Exceedance modelling in the context of the data fusion mapping under ETC/ACM

Jan Horálek (CHMI)

Peter de Smet, Frank de Leeuw (RIVM), Pavel Kurfürst, Nina Benešová (CHMI), Philipp Schneider (NILU)

European Topic Centre
on Air Pollution and
Climate Change Mitigation

CZECH
HYDROMETEOROLOGICAL
INSTITUTE

National Institute for Public Health
and the Environment

NILU

1. Mapping methodology

Estimation of area in exceedance
Estimation of population living in exceedance areas

2. Influence of grid resolution

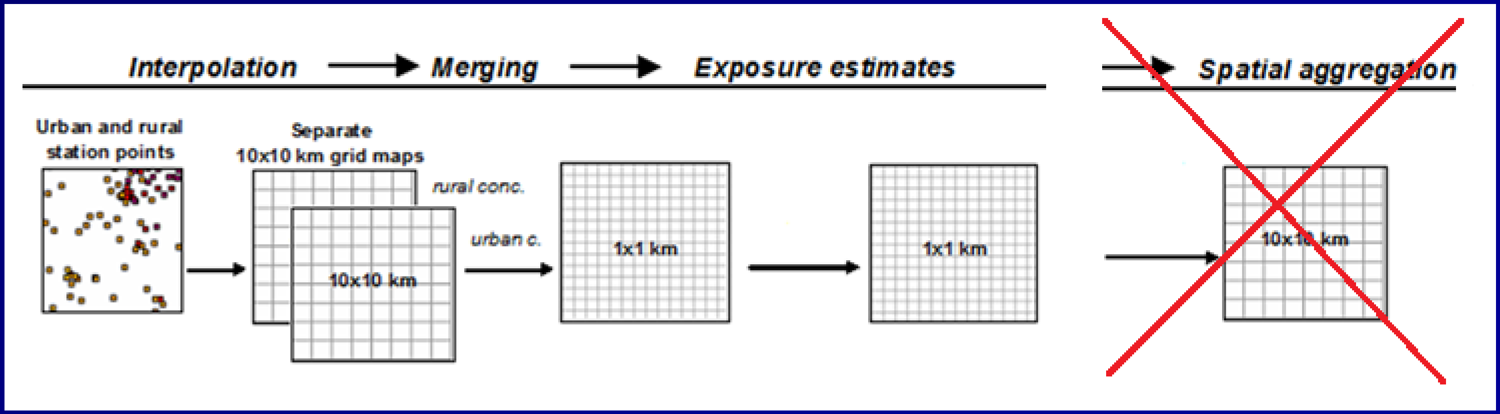
3. Case study: NO_2 improved mapping incl. traffic

4. Conclusion and recommendation

Data fusion mapping methodology

Regression – Interpolation – Merging Mapping

Linear regression model of monitoring data (as a primarily source of information), CTM output and different other supplementary data followed by ***interpolation of its residuals*** by kriging (so-called residual kriging).
Rural and urban background map layers created separately (based on rural resp. urban/suburban background stations) and ***merged*** into the final maps using population density.

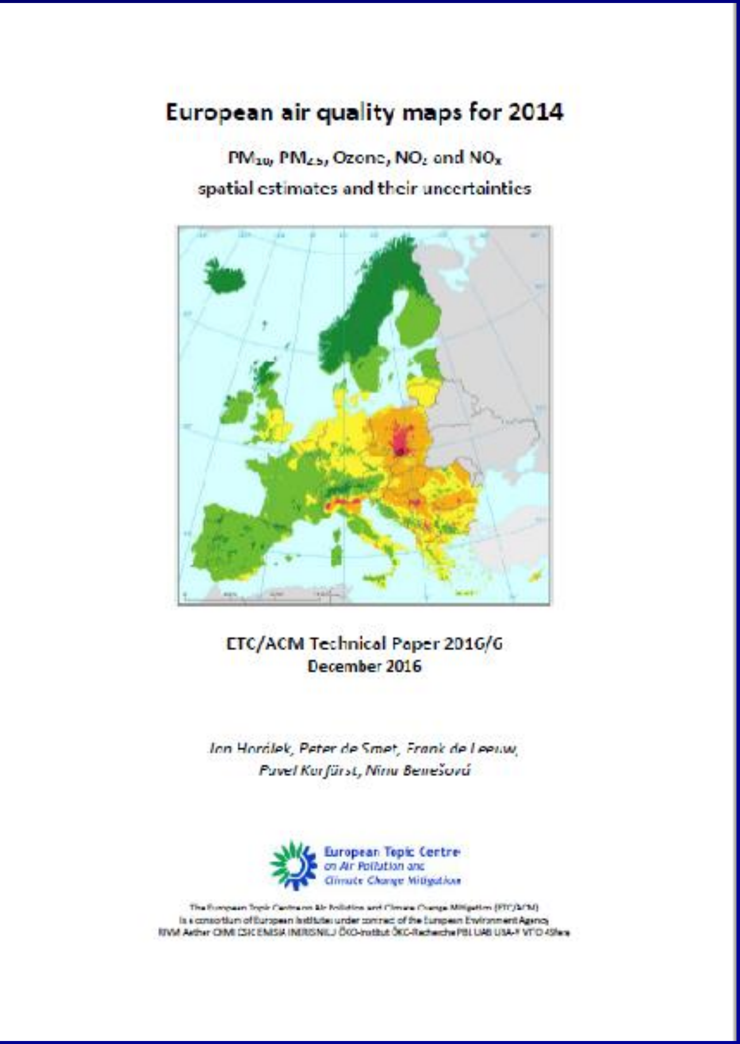


Annual air quality maps

Regular annual product:
ETC/ACM Technical Paper
„European air quality maps"

Concentration maps (PM_{10}, $PM_{2.5}$, O_3, NO_2, NO_x), probability of exceedance maps, exposure tables, uncertainty analysis, inter annual difference maps, trends.

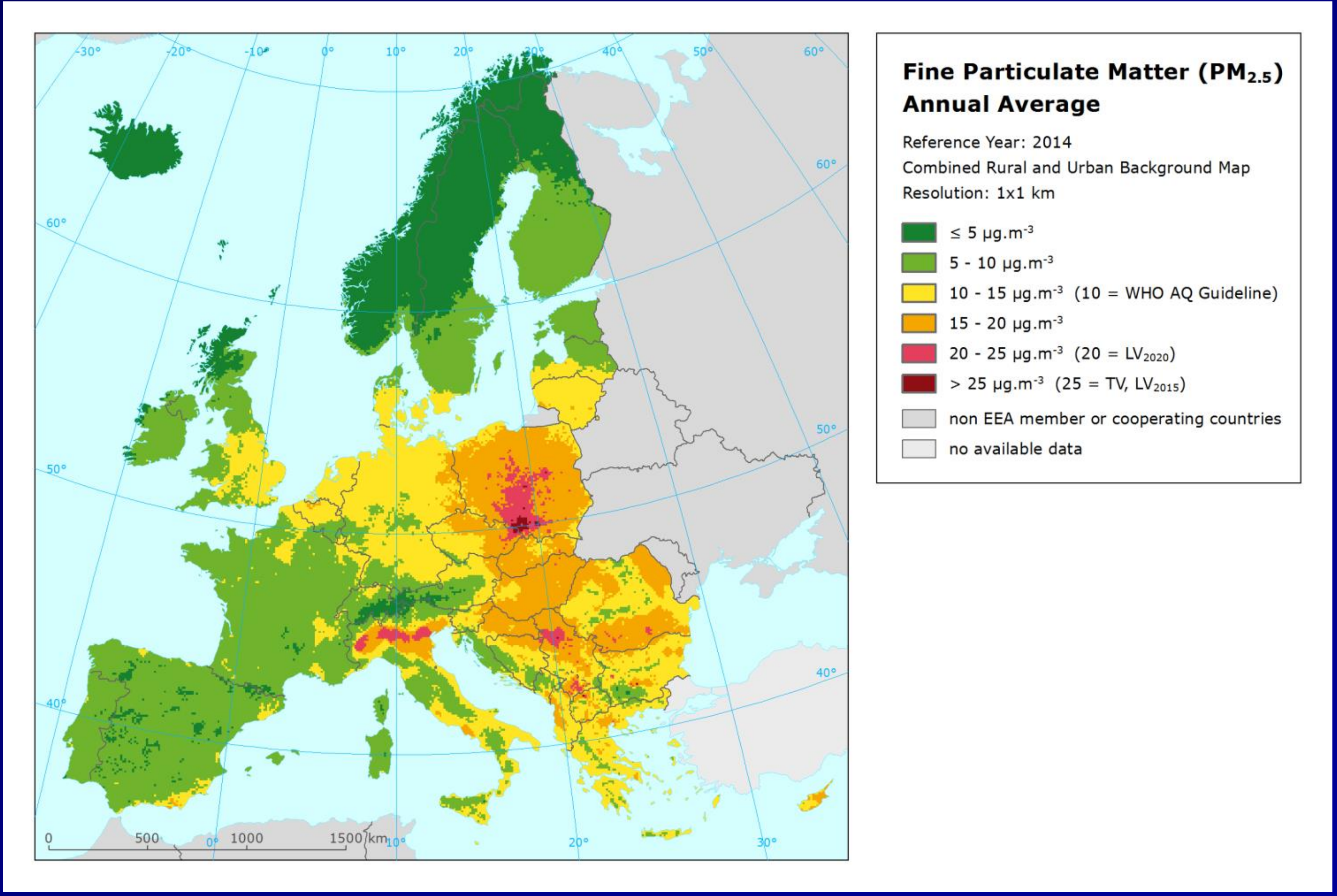
Most recent: ***ETC/ACM TP 2016/6***, maps and tables for 2014



