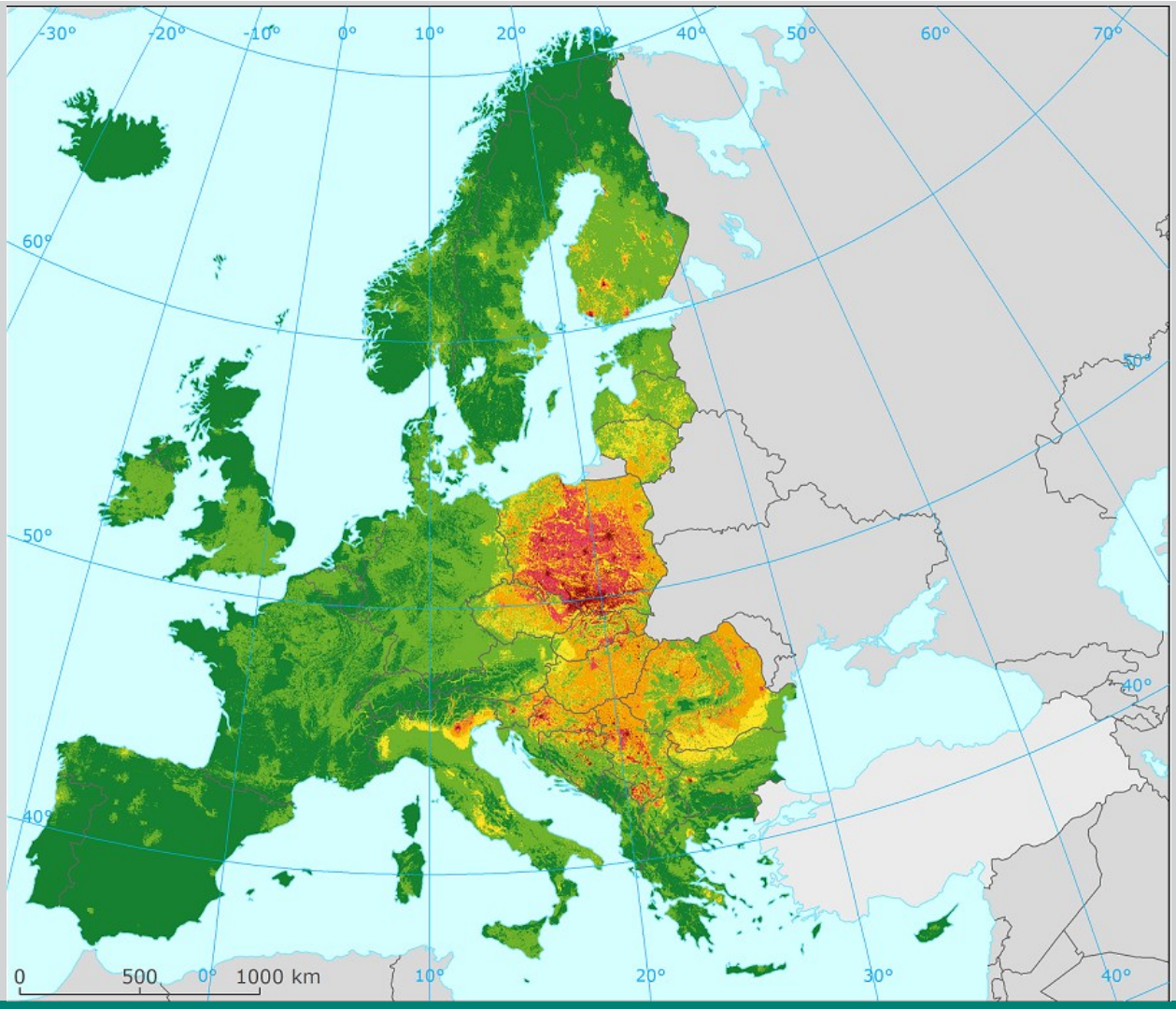
Benzo(a)pyrene (BaP) annual mapping

Evaluation of its potential regular updating

September 2021



Authors:
Jan Horálek (CHMI), Markéta Schreiberová (CHMI), Philipp Schneider (NILU)

