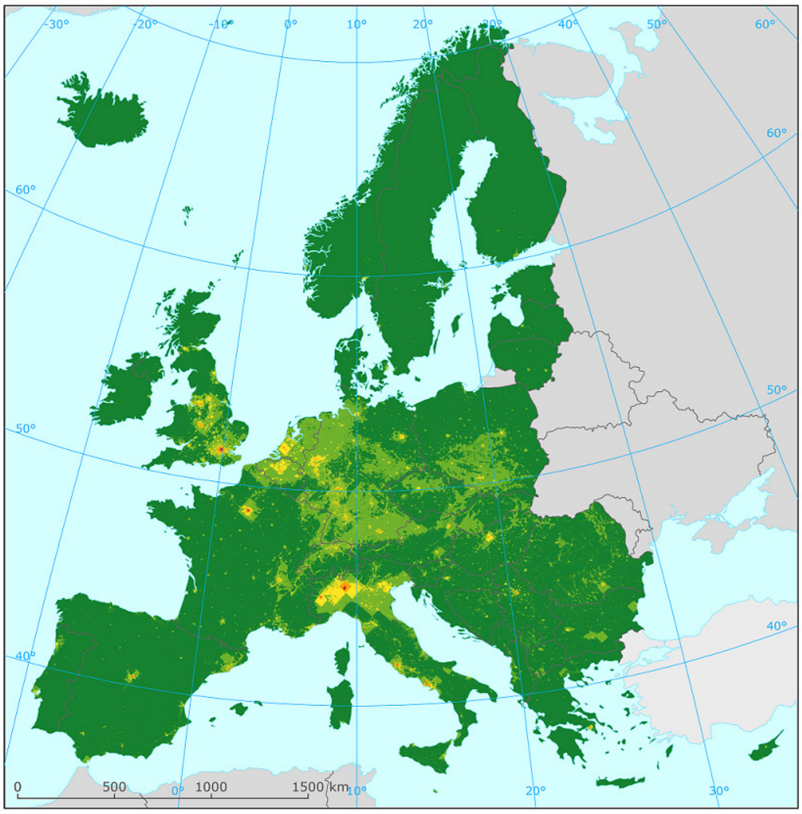
European NO_2 air quality map for 2014

Improved mapping methodology using land cover and traffic data



ETC/ACM Technical Paper 2017/6

October 2017

Jan Horálek, Peter de Smet, Frank de Leeuw, Pavel Kurfürst

European Topic Centre on Air Pollution and Climate Change Mitigation

The European Topic Centre on Air Pollution and Climate Change Mitigation (ETC/ACM)
is a consortium of European institutes under contract of the European Environment Agency
RIVM Aether CHMI CSIC EMISIA INERIS NILU ÖKO-Institut ÖKO-Recherche PBL UAB UBA-V VITO 4Sfera

Front page picture:
Concentration map of NO_2 annual average for the year 2014. Spatially interpolated concentration field. Improved mapping methodology using land cover and traffic data. Units: µg.m^{-3}. (Map 2.1 of this paper.)

Author affiliation:
Jan Horálek, Pavel Kurfürst: Czech Hydrometeorological Institute (CHMI), Prague, Czech Republic
Peter de Smet, Frank de Leeuw: National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

DISCLAIMER

ETC/ACM Technical Paper 2017/6
European Topic Centre on Air Pollution and Climate Change Mitigation
PO Box 1
3720 BA Bilthoven
The Netherlands
Phone +31 30 2748562
Fax +31 30 2744433
Email etcacm@rivm.nl
Website http://acm.eionet.europa.eu/

Contents

1 Introduction

As a part of the work carried out by the European Topic Centre on Air Quality and Climate Change Mitigation (ETC/ACM), annual European air quality maps of PM and ozone have been produced for many years (Horálek et al., 2017a and references therein). Nitrogen dioxide (NO_2) maps started to be produced regularly within the framework of the ETC/ACM under the mapping report Horálek et al. (2017a), in which the annual average 2014 map for NO_2 is presented. This map has been prepared based on methodology routinely used for PM_{10} and ozone. However, its uncertainty analysis gives poorer results compared to the maps of PM_{10} and ozone. Next to this, it concerns only rural and urban background areas not accounting for hot spot location (traffic), although traffic is the most important source of NO_2. In order to produce more advanced European-scale maps of annual average NO_2 concentrations, an improved methodology has been developed in Horálek et al. (2017b).

In the improved methodology, land cover and road data are included in the rural and urban background NO_2 mapping, namely the CLC land cover data and the Global Road Inventory Project (GRIP) data (Meijer et al., 2016). Simultaneously, the application of the 1x1 km resolution is moved from the combined final merging process-step to the early process-step of creation of the separate rural and urban background map layers. Next to this, the improved methodology incorporates the traffic map layer – created based on the urban and suburban traffic stations and selected supplementary data – with the background map, using the road data showing the area influenced by urban traffic. Next to the NO_2 concentration map, the improved methodology has been developed also for population exposure estimates, in order to better reflect the population exposed to traffic.