http://acm.eionet.europa.eu/reports/

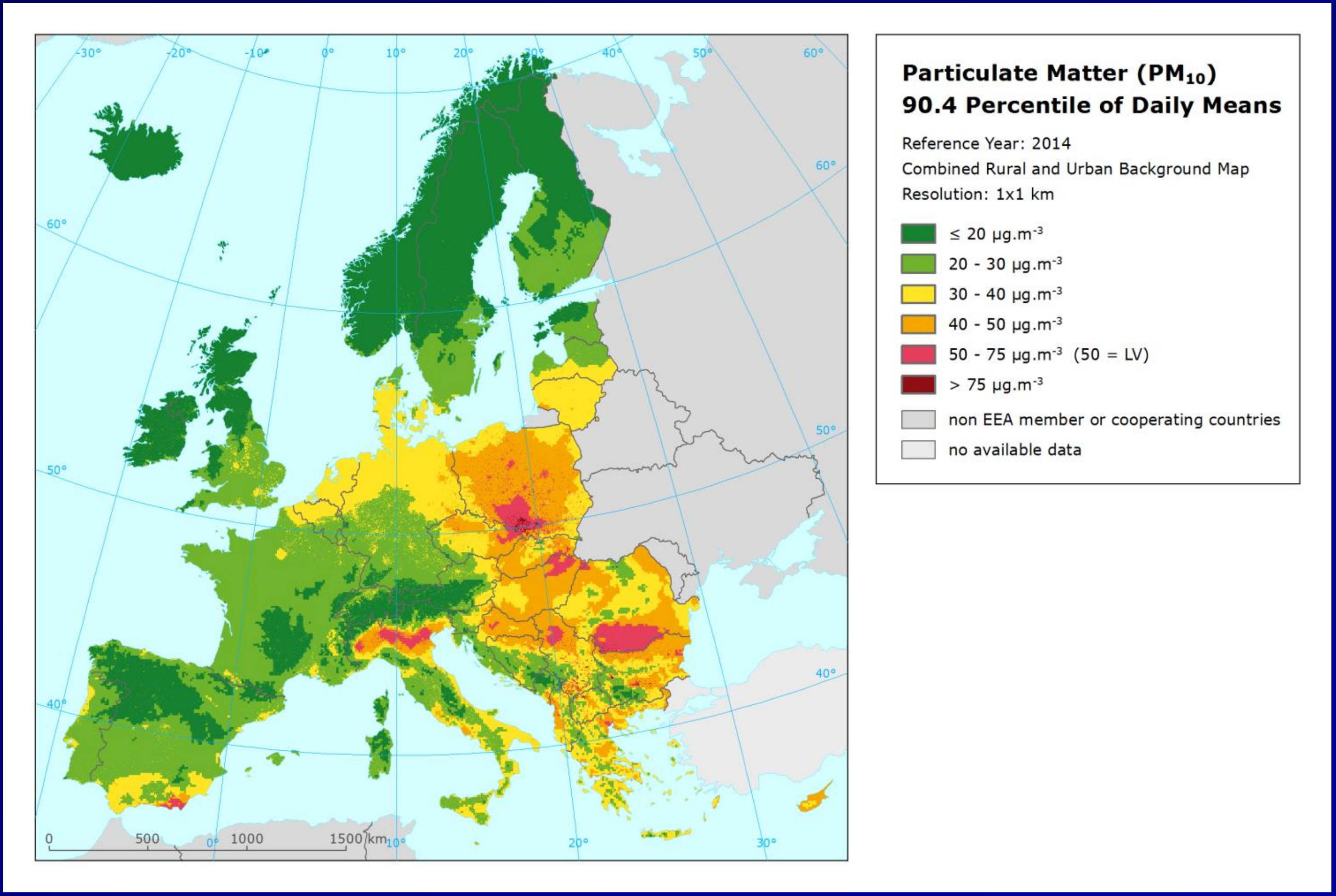
Area in exceedance – concentration maps

$PM_{2.5}$ – annual average, 2014



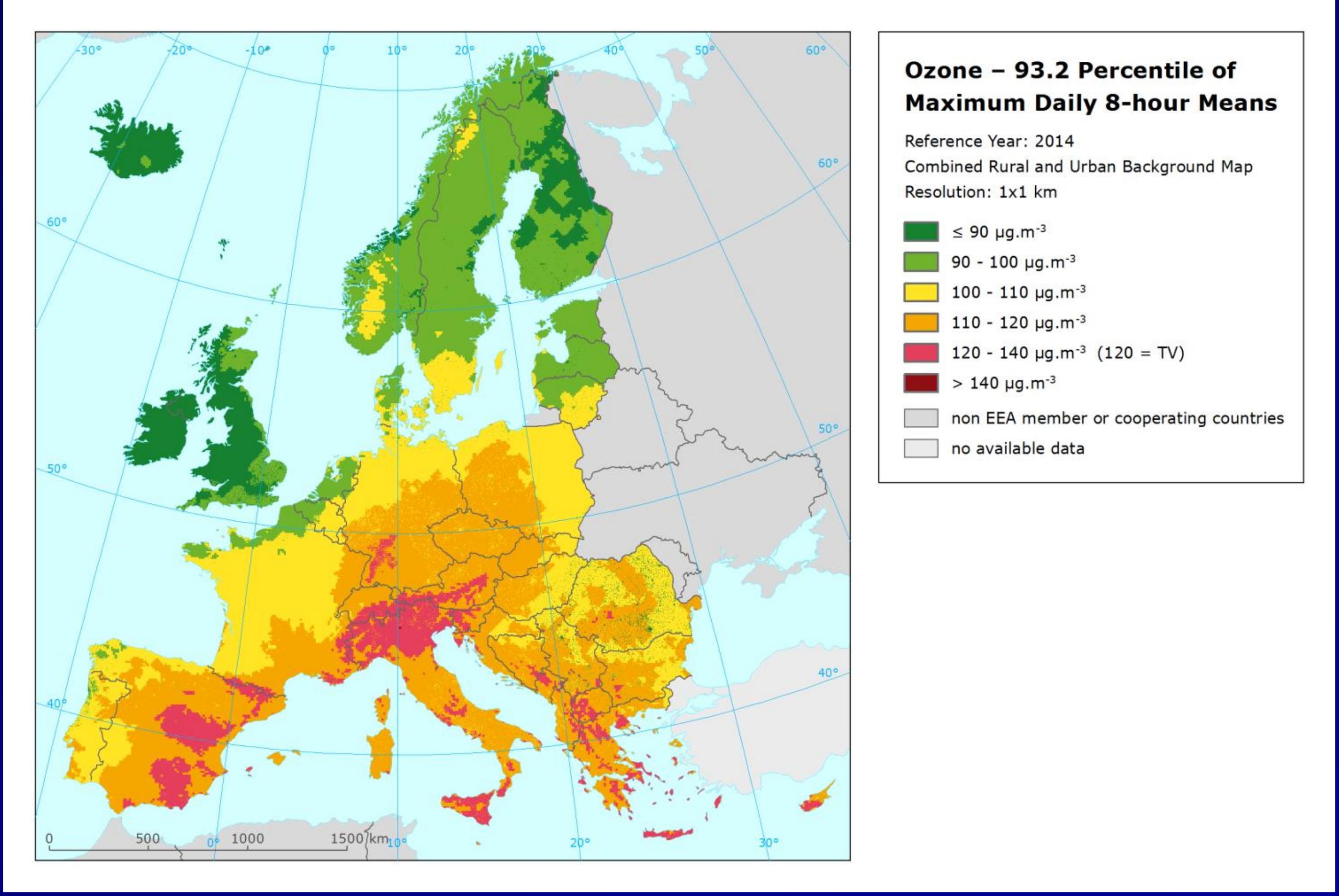
Area in exceedance – concentration maps

PM_{10} – 90.4 percentile of daily means, 2014



Area in exceedance – concentration maps

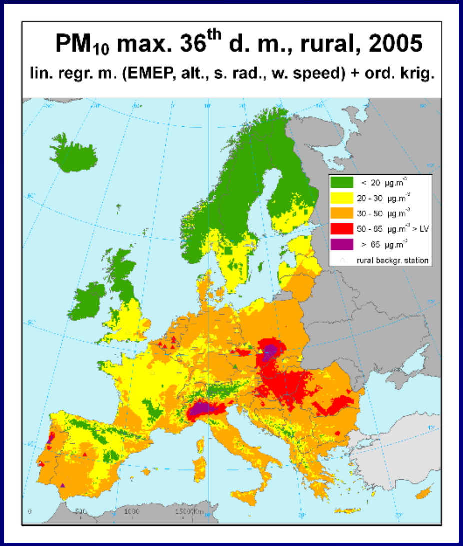
O_3 – 93.2 percentile of max. daily 8-hour means, 2014