ETC/ATNI consortium partners:
NILU – Norwegian Institute for Air Research, Aether Limited, Czech Hydrometeorological Institute (CHMI), EMISIA SA, Institut National de l'Environnement Industriel et des risques (INERIS), Universitat Autònoma de Barcelona (UAB), Umweltbundesamt GmbH (UBA-V), 4sfera Innova, Transport & Mobility Leuven NV (TML)

European Environment Agency
European Topic Centre on Air pollution, transport, noise and industrial pollution

Cover design: EEA
Cover picture: Concentration map of BaP annual average 2019. Units: ng·m^{-3}. (Map 4.6 of this report.)
Layout: ETC/ATNI

The withdrawal of the United Kingdom from the European Union did not affect the production of the report.
Data reported by the United Kingdom are included in all analyses and assessments contained herein, unless otherwise indicated.

Authors
Jan Horálek, Markéta Schreiberová, Czech Hydrometeorlogical Institute (CHMI, Czechia)
Philipp Schneide, NILU - Norwegian Institute for Air Research (NILU, Norway)

ETC/ATNI c/o NILU
ISBN 978-82-93752-44-8

European Topic Centre on Air pollution, transport, noise and industrial pollution
c/o NILU – Norwegian Institute for Air Research
P.O. Box 100, NO-2027 Kjeller, Norway
Tel.: +47 63 89 80 00
Email: etc.atni@nilu.no
Web : https://www.eionet.europa.eu/etcs/etc-atni

Contents

Summary 5

Acknowledgements 6

1 Introduction 7

2 Methodology 8
- 2.1 Mapping method 8
- 2.2 Pseudo BaP stations 8
- 2.3 Uncertainty estimates of the concentration maps 9

3 Input data 10
- 3.1 Air quality monitoring data 10
- 3.2 Chemical transport modelling data 11
- 3.3 Other supplementary data 12

4 Analysis 13
- 4.1 Initial BaP mapping 13
- 4.2 Pseudo BaP data 16
- 4.3 Final BaP mapping 18

5 Conclusions and recommendations 22

6 References 23

Annex Concentration maps including station points 25

Summary

The report explores potential regular production of benzo(a)pyrene (BaP) maps at the European scale in line with the operational production of other air quality maps.

The analysis has been done based on 2018 and 2019 data. In order to obtain spatial estimates of BaP, datasets of station measurements have been combined with chemical transport modelling and other supplementary data using the Regression – Interpolation – Merging Mapping (RIMM) method.

Initially, a mapping approach using only actual BaP measurements (and supplementary data) has been examined. However, due to the poor spatial coverage of BaP measurement stations at the European scale, such maps show large areas with high interpolation uncertainty, mainly in rural regions, but to some extent also in urban areas. Trying to overcome this limitation, so-called pseudo BaP data have been introduced in areas with a lack of BaP stations. In such regions, estimates of BaP were computed based on $PM_{2.5}$ measurements (or with pseudo $PM_{2.5}$ data based on PM_{10} measurements), using exponential regression of the observed BaP concentrations with the $PM_{2.5}$ data, as well as geographical coordinates and land cover as predictor variables.

Based on both actual BaP measurements and pseudo BaP data, maps of annual average BaP concentration have been prepared for all of Europe. The uncertainty of the final maps was evaluated by leave-one-out cross-validation and by the interpolation relative standard uncertainty. The cross-validation uncertainty has been estimated at a level of about 120% (for 2018) and 100% (for 2019) in rural areas and about 70% (in both years) in urban areas. The interpolation uncertainty has been estimated to be below 60% (in both years) in the majority of the study region. Based on the results, the report recommends the regular production of BaP maps. However, due to the relatively high cross-validation uncertainty (particularly in the rural areas), it is recommended to label them as experimental maps to indicate that they do not yet meet the same accuracy standards as the regularly produced maps of other pollutants.

Acknowledgements

The EEA task manager was Alberto González Ortiz.

The daily BaP data were extracted from the AQ e-reporting database by 4sfera. The modelling data were spatially transformed to the regular EEA grid by Jana Schovánková (CHMI). The report also benefited from the expertise of Cristina Guerreiro (NILU).