In this paper, the improved methodology is applied for NO_2 annual average 2014. The analysis is based on interpolation of annual average of NO_2 monitoring data from 2014, reported by EEA member and cooperating countries in 2015. In the map creation, monitoring data (considered as primarily data) are combined with chemical transport model results and other supplementary data (including land cover). Chapter 2 presents the final concentration map and exposure estimates. Annex 1 describes briefly the applied improved methodology. Annex 2 documents the input data applied. Annex 3 presents the technical details of the map and its uncertainty analysis including the cross-validation results. Annex 4 presents separate urban traffic map applicable for urban hotspots areas only.

2 NO_2 improved map

2.1 NO_2 – Annual mean

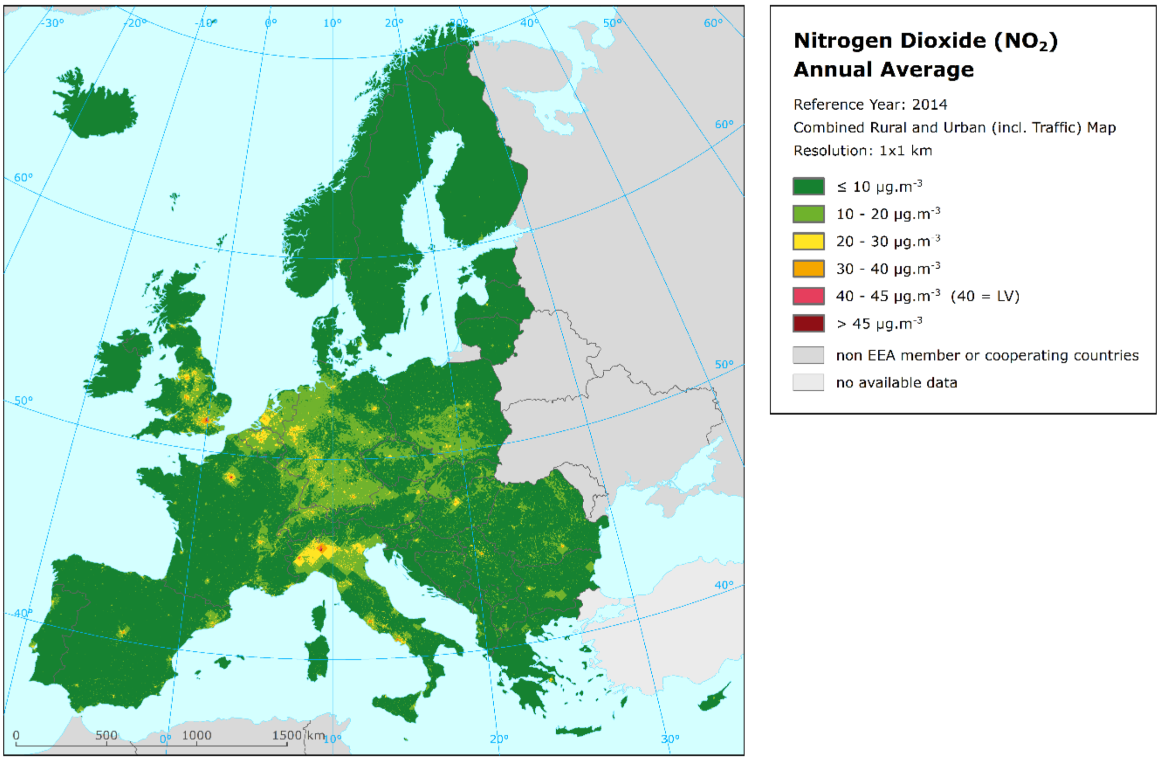
2.1.1 Concentration map

The Ambient Air Quality Directive (EU, 2008) sets the limit value (LV) for the NO_2 annual average at the level of 40 $\mu g.m^{-3}$. This is the same concentration level as recommended by the World Health Organization for the NO_2 annual average as the Air Quality Guideline (WHO, 2005).

Map 2.1 presents the final combined concentration 1x1 km gridded map for the 2014 NO_2 annual average as the result of interpolation and merging of the separate maps as described in Annex 1.

Supplementary data used in the linear regression are in principle similar as in Horálek et al. (2017a): for rural areas they consist of EMEP model output (Simpson et al., 2012), altitude, wind speed, population density and land cover; for urban background areas the EMEP model output, altitude, wind speed, population density, and land cover are supplemented with GRIP road data; for traffic areas the EMEP model output, altitude, wind speed and land cover are used (Annex 2).

Map 2.1 Concentration map of NO_2 annual average, 2014, improved methodology



The most of the European area shows NO_2 levels below 20 $\mu g.m^{-3}$, with most of the rural areas below 10 $\mu g.m^{-3}$. Some areas above 20 $\mu g.m^{-3}$ can be found in the Po valley, the Benelux, the German Ruhr region, in central and southern England, in the Île de France and around Rome. According to the measurements from monitoring stations, the limit value (LV) of 40 $\mu g.m^{-3}$ is exceeded specifically in several large agglomerations all over Europe. However, it should be noted that the interpolated map is created at 1x1 km only and as such refers to the rural and urban *background* situations only, while the exceedances of the NO_2 limit values occur mostly at local *hotspots* such as densely urbanised and industrialised areas, and dense traffic locations. In order to visualise the actual urban traffic concentration levels at the hotspots areas, a separate urban traffic map is presented in Annex 4, Figure A4.1. It should be emphasized that this map is applicable for urban traffic areas only.