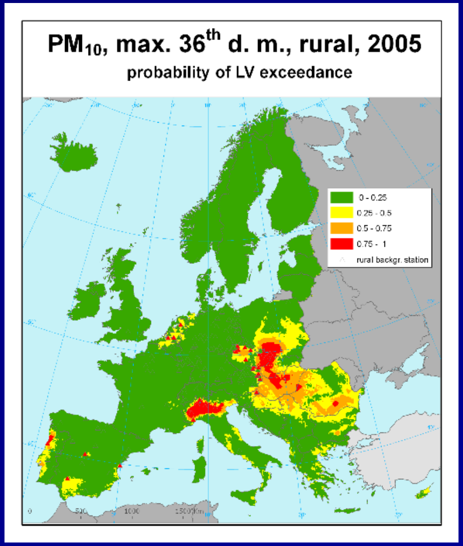


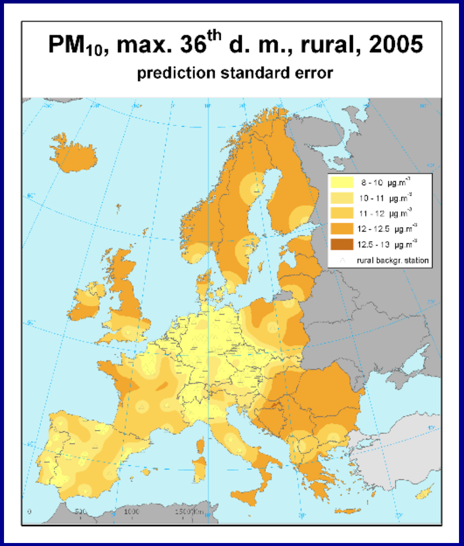
Area in exceedance – PoE maps

Probability of exceedance (PoE) maps

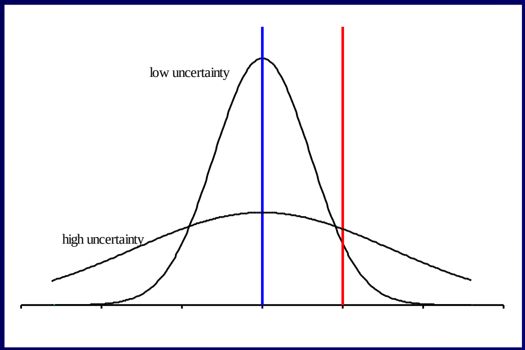
Based on ***concentration*** and ***uncertainty maps*** estimated based on geostatistic theory.







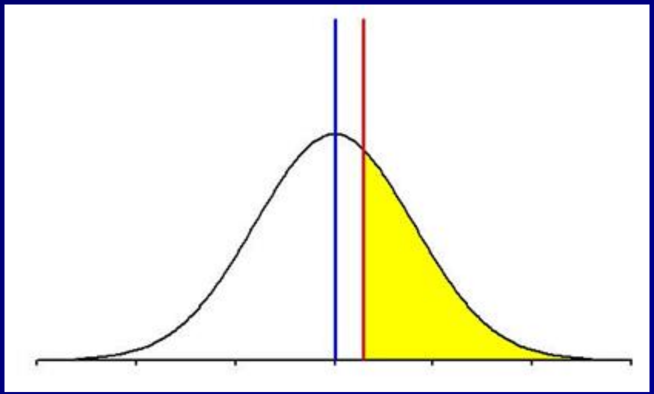
Area in exceedance – PoE maps

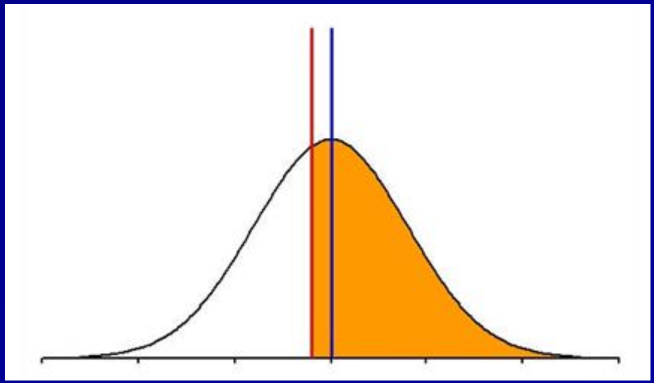


$$PoE(x) = 1 - \Phi\left(\frac{LV - C_c(x)}{\sigma_c(x)}\right)$$

$PoE(x)$	is the probability of limit/target value (LV/TV) exceedance in the grid cell x,
$\Phi()$	is the cumulative distribution function of the normal distribution,
LV	is the limit or target value of the relevant indicator,
$C_c(x)$	is the interpolated concentration in the grid cell x,
$\sigma_c(x)$	is the standard error of the estimation in the grid cell x.

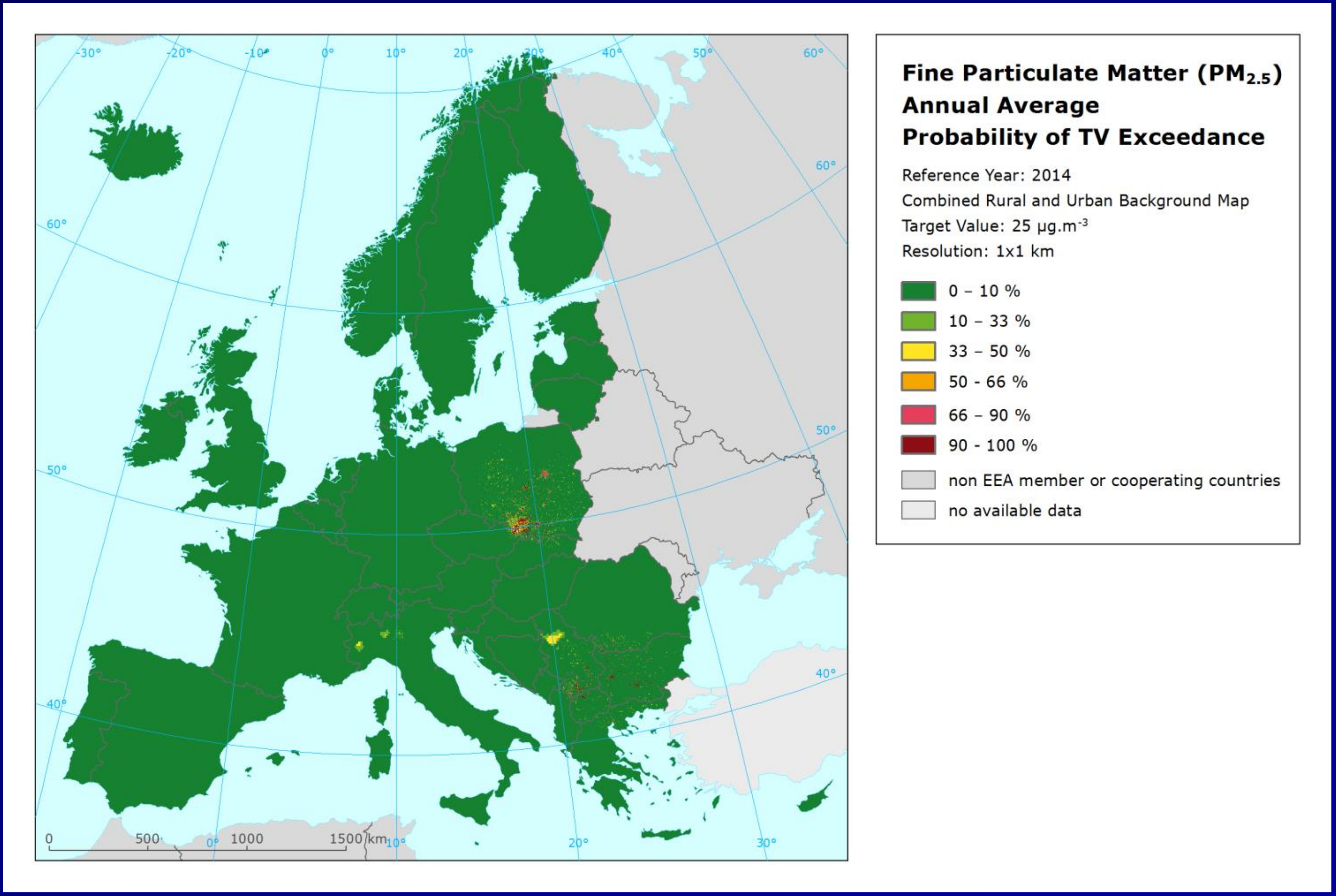
Map class colour	Percentage probability of threshold exceedance	Degree of probability (or likelihood) of exceedance	Likelihood of exceedance	
Green	0 – 10	Little	Very unlikely	More unlikely than likely
Light green	10 – 33	Low	Unlikely	
Yellow	33 – 50	Modest	About as likely as not	
Orange	50 – 66	Moderate		More likely than not
Light red	66 – 90	Large	Likely	
Dar red	90 – 100	High	Very likely	





