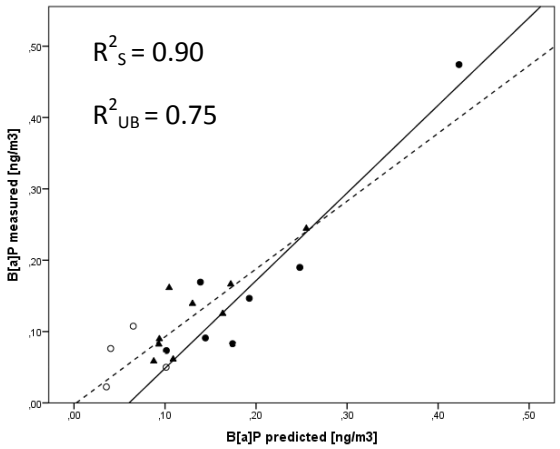
Netherlands



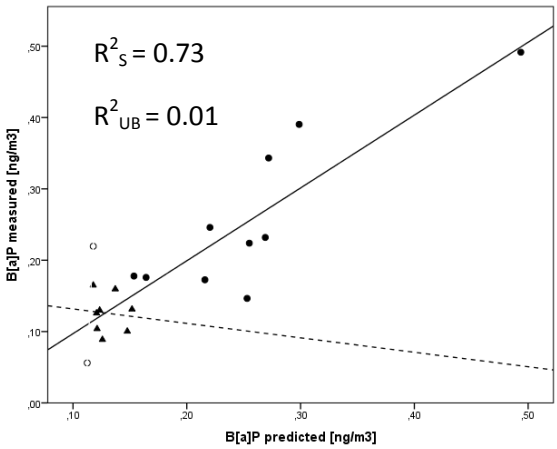
Munich/Augsburg



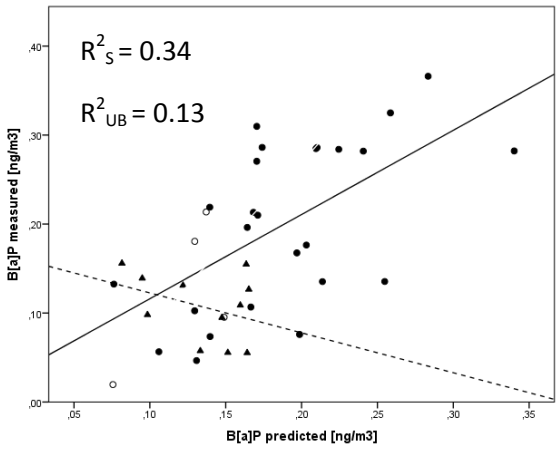
Paris



Rome



Catalonia



Athens

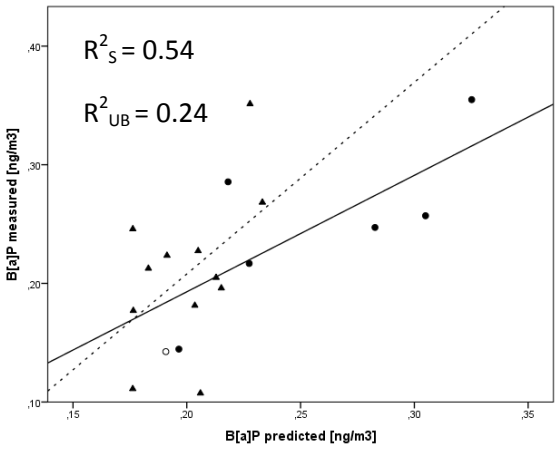
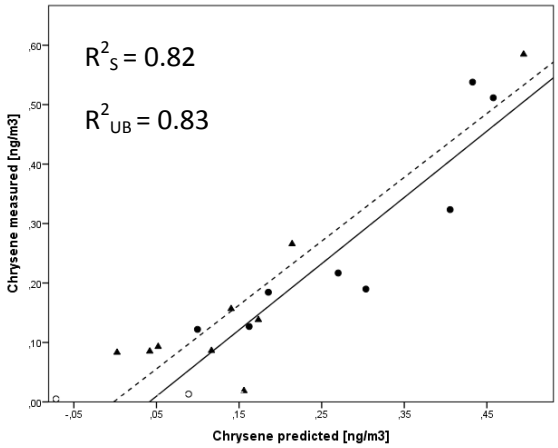
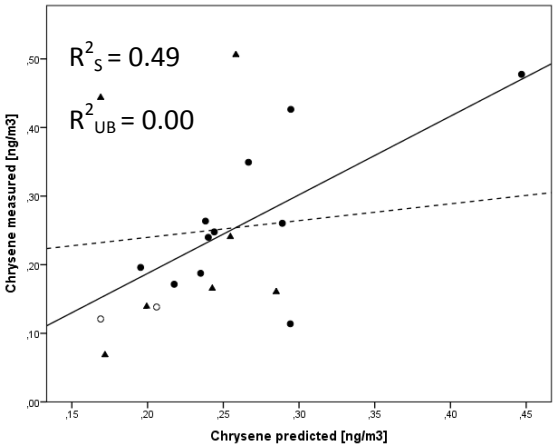