The relative mean uncertainty of the NO_2 annual average map is 37 % for rural and 24 % for urban background areas (Annex 3). Compared with the results of Horálek et al. (2017a) one can conclude that the updated methodology improves the mapping uncertainty by 7 percentage units for rural and by 2 percentage units for urban background areas.

2.1.2 Population exposure

Table 2.1 gives the population frequency distribution for a limited number of exposure classes calculated on a grid of 1x1 km resolution, as well as the population-weighted concentration for individual countries and for Europe as a whole according to Equations A1.3 and A1.4 of Annex 1.

It has been estimated that in 2014 about 3 % of the European population and also the EU-28 population lived in areas with NO_2 annual average concentrations above the EU limit value of 40 $\mu g.m^{-3}$. CSI004 (EEA, 2016b) estimates that about 7 % of the population in urban agglomerations in the EU-28 was exposed in 2014 to levels above the EU limit value. The difference between the two estimated fractions is mainly because CSI004 accounts only for the urban population in the agglomerations with a total population of about 180 million (where, as pointed out above, hotspots are most frequent). Whereas, Table 2.1 provides estimates for the total population (about 530 million) including the population in rural areas, smaller cities and villages.

The European-wide population-weighted concentration of the NO_2 annual average for 2014 is estimated to be about 19 $\mu g.m^{-3}$, for both EU-28 and Europe as a whole.

Figure 2.1 shows, for the whole mapped area, the population frequency distribution for exposure classes of 2 $\mu g.m^{-3}$. The median concentration is about 17.5 $\mu g.m^{-3}$: half of the population is exposed to concentrations lower than 17.5 $\mu g.m^{-3}$ while the other half is exposed to higher concentrations.

Figure 2.2 Population frequency distribution, NO_2 annual average, 2014, improved methodology

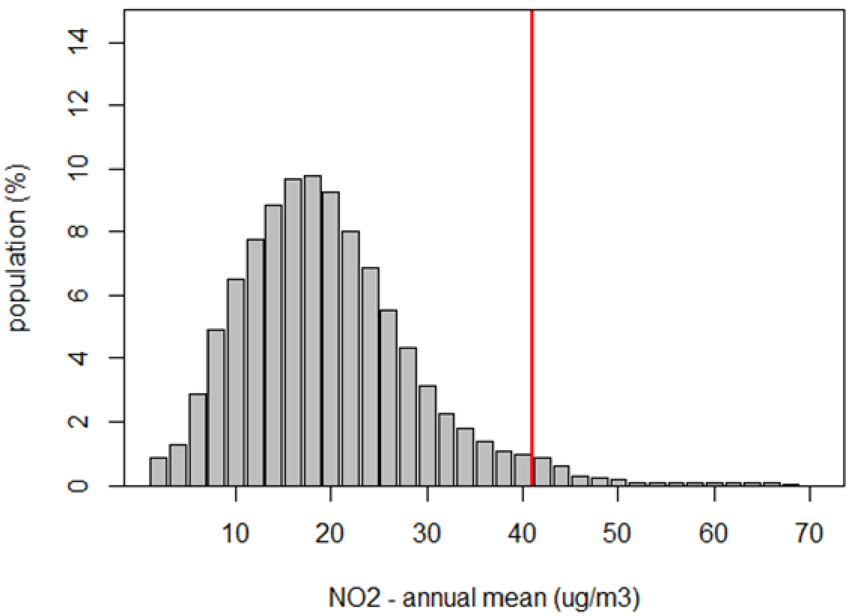


Table 2.1 Population exposure and population-weighted concentration, NO_2 annual average, 2014, improved methodology