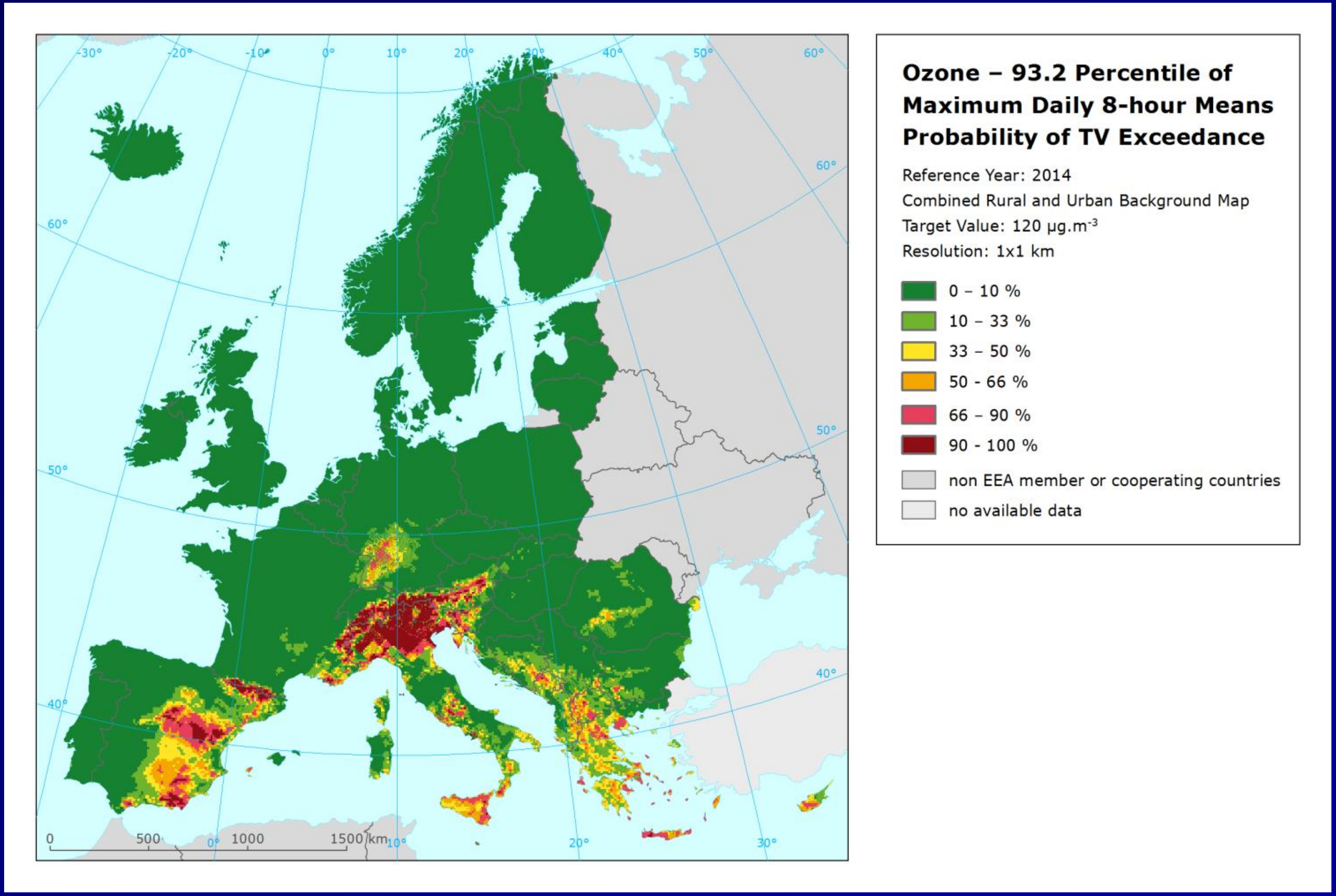
Area in exceedance – PoE maps

$PM_{2.5}$ – annual average, 2014



Area in exceedance – PoE maps

O_3 – 93.2 percentile of max. daily 8-hour means, 2014



Population living in exceedance areas

$PM_{2.5}$ annual average, 2014 exposure table

Country		Population	PM2.5 annual average, exposed population [%]						Population weighted conc.
			< LV2020				> LV2020		
			< TV					> TV	
		[inhbs . 1000]	< 5 µg.m-3	5 - 10 µg.m-3	10 - 15 µg.m-3	15 - 20 µg.m-3	20 - 25 µg.m-3	> 25 µg.m-3	[µg.m-3]
Albania	AL	2 896		0.0	24.2	72.8	2.2	0.8	16.5
Andorra	AD	73	0.5	2.7	96.7				10.0
Austria	AT	8 507	0.4	13.0	54.0	32.5			12.9
Belgium	BE	11 204		2.7	67.2	30.0			13.7
Bosnia & Herzegovina	BA	3 831		1.7	47.9	47.4	3.0		15.3
Bulgaria	BG	7 246		0.9	8.4	12.0	33.4	45.2	24.0
Croatia	HR	4 247		4.8	31.8	63.1	0.3		15.6
Cyprus	CY	858			5.0	95.0			17.0
Czech Republic	CZ	10 512		0.2	9.5	69.3	13.8	7.2	18.7
Denmark	DK	5 627	0.4	4.2	95.4				11.6
Estonia	EE	1 316		94.9	5.1				8.7
Finland	FI	5 451	1.1	98.4	0.5				7.4
France (metropolitan)	FR	63 989	0.0	31.5	67.4	1.1			11.0
Germany	DE	80 767	0.0	2.3	85.5	12.3			13.4
Greece	GR	10 927		0.4	29.1	50.6	11.9	8.0	17.0
Spain (excl. Canarias)	ES	44 397	0.1	43.6	55.1	1.2	0.0		10.7
Sweden	SE	9 645	4.0	82.0	13.9				7.6
Switzerland	CH	8 140	0.8	16.6	81.1	1.6			11.6
United Kingdom (& dep.)	UK	64 351	0.2	13.9	86.0				11.6
Total		**532 738**	**0.4**	**16.4**	**52.8**	**18.6**	**7.7**	**4.2**	**14.1**
			16.7		**71.4**		**11.9**		
EU-28		**502 424**	**0.1**	**16.3**	**54.2**	**18.3**	**7.3**	**3.7**	**14.0**
			16.4		**72.5**		**11.0**		

1. Mapping methodology

Estimation of area in exceedance
Estimation of population living in exceedance areas