Figure S5. Correlation between measured and predicted B[a]P annual averages for S - ●, UB - ▲, RB - ○. Solid line – S, dotted line - UB

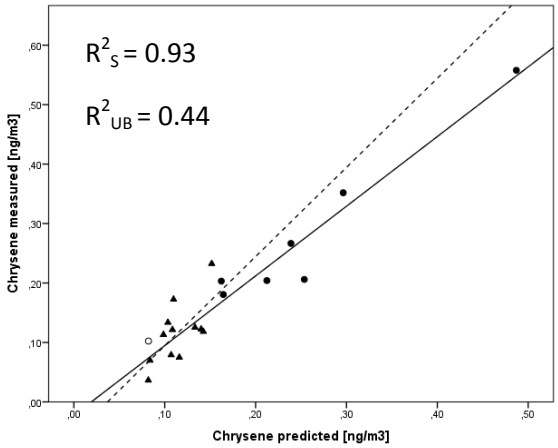
Oslo



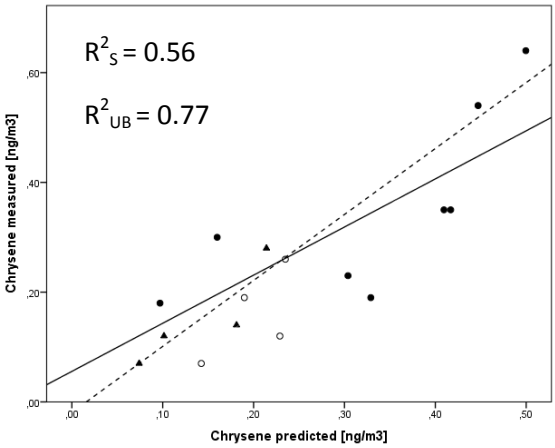
Helsinki/Turku



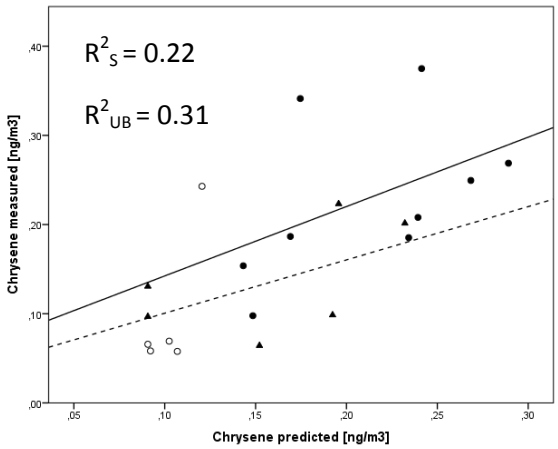
London/Oxford



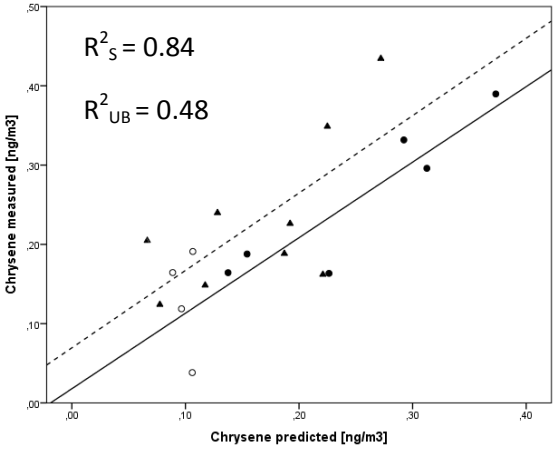
Netherlands



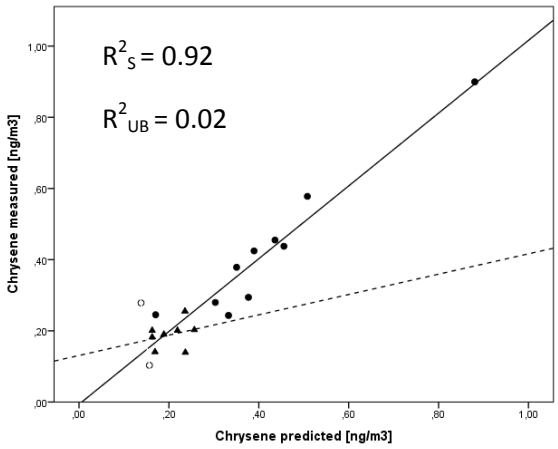
Munich/Augsburg



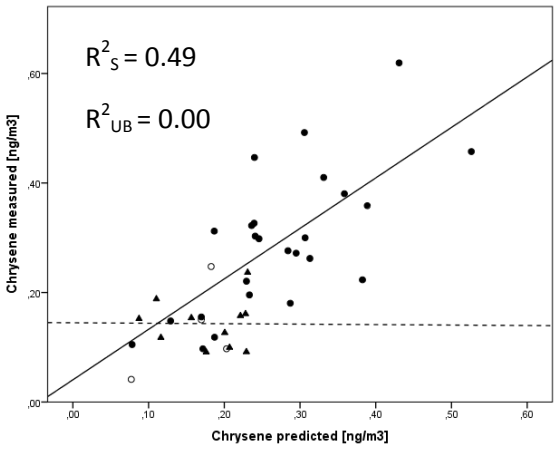
Paris



Rome



Catalonia



Athens

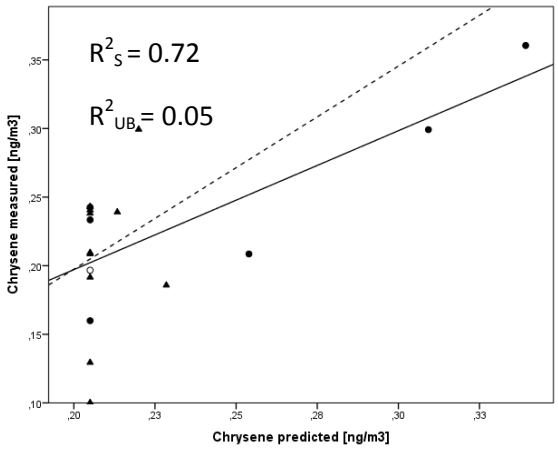