Country		Population	NO_2 annual average, exposed population [%]						Population weighted conc.
			< LV				> LV		
		[inhbs . 1000]	< 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 $\mu g.m^{-3}$	[$\mu g.m^{-3}$]
Albania	AL	2 896	24.2	54.3	18.7	2.8			14.8
Andorra	AD	73	3.9	95.7	0.4				15.0
Austria	AT	8 507	11.7	47.7	28.2	11.9	0.4		19.2
Belgium	BE	11 204	1.8	40.0	45.7	11.5	1.0		21.9
Bosnia & Herzegovina	BA	3 831	25.9	52.3	21.7	0.0			15.1
Bulgaria	BG	7 246	15.8	58.1	23.4	2.7			16.5
Croatia	HR	4 247	25.5	48.8	24.8	1.0			15.7
Cyprus	CY	858	37.8	55.2	4.9	2.0			12.8
Czech Republic	CZ	10 512	8.7	69.3	19.9	2.0	0.1		16.8
Denmark	DK	5 627	54.9	32.3	11.9	0.7	0.3		11.0
Estonia	EE	1 316	59.6	36.5	3.9				9.0
Finland	FI	5 451	64.9	31.9	3.0	0.2			8.3
France (metropolitan)	FR	63 989	27.6	40.5	19.0	7.7	2.9	2.3	17.7
Germany	DE	80 767	4.2	50.8	37.3	4.8	1.3	1.6	20.2
Greece	GR	10 927	41.7	26.2	21.9	7.8	1.8	0.5	14.9
Hungary	HU	9 877	10.2	62.6	20.6	5.8	0.6	0.2	17.1
Iceland	IS	326	36.7	57.3	5.9	0.0			10.9
Ireland	IE	4 606	76.2	21.3	2.5				6.1
Italy	IT	60 783	8.0	36.6	35.4	14.1	3.6	2.3	22.5
Latvia	LV	2 001	39.1	49.0	10.5	1.5			12.3
Liechtenstein	LI	37	1.4	72.7	24.7	0.8	0.5		18.5
Lithuania	LT	2 943	37.5	54.6	7.5	0.5			12.5
Luxembourg	LU	550	4.8	50.2	35.5	9.5			19.9
Macedonia, FYROM of	MK	2 066	6.7	73.2	18.8	1.3			16.0
Malta	MT	425	9.4	76.1	11.6	3			16.0
Monaco	MC	38		1.7	77	22			24.5
Montenegro	ME	622	24.3	65.3	10.5				13.9
Netherlands	NL	16 829	0.5	41.5	46.8	10.4	0.6	0.2	21.9
Norway	NO	5 108	43.2	38.7	15.3	2.1	0.6		12.4
Poland	PL	38 018	21.4	58.4	18.4	1.2	0.4	0.1	15.1
Portugal (excl. Az., Mad.)	PT	9 922	41.9	35.8	19.4	2.6	0.3		13.7
Romania	RO	19 947	14.4	60.5	20.6	3.8	0.4	0.2	16.5
San Marino	SM	33	3.1	92.4	4.4	0			14.7
Serbia (incl. Kosovo*)	RS	7 147	15.5	46.5	28.6	8.8	0.6		18.5
Slovakia	SK	5 416	8.3	81.4	8.8	1.5			15.2
Slovenia	SI	2 061	26.9	53.1	19.6	0.4			14.9
Spain (excl. Canarias)	ES	44 397	12.0	44.3	30.6	10.0	2.7	0.4	19.9
Sweden	SE	9 645	55.9	39.6	3.3	1.2	0.0		9.9
Switzerland	CH	8 140	2.8	48.3	41.0	4.1	1.5	2.3	20.9
United Kingdom (& dep.)	UK	64 351	9.0	36.1	36.0	14.3	2.4	2.3	22.2
Total		**532 738**	**16.4**	**45.3**	**28.0**	**7.5**	**1.6**	**1.2**	**18.6**
			97.2				**2.8**		
EU-28		**502 424**	**16.3**	**45.1**	**28.1**	**7.7**	**1.7**	**1.2**	**18.7**
			97.1				**2.9**		

Kosovo*	KS	1 821	21.0	73.1	5.9				13.6
Serbia (excl. Kosovo*)	RS	7 147	14.1	40.0	34.2	10.9	0.7		19.6

*) under the UN Security Council Resolution 1244/99

Note 1: Turkey is not included in the calculation due to the lack of air quality data.
Note 2: Empty cells mean: no population in exposure.

Comparing Table 2.1 with the results of Horálek et al. (2017a), one can conclude that the overall European population-weighted concentration calculated by the improved approach is only slightly higher compared to the population exposure calculated based on the earlier used method: the inclusion of land cover lowers slightly the estimated value, while the inclusion of the traffic layer raises this. The main difference is in the higher fraction of population exposed to high concentrations, caused by the inclusion of the traffic layer, which is related to the hot spot locations. Next to this, one can see a higher fraction of population exposed to concentrations below 10 $\mu g.m^{-3}$, which is related with a more realistic estimation of the population exposure due to taking into account the land cover data source. All these results are in agreement with findings of Horálek et al. (2017b).

References

Danielson JJ, Gesch DB (2011). Global multi-resolution terrain elevation data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011–1073. https://lta.cr.usgs.gov/GMTED2010

ECMWF: Meteorological Archival and Retrieval System (MARS). http://www.ecmwf.int/

EEA (2008). ORNL Landscan 2008 Global Population Data conversion into EEA ETRS89-LAEA5210 1km grid (eea_r_3035_1_km_landscan-eurmed_2008, by Hermann Peifer of EEA).

EEA (2011). Guide for EEA map layout. EEA operational guidelines. August 2011, version 4. http://www.eionet.europa.eu/gis/docs/GISguide_v4_EEA_Layout_for_map_production.pdf

EEA (2013a). Corine land cover 2000 (CLC2000) raster data. 100x100m gridded version 17 (12/2013). http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-raster-3

EEA (2013b). Corine land cover 2006 (CLC2006) raster data. 100x100m gridded version 17 (12/2013) http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster-3

EEA (2016a). Air Quality e-Reporting. Air quality database. http://www.eea.europa.eu/data-and-maps/data/aqereporting-1

EEA (2016b). Exceedance of air quality limit values in urban areas. CSI004 indicator assessment. http://www.eea.europa.eu/data-and-maps/indicators/exceedance-of-air-quality-limit-3/assessment-2

Eurostat (2014). GEOSTAT 2011 grid dataset. Population distribution dataset. http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography

Eurostat (2016). Total population for European states for 2014. http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&language=en&pcode=tps00001&tableSelection=1&footnotes=yes&labeling=labels&plugin=1

Horálek J, Denby B, de Smet P, de Leeuw F, Kurfürst P, Swart R, van Noije T (2007). Spatial mapping of air quality for European scale assessment. ETC/ACC Technical paper 2006/6. http://acm.eionet.europa.eu/reports/ETCACC_TechnPaper_2006_6_Spat_AQ