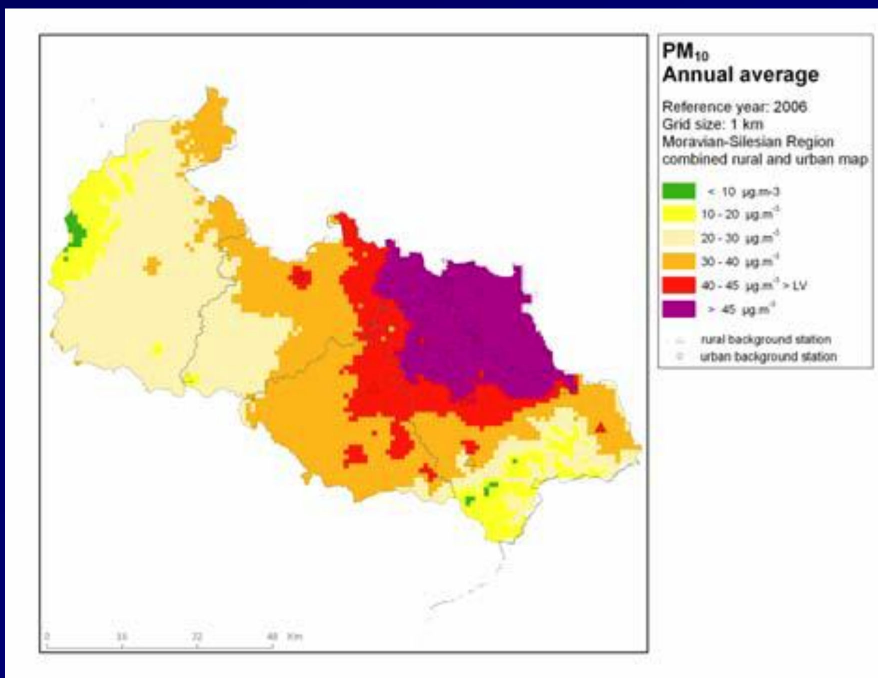
2. Influence of grid resolution

3. Case study: NO_2 improved mapping incl. traffic

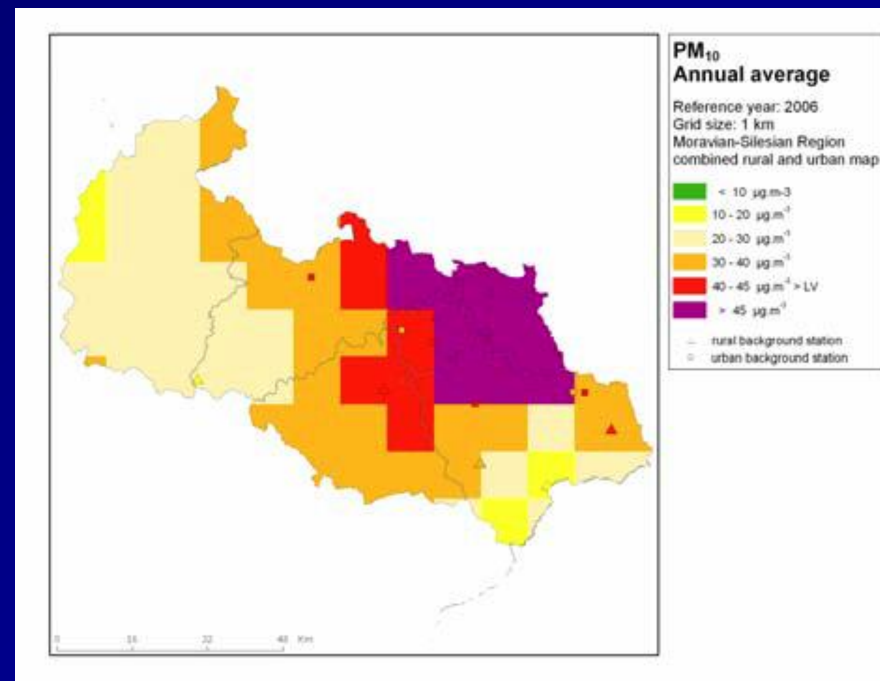
4. Conclusion and recommendation

Influence of grid resolution – concentration map

PM_{10} – annual average, 2006, Moravian-Silesian Region, CZ



1x1 km

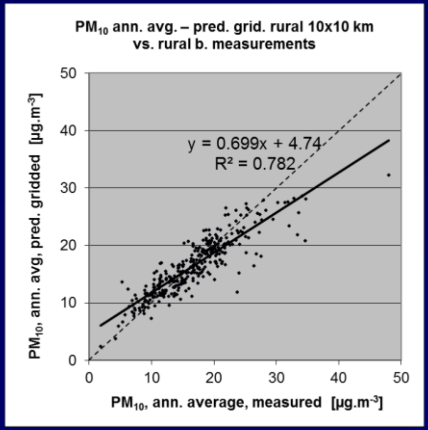


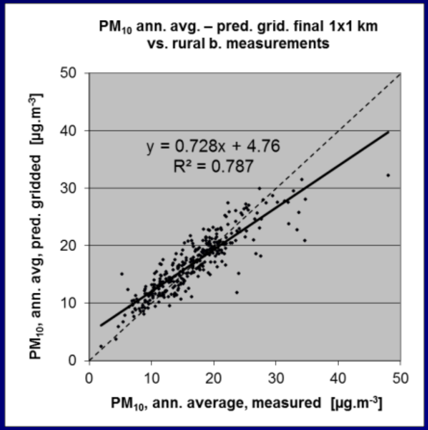
10x10 km (spatially aggregated)

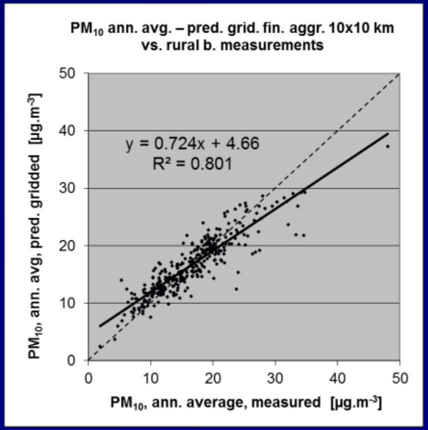
Influence of grid resolution – concentration map

PM_{10}, annual average, 2014

simple comparison – rural areas







Good representation in both 1x1 km and 10x10 km maps.

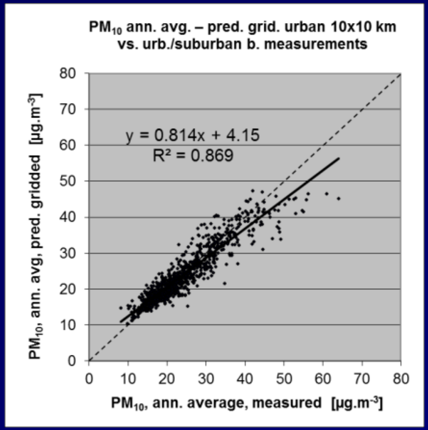
rural 10x10 ***final merged 1x1*** ***final, aggr. 10x10***

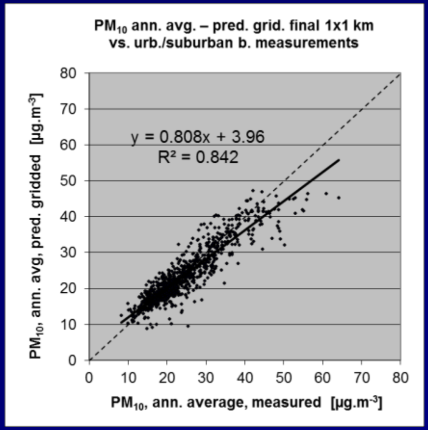
PM_{10}	rural backgr. stations			
	RMSE	bias	R^2	lin. r. equation
Annual average				
cross-valid. prediction, separate (r or ub) map	3.5	0.1	0.682	y = 0.666x + 5.71
grid prediction, 10x10 km separate (r or ub) map	2.9	-0.3	0.782	y = 0.699x + 4.74
grid prediction, 1x1 km final combined map	2.9	0.2	0.787	y = 0.728x + 4.76
grid prediction, aggr. 10x10 km final comb. map	2.8	0.0	0.801	y = 0.724x + 4.66

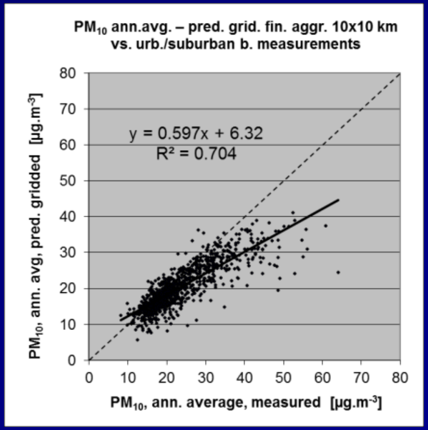
Influence of grid resolution – concentration map

PM_{10}, annual average, 2014

simple comparison – urban background areas







Good representation in1x1 km map, but not in 10x10 km map (bias, RMSE, R^2).

rural 10x10 *final merged 1x1* *final, aggr. 10x10*

PM_{10}	urban/suburban backgr. stations			
	RMSE	bias	R^2	lin r. equation
Annual average				
cross-valid. prediction, separate (r or ub) map	4.2	0.0	0.757	y = 0.761x + 5.60
grid prediction, 10x10 km separate (r or ub) map	3.1	-0.2	0.869	y = 0.814x + 4.15
grid prediction, 1x1 km final combined map	3.4	-0.6	0.842	y = 0.808x + 3.96
grid prediction, aggr. 10x10 km final comb. map	5.7	-3.2	0.704	y = 0.597x + 6.32

Influence of grid resolution – PoE map

The map shows the probability that the spatial average of the relevant grid (e.g. 10x10 km or 1x1 km) exceeds the limit value.

10x10 km resolution map gives a realistic result for rural background areas only. (Not for urban areas.)

1x1 km resolution map gives a realistic result for both rural and urban background areas. (Not for hotspots.)

Influence of grid resolution – exposure table

PM_{10}, annual average, 2006 – population weighted concentration

Country	PM_{10} – annual average			PM_{10} – 36th highest		
	C. 10 - 1 - 1	D. 1 - 1 - 1	F. aggr.D-10	C. 10 - 1 - 1	D. 1 - 1 - 1	F. aggr.D-10
Albania	31.8	31.8	27.5	54.0	53.5	46.1
Andorra	22.5	21.6	10.6	35.7	34.9	18.7
Austria	26.0	26.1	22.9	47.1	47.3	41.7
Belgium	31.3	31.2	30.1	51.3	51.2	50.0
Bosnia-Herzeg.	33.1	33.1	29.2	57.4	57.6	50.5
Bulgaria	41.6	41.8	31.9	74.2	74.5	55.2
Croatia	31.5	31.6	29.2	53.7	53.8	49.2
Cyprus	35.4	35.9	33.1	58.2	57.8	50.4
Czech Republic	33.5	33.5	30.6	57.5	57.7	53.2
Denmark	23.5	23.6	21.5	37.0	37.5	34.2
Estonia	19.7	19.7	17.4	34.1	34.4	30.6
Finland	17.0	16.6	16.0	29.5	29.5	28.4
France	20.4	20.3	19.6	32.9	32.7	31.7
Germany	24.2	24.3	22.6	41.3	41.5	38.8
Greece	33.6	33.6	29.8	54.3	54.0	47.8
Slovenia	29.0	29.2	25.9	49.2	49.4	43.4
Spain	31.4	31.6	24.2	49.3	49.6	39.3
Sweden	19.0	18.8	16.7	32.0	32.2	28.2
Switzerland	23.2	23.2	20.6	43.9	43.8	39.3
United Kingdom	23.2	23.3	20.8	35.5	35.7	32.1
Total	**28.5**	**28.6**	**25.5**	**47.8**	**47.9**	**42.9**

1. Mapping methodology

Estimation of area in exceedance
Estimation of population living in exceedance areas

2. Influence of grid resolution

3. Case study: NO_2 improved mapping incl. traffic

4. Conclusion and recommendation

Case study: Improved NO_2 mapping

Detailed analysis presented in ***ETC/ACM Technical Paper 2016/12 „Inclusion of LC and traffic data in NO_2 mapping“***

Inclusion of land cover and road data in NO_2 background mapping.

Creation of traffic layer map, based on traffic stations. Subsequently, inclusion of this traffic map layer in NO_2 map and exposure estimate.