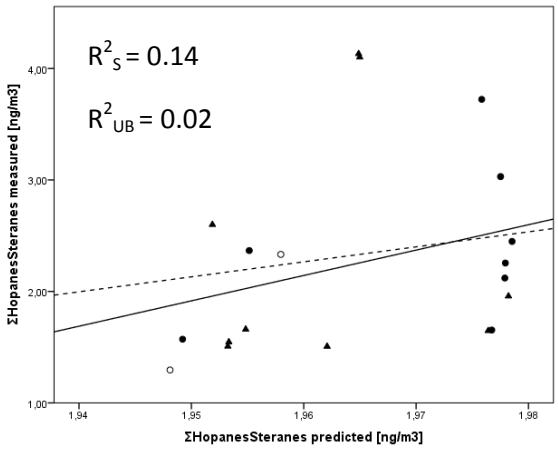
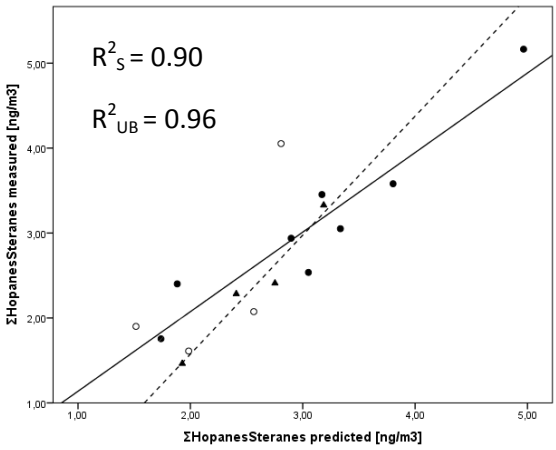


Figure S6. Correlation between measured and predicted Chrysene annual averages for S - ●, UB - ▲, RB - ○. Solid line – S, dotted line – UB

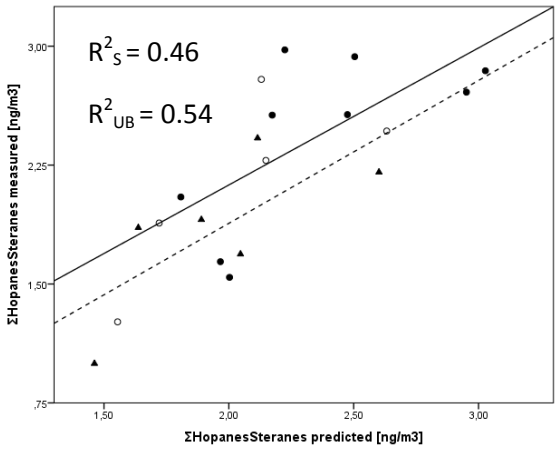
Oslo



Netherlands



Munich/Augsburg



Catalonia

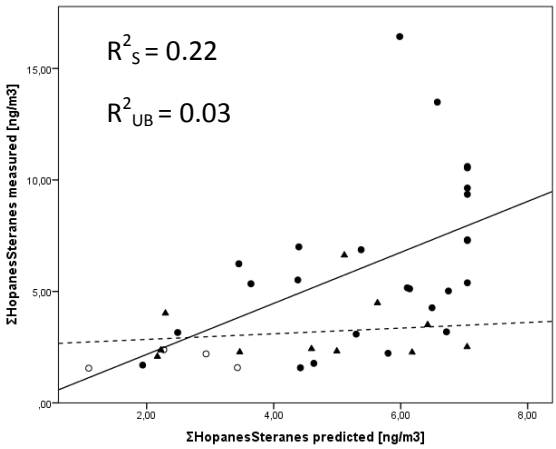


Figure S7. Correlation between measured and predicted ΣHopanesSteranes annual averages for S - ●, UB - ▲, RB - ○. Solid line – S, dotted line - UB