Horálek J, de Smet P, de Leeuw F, Kurfürst P, Benešová N (2017a). European air quality maps for 2014. ETC/ACM Technical Paper 2016/6. http://acm.eionet.europa.eu/reports/ETCACM_TP_2016_6_AQMaps2014

Horálek J, de Smet P, Schneider P, Kurfürst P, de Leeuw F (2017b). Inclusion of land cover and traffic data in NO_2 mapping methodology. ETC/ACM Technical Paper 2016/12. http://acm.eionet.europa.eu/reports/ETCACM_TP_2016_12_LC_and_traffic_data_in_NO2_mapping

JRC (2009). Population density disaggregated with Corine land cover 2000. 100x100 m grid resolution, EEA version popu01clcv5.tif of 24 Sep 2009. http://www.eea.europa.eu/data-and-maps/data/population-density-disaggregated-with-corine-land-cover-2000-2

Mareckova K, Pinterits M, Wankmüller R, Tista M (2016). Inventory Review 2016. Review of emission data reported under the LRTAP Convention and NEC Directive. Stage 1 and 2 review & Status of gridded and LPS data. EEA/CEIP Technical Report 1/2016. http://www.ceip.at/review_proces_intro/review_reports

NILU (2016). EBAS, database of atmospheric chemical composition and physical properties (NILU, Norway). http://ebas.nilu.no/

NMI (2016). EMEP/MSC-W modelled air concentrations and depositions. Yearly, monthly, daily and hourly gridded data. http://thredds.met.no/thredds/catalog/data/EMEP/2016_Reporting/catalog.html

ORNL (2008). ORNL LandScan high resolution global population data set. http://www.ornl.gov/sci/landscan/landscan_documentation.shtml

Simpson D, Benedictow A, Berge H, Bergström R, Emberson LD, Fagerli H, Hayman GD, Gauss M, Jonson JE, Jenkin ME, Nyíri A, Richter C, Semeena VS, Tsyro S, Tuovinen J-P, Valdebenito A, Wind P (2012). The EMEP MSC-W chemical transport model – technical description. Atmospheric Chemistry and Physics, 12, 7825–7865, doi:10.5194/acp-12-7825-2012. http://www.atmos-chem-phys.net/12/7825/2012/acp-12-7825-2012.html

UN (2015). World Population Prospects, the 2015 Revision. United Nations. Department of Economic and Social Affairs, Population Division. http://esa.un.org/unpd/wpp/Download/Standard/Population/

WHO (2005). WHO Air quality guidelines for particulate matters, ozone, nitrogen dioxide and sulphur dioxide. Global update 2005. http://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/index.html

Annex 1 Methodology

A1.1 Mapping method

The mapping methodology is in principle the same as documented in Horálek et al. (2017b). It is an improved variant of the regression – interpolation – merging mapping. Separate map layers are created for rural, urban background and urban traffic areas on a grid at 1x1 km resolution. The rural background map layer is based on the rural background stations, the urban background map layer on the urban and the suburban background stations, and the urban traffic map layer on the urban and the suburban traffic stations. All the map layers are created using a linear regression model followed by kriging of the residuals produced from that model. Interpolation is therefore carried out according to the relation:

$$\hat{Z}(s_0) = c + a_1.X_1(s_0) + a_2.X_2(s_0) + \ldots + a_n.X_n(s_0) + \eta(s_0) \qquad \text{(A1.1)}$$

where $\hat{Z}(s_0)$ is the estimated value of the air pollution indicator at the point s_o,

$X_1(s_0), X_2(s_0), \ldots, X_n(s_0)$ are the n number of individual supplementary variables at the point s_o

$c, a_1, a_2, \ldots, a_n$ are the $n+1$ parameters of the linear regression model calculated based on the data at the points of measurement,

$\eta(s_0)$ is the spatial interpolation of the residuals of the linear regression model at the point s_o calculated based on the residuals at the points of measurement.

For different map layers (rural, urban background, urban traffic) different supplementary data are used, depending on their improvement to the fit of the regression. The three map layers are merged into one final map using a weighting procedure

$$\hat{Z}_F(s_0) = (1 - w_U(s_0)) \cdot \hat{Z}_R(s_0) + w_U(s_0)(1 - w_T(s_0)) \cdot \hat{Z}_{UB}(s_0) + w_U(s_0) w_T(s_0) \cdot \hat{Z}_T(s_0) \qquad \text{(A1.2)}$$

where $\hat{Z}_F(s_0)$ is the resulting estimated concentration in a grid cell s_o for the final map,

$\hat{Z}_{UB}(s_0)$ is the estimated concentration in a grid cell s_o for the urban background map layer,

$\hat{Z}_R(s_0)$ is the estimated concentration in a grid cell s_o for the rural background map layer,

$\hat{Z}_T(s_0)$ is the estimated concentration in a grid cell s_o for the urban traffic map layer,

$w_U(s_0)$ is the weight representing the ratio of the urban character of the a grid cell s_o,

$w_T(s_0)$ is the weight representing the ratio of areas exposed to traffic air quality in a grid cell s_o.

The weight $w_U(s_0)$ is based on the population density grid, while $w_T(s_0)$ is based on the buffers around the roads.

For further details, see Horálek et al. (2017b and references therein).