NO_2 annual average, 2013 – rural and urban background map layers improvement by land cover data inclusion

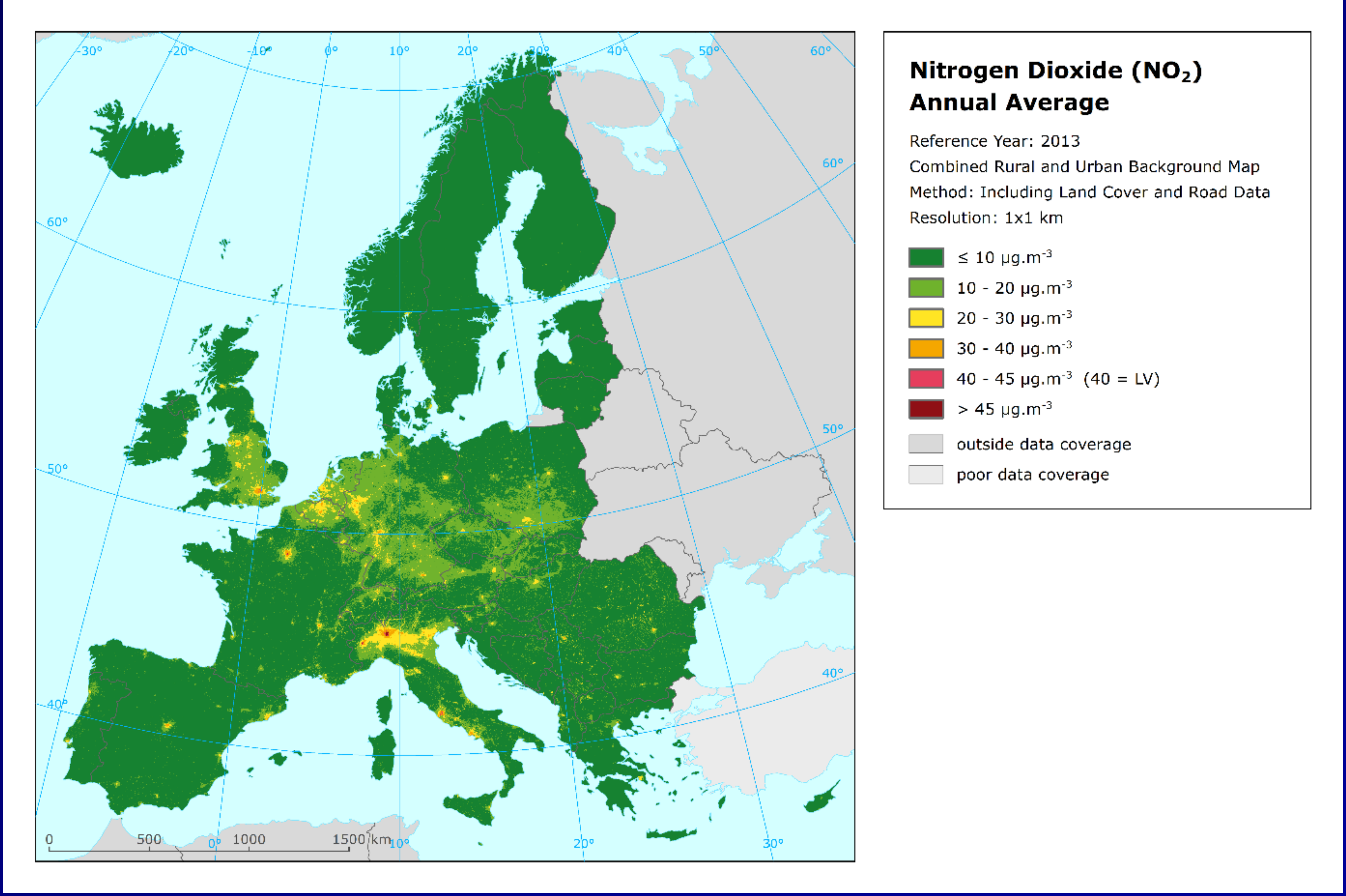
Inclusion of LC in LRM: 8 general CLC classes, with different radius. Together with other data – input to stepwise regression. For best variants: ***spatial interpolation*** of residuals, ***1x1 km resolution.*** Different variants compared by cross-validation.

	spatial interpolation variant + supplementary data used	rural areas				
		RMSE	RRMSE	bias	R^2	regr. eq.
(i)	current (EMEP, wind speed; 10x10 km)	3.4	35.8%	0.1	0.699	y = 0.677x + 3.2
(ii)	current 1km (EMEP, wind speed; 1x1 km)	3.4	35.2%	0.1	0.681	y = 0.694x + 2.0
(iii)	impr. current 1km (EMEP, altitude, wind speed; 1x1 km)	3.2	33.8%	0.1	0.706	y = 0.725x + 2.7
(iv)	including LC (EMEP, altitude, w. sp., land cover; 1x1 km)	2.8	29.2%	0.1	0.782	y = 0.810x + 1.9
(v)	without CTM (alt., w.sp., s.s. rad., temp., pop., LC; 1x1 km)	3.3	34.6%	0.2	0.698	y = 0.760x + 2.5

	spatial interpolation variant + supplementary data used	urban background areas				
		RMSE	RRMSE	bias	R^2	regr. eq.
(i)	current (EMEP, wind speed; 10x10 km)	5.1	23.9%	0.0	0.557	y = 0.572x + 9.2
(ii)	current 1km (EMEP, wind speed; 1x1 km)	5.1	23.6%	0.0	0.568	y = 0.586x + 8.9
(iii)	impr. current 1km (EMEP, wind speed, population; 1x1 km)	4.8	22.7%	0.0	0.603	y = 0.624x + 8.1
(iv)	including LC (EMEP, alt., w. sp., pop., road, LC; 1x1 km)	4.6	21.3%	0.0	0.645	y = 0.670x + 7.1
(v)	without CTM (alt., w.sp., pop., road, land cover; 1x1 km)	4.7	22.1%	0.0	0.624	y = 0.659x + 7.4

Inclusion of land cover brings clear map improvement.

NO_2 annual avg., 2013 – rural and urban background map



NO_2 annual average, 2013 – traffic map layer creation

Based on 855 urban traffic stations. (Rural traffic stations not considered, due to their small number, i.e. 19). Supplementary data – 37 variables (EMEP model, altitude, meteo, LC, …): input to ***linear regression model*** analysis (stepwise regression + backwards elimination). 4 variants selected for further analysis.

linear regr. model + OK of its residuals	(ii) current 1k urban traffic coeff.	(iii) altern. curr. urban traffic coeff.	(iv) incl. LC urban traffic coeff.	(v) without CTM urban traffic coeff.
c (constant)	42.58	36.16	27.39	-109.71
a1 (EMEP model)	0.654	0.685	0.550	
a2 (GTOPO_1km)		-0.0172	-0.0254	-0.0204
a3 (GTOPO_5km_rad)		0.0168	0.0213	0.0176
a4 (wind speed)	-1.278		-1.529	1.628
a5 (s. solar radiation)	-1.100	-1.019		-3.414
a6 (temperature)				-1.550
a7 (relative humidity)				1.206
a8 (population_5km_rad)				0.000017
a8 (T2buf75m_1km)			7.263	
a10 (LDR_5km_rad)			0.0208	0.0248
adjusted R^2	**0.33**	**0.33**	**0.36**	**0.31**
st. err. [$\mu g.m^{-3}$]	**10.80**	**10.80**	**10.53**	**10.93**