In all calculations and map presentations the EEA standard projection ETRS89-LAEA5210 (also known as ETRS89 / LAEA Europe, see www.epsg-registry.org) is used. The interpolation and mapping domain consists of the areas of all EEA member and cooperating countries, as far as they fall into the EEA map extent *Map_1c* (EEA, 2011). The mapping area covers the whole Europe apart from Belarus, Moldova, Ukraine and the European parts of Russia and Kazakhstan.

A1.2 Calculation of population exposure

Population exposure for individual countries and for Europe as a whole is calculated from the air quality map layers and population density data, both at 1x1 km resolution. It is calculated separately for the areas where the air quality is considered to be directly influenced by traffic and for the background (both rural and urban) areas. For each concentration class '*j*', the percentage population per country as well as the European-wide total is determined:

$$P_j = \frac{\sum_{i=1}^{N} I_{Bij}\left(1 - w_U(i)w_T(i)\right)p_i + \sum_{i=1}^{N} I_{Tij} w_U(i) w_T(i) p_i}{\sum_{i=1}^{N} p_i} \cdot 100 \qquad \text{(A1.3)}$$

where P_j is the percentage population living in areas of the *j*-th concentration class in either the country or in Europe as a whole,

p_i is the population in the *i*-th grid cell,

I_{Bij} is the Boolean 0-1 indicator showing whether the background air quality concentration (estimated by the combined rural/urban background map layer) in the *i*-th grid cell is within the *j*-th concentration class ($I_{Bij} = 1$), or not ($I_{Bij} = 0$),

I_{Tij} is the Boolean 0-1 indicator showing whether the traffic air quality concentration in the

i-th grid cell is within the *j*-th concentration class ($I_{Tij} = 1$), or not ($I_{Tij} = 0$),

N is the number of grid cells in the country or in Europe as a whole.

In addition, we express per-country and European-wide exposure as the population-weighted concentration, i.e. the average concentration weighted according to the population in a grid cell:

$$\hat{c} = \frac{\sum_{i=1}^{N} c_i p_i}{\sum_{i=1}^{N} p_i} \qquad \text{(A1.4)}$$

where $\hat{c}$ is the population-weighted average concentration in the country or in the whole Europe,

c_i is the concentration in the i^{th} grid cell.

A1.3 Methods for uncertainty analysis

The uncertainty estimation of the European map is based on cross-validation. The cross-validation method computes the spatial interpolation for each measurement point from all available information except from the point in question. The predicted and measurement values at these points are plotted in the form of a scatter plot. With help of statistical indicators the quality of the predictions is demonstrated objectively. The advantage of the nature of this cross-validation technique is that it enables evaluation of the quality of the predicted values at locations without measurements, as long as they are within the area covered by the measurements. The main statistical indicator used are *root mean squared error* (RMSE), *relative root mean squared error* (RRMSE), *bias* (or mean prediction error, MPE). Other indicators are R^2 and the regression equation parameters *slope* and *intercept*, following from the scatter plot between the predicted (using cross-validation) and the observed concentrations.

In addition, we make a simple comparison between the point measurements and interpolated values of the 1x1 km grid for separate map layers and for the final combined map. Note that the grid cell value is the averaged result of the interpolation in this grid cell area. The interpolated value within a grid cell will only approximate the predicted value(s) at the station(s) lying within that cell.

For further details, see Horálek et al. (2017a).

Annex 2 Input data

The input data in this paper are mostly the same as in Horálek et al. (2017a, 2017b).

A2.1 Air quality monitoring data

Air quality station monitoring data for the relevant year are extracted from the Air Quality e-Reporting database, EEA (2016a). This data set is supplemented with 11 additional rural background stations from the database EBAS (NILU, 2016) not reported to the Air Quality e-Reporting database. Only data from stations classified as *background* and *traffic* (for the three types of area, *rural*, *suburban* and *urban*) are used. Station type *industrial* is not considered; it represents local scale concentration levels not applicable at the mapping resolution employed. The following pollutant and its indicator is considered:

NO_2 – annual average [$\mu g.m^{-3}$], year 2014

Only the stations with annual data coverage of at least 75 percent are used. We excluded the stations located outside the interpolation and mapping domain. In total, 410 rural background stations, 1126 urban/suburban background stations and 704 urban/suburban traffic stations are used. Rural traffic stations are not considered due to their small number (i.e. 16), in agreement with Horálek et al. (2017b).

A2.2 EMEP MSC-W model output

The chemical dispersion model used is the EMEP MSC-W (formerly called Unified EMEP) model (version rv4.9), which is an Eulerian model. Simpson et al. (2012) and https://wiki.met.no/emep/page1/emepmscw_opensource (web site of Norwegian Meteorological Institute) describe the model in more detail. Emissions for the relevant year 2014 (Mareckova et al., 2016) are used and the model is driven by ECMWF meteorology for the relevant year 2014. EMEP (2016) provides details on the EMEP modelling for 2014. The resolution of the model is cc. 50x50 km.

We downloaded the EMEP data from NMI (2016) in the form of NO_2 daily means. We aggregated these primary data to the NO_2 annual average 2014 values. We converted the data to 1x1 km grid resolution. For details of conversion, see Horálek et al. (2017a).

A2.3 Other supplementary data

Altitude

We use the altitude data field (in meters) of *Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010)*, with an original grid resolution of 15x15 arcseconds, see Danielson et al. (2011). We converted the field into the ETRS 1989 LAEA projection. In the following step, we resampled the raster dataset to 100x100 m resolution and shifted it to the extent of EEA reference grid. As a final step, two spatial aggregation were executed into 1x1 km grid resolution: the spatial averaging of 1x1 km grid, and the floating averaging of the circle with radius of 5 km around all relevant 1x1 km grid cells.

Meteorological parameter – wind speed

In this paper we use the following ECMWF variable on a 0.25x0.25 degrees resolution: wind speed – annual average [$m.s^{-1}$], year 2014 (aggregated from 6-hour means).

We converted the data to 1x1 km grid resolution. For details of deriving, aggregation and conversion, see Horálek et al. (2007, 2017a).

Population density and population totals

Population density (in inhbs.km^{-2}, census 2011) is based on *Geostat 2011* grid dataset, Eurostat (2014). The dataset is in 1x1 km resolution, in the EEA reference grid. For regions not included in the Geostat 2011, alternative sources were used, namely JRC (2009) for Gibraltar and ORNL (2008) for Faroe Islands, British crown dependencies and northern Cyprus. For details, see Horálek et al. (2017a).

National population totals presented in the exposure table of this paper are based on Eurostat (2016) national population data for 2014. For Andorra and Monaco, the population total is based on UN population data (UN, 2015) for 2014. For details, see Horálek et al. (2017a).

Land cover

CORINE Land Cover 2006 – grid 100 x 100 m, Version 17 (12/2013) is used (CLC2006 – 100m, g100_06.zip; EEA, 2013b). Greece is missing in the CLC2006, Therefore, we inserted for Greece the CLC2000 data (grid 100 x 100 m, Version 17, 12/2013 EEA, 2013a). The countries and regions missing in both CLC2006 and CLC2000 are Andorra, Faroe Islands and British crown dependencies.

In agreement with Horálek et al. (2017b), the 44 CLC classes have been re-grouped into the 8 more general classes, see. In this paper we use the following general classes:

Table 2.1 General land cover classes, based on CLC2006 classes, used in mapping

Label	General class description	CLC classes grid codes	CLC classes codes	CLC classes description
HDR	High density residential areas	1	111	Continuous urban fabric
LDR	Low density residential areas	2	112	Discontinuous urban fabric
AGR	Agricultural areas	12 – 22	211 – 244	Agricultural areas
NAT	Natural areas	23 – 34	311 – 335	Forest and semi natural areas

Two aggregations are used, i.e. into 1x1 km grid and into the circle with radius of 5 km. For each general CLC class we spatially aggregated the high land use resolution into the 1x1 km EEA standard grid resolution. The aggregated grid square value represents for each general class the total area of this class as percentage of the total 1x1 km square area. For details, see Horálek (2017b).

Road type vector data

GRIP (Meijer et al., 2016) vector road type data base provided by PBL is used.

Percentage of the area influenced by traffic is represented by buffers around the roads: for the individual classes 1 – 4, for all classes together and for classes 1 – 3 together, at all 1x1 grid cells; a buffer of 75 metres distance at each side from each road vector is taken for the roads of classes 1 and 2, while a buffer of 50 metres is taken for the roads of classes 3 and 4,

Annex 3 Technical details and mapping uncertainties

A3.1 NO_2

This annex contains technical details on the linear regression models and the residual kriging, including the performance. Furthermore, uncertainty estimates for the NO_2 annual average map are given, both for the separate (i.e. rural, urban background and urban traffic) map layers and for the final combined map (i.e. Map 2.1). As stated in Section A1.1, the final combined map is constructed by merging of the three separate map layers. Next to this, the urban traffic map layer is presented separately in Map A4.1.

Technical details on the interpolation model

Table A3.1 presents the estimated parameters of the linear regression models and of the residual kriging and includes the statistical indicators of both the regression and the kriging.

Table A3.1 Parameters and statistics of linear regression model and ordinary kriging of NO_2 annual average for 2014 in rural (left), urban background (middle) and urban traffic (right) areas for the final combined map

		NO_2 Annual average		
		Rural areas	**Urban backgr. areas**	**Urban traffic areas**
Linear regresion model (LRM, Eq. A1.1)	c (constant)	9.7	17.38	23.7
	a1 (EMEP model)	0.698	0.54	0.693
	a2 (altitude_1km)	-0.0085	-0.0087	-0.0229
	a3 (altitude_5km_radius)	0.0067	0.0089	0.0220
	a4 (wind speed)	-1.13	-1.84	-1.27
	a5 (population*1000)	0.00064	0.00022	*n. sign.*
	a6 (NAT_1km)	*n. sign.*	-0.0562	*n. sign.*
	a7 (AGR_1km)	*n. sign.*	-0.0242	*n. sign.*
	a8 (LDR_5km_radius)	0.00224	0.00164	0.00363
	a9 (HDR_5km_radius)		0.00245	0.00352
	a10 (NAT_5km_radius)	-0.00050	*n. sign.*	*n. sign.*
	a11 (T1buf75m_1km)		13.736	*n. sign.*
	Adjusted R^2	**0.67**	**0.52**	**0.38**
	Standard Error [μg.m^{-3}]	**3.4**	**5.3**	**10.3**
Ordinary kriging (OK) of LRM residuals	nugget	12	16	58
	sill	12	24	134
	range [km]	190	300	570
LRM + OK of its residuals	**RMSE [μg.m^{-3}]**	**3.3**	**4.8**	**8.9**
	Relative RMSE [%]	**36.5**	**23.6**	**25.5**
	Bias (MPE) [μg.m^{-3}]	**0.1**	**0.0**	**0.0**

It should be noted that for the limited areas with the lack of CLC land cover data (see Section A2.3), we applied in the final combined map (see Map 2.1) the NO_2 annual average 2014 concentration levels as presented in Horálek et al. (2017a), i.e. based on the old methodology.