NO_2 annual average, 2013 – traffic map layer creation

Spatial interpolation, 1x1 km resolution

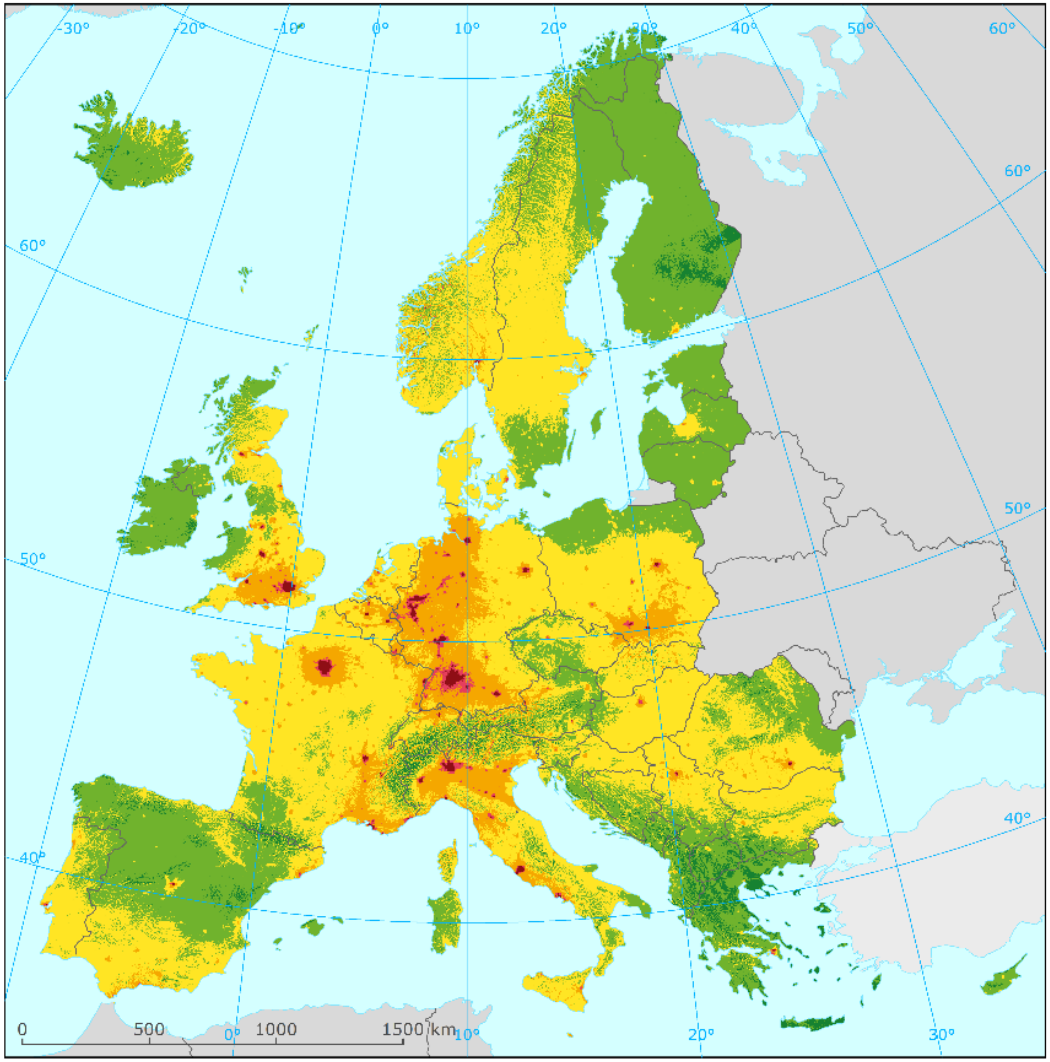
Different LRM variants, comparison based on cross-validation.

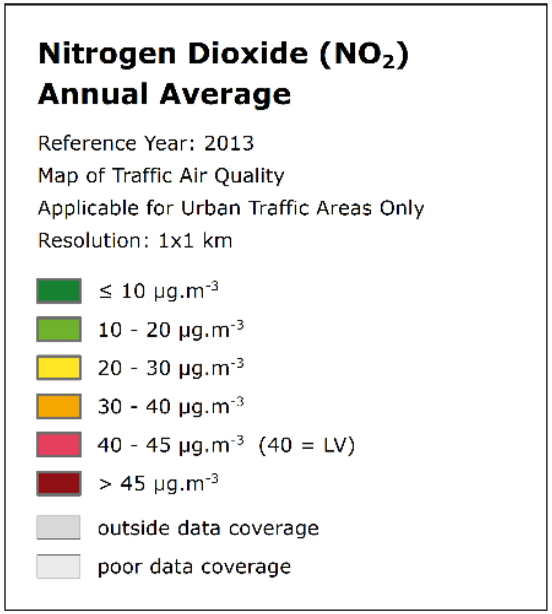
	spatial interpolation variant + supplementary data used	urban traffic areas				
		RMSE	RRMSE	bias	R^2	regr. eq.
(ii)	current 1km (EMEP, wind speed, s. solar radiation; 1x1 km)	9.5	25.1%	0.0	0.470	y = 0.485x + 19.6
(iii)	alternative current (EMEP, alt., s. solar rad.; 1x1 km)	9.5	25.1%	0.1	0.470	y = 0.483x + 19.7
(iv)	including LC (EMEP, alt., w. sp., road data, LC; 1x1 km)	9.2	24.3%	0.1	0.505	y = 0.529x + 17.9
(v)	without CTM (alt., w.sp., s.s.r., r.h., temp., pop., LC; 1x1 km)	9.2	24.3%	0.1	0.504	y = 0.534x + 17.7

Different variants give quite similar results.

Based on RRMSE (24%) and R^2 of cross-validation scatterplot (0.51): estimates of urban traffic air quality is reasonable.

NO_2 annual average, 2013 – urban traffic map layer





Map can be applied for urban traffic areas only.

$\hat{Z}_U(s_0) = (1 - w_T(s_0)) \cdot \hat{Z}_{UB}(s_0) + w_T(s_0) \cdot \hat{Z}_T(s_0)$

Case study: Improved NO_2 mapping - continuation

Inclusion of traffic map layer in background map

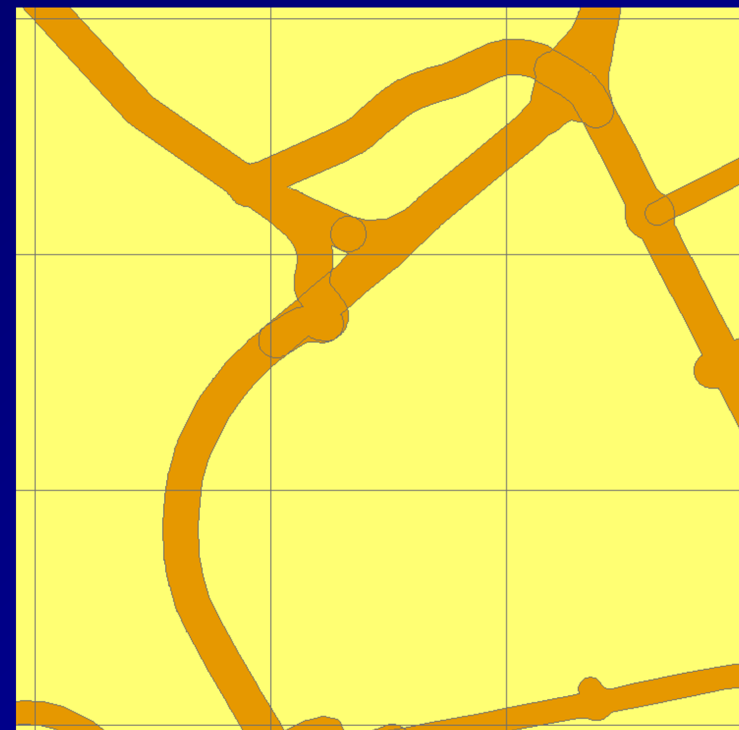
Inclusion in urban map layer:

$$\hat{Z}_U(s_0) = (1 - w_T(s_0)) \cdot \hat{Z}_{UB}(s_0) + w_T(s_0) \cdot \hat{Z}_T(s_0)$$

Weight: based on buffer around the streets/roads (GRIP database, source: PBL).

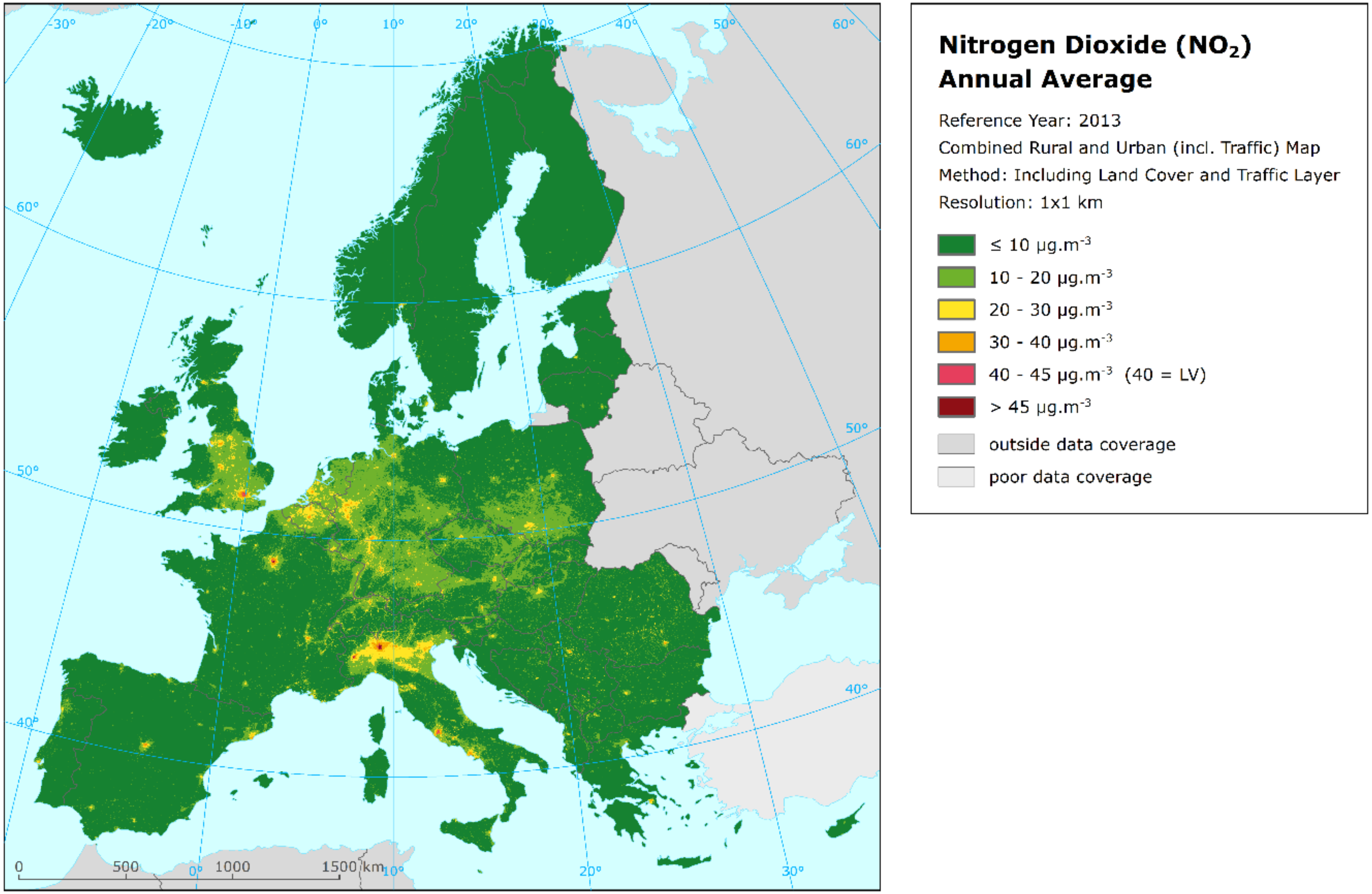
Buffer: tentative. 75 meters around roads of class 1 and 2, 50 meters around roads of class 3. (Should be refined.)