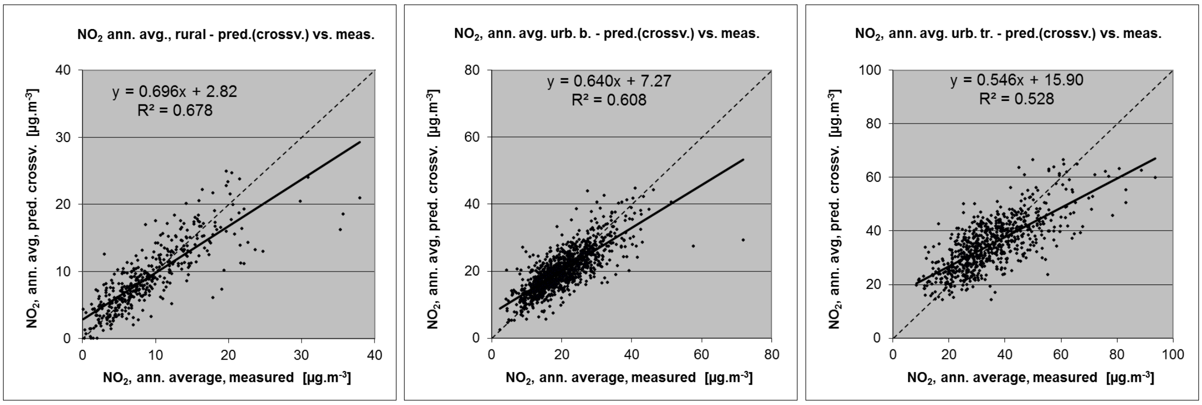
Uncertainty estimated by cross-validation

Table A3.1 shows both absolute and relative mean uncertainty, expressed by RMSE and Relative RMSE. The absolute mean uncertainty of the NO_2 annual average expressed as RMSE is 3.3 μg.m^{-3} for the rural areas and 4.8 μg.m^{-3} for the urban background areas. For the NO_2 urban traffic areas it is 8.9 μg.m^{-3}.

The relative mean uncertainty of the NO_2 annual average map is 37 % for rural and 24 % for urban background areas. For the NO_2 urban traffic areas it is 25 %.

Figure A3.1 shows the cross-validation scatter plots for NO_2 annual average. The R^2 indicates that the variability is attributable to the interpolation for about 68 % at the rural areas, for about 61 % at the urban background areas and for about 68 % at the urban traffic areas.

Figure A3.1 Correlation between cross-validated predicted and measurement values for NO_2 annual average 2014 for rural (left), urban background (middle) and urban traffic (right) areas



Like in the case of other pollutants, the cross-validation scatter plots indicate the underestimation of high concentrations in the places with no measurements. For example, in urban background areas an observed value of 40 µg.m^{-3} is estimated in the interpolations to be about 33 µg.m^{-3}, about 18 % too low.

Comparison of point measurement values with the predicted grid value

Next to the above presented cross-validation, a simple comparison was made between the point observation values and interpolated predicted 1x1 km grid values. The comparison has been made primarily for the separate rural, separate urban background and separate urban traffic map layers. Beside this, the comparison has been done also for the final combined map.

Table A3.2 presents the results of this comparison. One can conclude that the final combined map in 1x1 km resolution is representative for rural and urban background areas, but not for urban traffic areas.

Table A3.2 Statistical indicators from the scatter plots for the predicted grid values from separate (rural, urban background or urban traffic) map layers and final combined map versus the measurement point values for rural (upper left), urban background (upper right) and urban traffic (bottom left) stations for NO_2 annual average 2014

	rural backgr. stations				urban/suburban backgr. stations			
	RMSE	bias	R^2	lin. r. equation	RMSE	bias	R^2	lin r. equation
grid prediction, separate (r or ub or ut) map	3.2	0.1	0.703	y = 0.708x + 2.72	4.0	0.0	0.719	y = 0.697x + 6.11
grid prediction, final merged map	3.4	0.5	0.685	y = 0.781x + 2.48	4.5	0.8	0.666	y = 0.752x + 5.85

	urban/suburban traffic stations			
	RMSE	bias	R^2	lin. r. equation
grid prediction, 1x1 km separate (r or ub) map	7.1	0.0	0.959	y = 0.966x + 1.21
grid prediction, 1x1 km final merged map	14.9	-11.1	0.528	y = 0.546x + 15.90

Annex 4 Urban traffic map

Map A4.1 presents the urban traffic map layer prepared on basis of the urban/suburban traffic stations (Annex 3), in order to visualise the actual urban traffic concentration levels at the hotspots areas. It should be emphasized that this map is applicable for urban traffic areas only.

The relative mean uncertainty of the urban traffic map is 25 % (Annex 3).

Map A4.1 Concentration map of NO_2 annual average, 2014, urban traffic air quality. Applicable for urban traffic areas only.