$$w_T(i) = T123buf\_1km(i) \ / \ 2$$

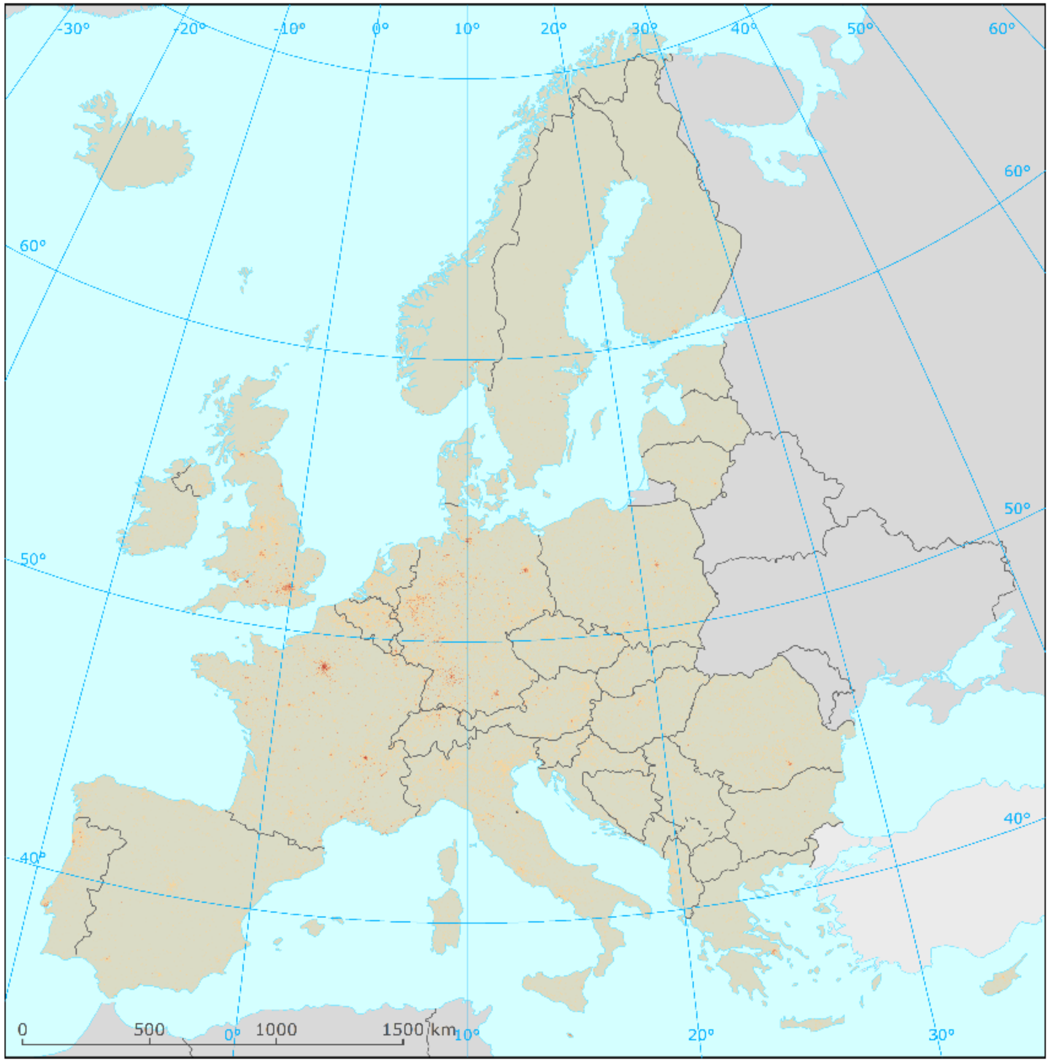


GRIP type class number	Class description
1	Highways
2	Primary roads
3	Secondary roads
4	Tertiary roads
5	Local, residential, urban roads

NO_2 annual average, 2013 – final merged map, 1x1 km



NO_2 annual average, 2013 – difference „final – background“



Nitrogen Dioxide (NO_2)
Annual Average
Difference Map

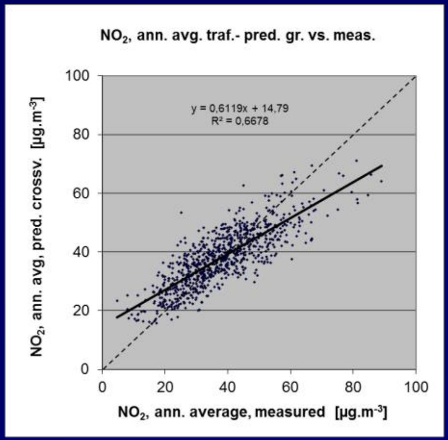
Reference Year: 2013
Difference: Method Including LC and Traffic Layer
- Method Including Land Cover
Map Resolution: 1x1 km

- ≤ - 5 $\mu g.m^{-3}$
- - 5 – - 2 $\mu g.m^{-3}$
- - 2 – 0 $\mu g.m^{-3}$
- 0 – 2 $\mu g.m^{-3}$
- 2 – 5 $\mu g.m^{-3}$
- > 5 $\mu g.m^{-3}$
- outside data coverage
- poor data coverage

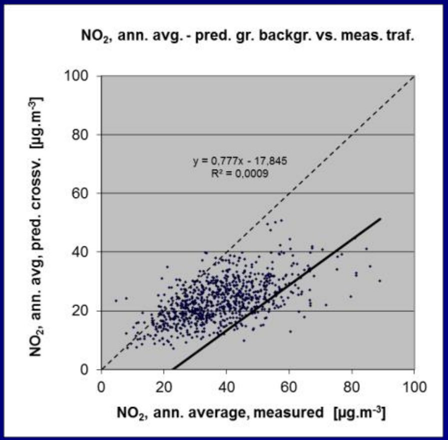
Case study: Improved NO_2 mapping – continuation

NO_2, annual average, 2013 – urban traffic areas

simple comparison – map layers vs. traffic stations



traffic map layer



background map layer

Bad representation of traffic areas in urban background layer.

NO_2	urban traffic stations			
	RMSE	bias	R^2	lin. r. equation
Annual average				
grid prediction, 1x1 km background map	7,6	-14,5	non significant	
grid prediction, 1x1 km urban traffic map layer	18,0	0,0	0,67	y = 0.612x + 14.8

Case study: Improved NO_2 mapping – continuation

Inclusion of traffic map layer in exposure estimate

$$\hat{c} = \frac{\sum_{i=1}^{N} c_{Bi}\left(1 - w_U(i) w_T(i)\right) p_i + \sum_{i=1}^{N} c_{Ti} w_U(i) w_T(i) p_i}{\sum_{i=1}^{N} p_i}$$

Weight: based on a buffer around the roads (GRIP database, source: PBL), similar as for map creation.

However: We go inside 1x1 km grid in the population exposure estimate.

Case study: Improved NO_2 mapping – continuation

NO_2 annual average 2013 – exposure estimate, comparison of estimate based on background (top) and final merged map (i.e. including traffic layer)

Country	NO_2 annual average, exposed population [%]						Population weighted conc. [$\mu g.m^{-3}$]
	< LV				> LV		
	< 10 $\mu g.m^{-3}$	10 - 20 $\mu g.m^{-3}$	20 - 30 $\mu g.m^{-3}$	30 - 40 $\mu g.m^{-3}$	40 - 45 $\mu g.m^{-3}$	> 45 $\mu g.m^{-3}$	
Total - based on background map	13,9	46,9	30,4	7,2	1,1	0,4	18,6
	98,5				1,5		
Total - based on final merged map, including traffic layer	13,7	44,8	29,8	8,5	1,6	1,6	19,4
	96,8				3,2		

Difference of population in exceedance: cc. 1.7% of European population.

1. Mapping methodology

Estimation of area in exceedance
Estimation of population living in exceedance areas

2. Influence of grid resolution

3. Case study: NO_2 improved mapping incl. traffic

4. Conclusion and recommendation

Conclusions

Grid resolution is of a great importance.

Rural map: 10x10 km resolution satisfactory.

Urban background map: 1x1 km resolution satisfactory.

Map taking in account traffic AQ: going inside 1x1 km resolution is needed.

Thank you for your attention.