We are grateful to Alexey Gusev (EMEP, MCS-E, Moscow, Russia) for providing information on the regular BaP EMEP modelling under the Meteorological Synthesizing Centre - East (MCS-E).

The analysis of the 2019 data has been performed in a cooperation with the Czech national project ARAMIS.

1 Introduction

Concentration values of benzo(a)pyrene (BaP), as an indicator for the polycyclic aromatic hydrocarbons, are regularly monitored and assessed in the context of the 2004 Ambient Air Quality Directive (EC, 2004). BaP is considered among four of the most relevant pollutants in Europe in terms of health impacts and population exposure to concentrations above EU standards, together with PM, NO_2 and ozone (Guerreiro et al., 2015, EEA, 2020). However, contrary to the other mentioned pollutants, BaP is so far not included among the pollutants for which the European-wide spatial maps are regularly produced (Horálek et al., 2021). Under the European Topic Centre on Air Pollution and Climate Change Mitigation (ETC/ACM), what was the ETC/ATNI predecesor, a first European map of BaP was produced (Guerreiro et al., 2015, 2016), by combining observed, modelled and other supplementary data. However, the uncertainty of this final BaP map was very high, due to the limited number of BaP measuring stations, and only limited areas of Europe could be mapped.

Trying to overcome this limitation, the use of a "pseudo stations" approach was introduced in Horálek et al. (2017). In this approach, the BaP measurement data are supplemented by estimates of BaP concentrations at the locations of $PM_{2.5}$ or PM_{10} stations with no BaP measurement. The pseudo stations approach is successfully applied in the regular $PM_{2.5}$ mapping (Denby et al., 2011; Horálek et al., 2021). The advantage of this approach is that it enlarges the number of locations with BaP data (either measured or estimated) that can be used for spatial mapping. However, although the area that could be mapped using both the true and the pseudo stations was larger than the area mapped using the true stations only, non-mapped area were still large (Horálek et al., 2017).

Since then, the number of true BaP stations was increased and the spatial resolution of the EMEP model used in the mapping was refined. This paper examines the suitability of the BaP mapping for regular annual production. In areas with lack of BaP stations, it evaluates the use of the pseudo stations. Pseudo stations based on $PM_{2.5}$ are primarily used (giving better uncertainty results compared to pseudo stations based on PM_{10}, see Horálek et al., 2017). In order to increase the spatial coverage, $PM_{2.5}$ measurements used for these estimates are complemented with $PM_{2.5}$ pseudo stations estimated at locations with PM_{10} measurements only, as used in $PM_{2.5}$ mapping (Horálek et al., 2021). Next to this, the daily BaP data is briefly checked for its potential use. The analysis is based on 2018 and 2019 data.

Chapter 2 describes the methodology applied. Chapter 3 documents the input data. Chapter 4 presents the analysis. Finally, chapter 5 provides the conclusions and recommendations.

2 Methodology

2.1 Mapping method

In agreement with Guerreiro et al. (2016), the *Regression – Interpolation – Merging Mapping (RIMM)* method as used in the regular mapping of other pollutants (Horálek et al., 2021) is used. It consists of a linear regression model followed by kriging of the residuals from that regression model (residual kriging):

$$\hat{Z}(s_0) = c + a_1 X_1(s_0) + a_2 X_2(s_0) + \cdots + a_n X_n(s_0) + \hat{\eta}(s_0) \quad (2.1)$$

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 n number of individual supplementary variables at the point s_o
$c, a_1, a_2,,\ldots, a_n$ are $n+1$ parameters of the linear regression model calculated based on the data at the points of measurement,
$\hat{\eta}(s_0)$ 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.

The spatial interpolation of the regression's residuals is carried out using ordinary kriging, according to

$$\hat{\eta}(s_i) = \sum_{i=1}^{N} \lambda_i \eta(s_i),\ \sum_{i=1}^{N} \lambda_i = 1 \quad (2.2)$$

where $\eta(s_i)$ are the residuals of the linear regression model at N points of measurement s_i, $i = 1, \ldots, N$,
$\lambda_1,\ldots, \lambda_N$ are the estimated weights based on the variogram, which is a measure of a spatial correlation, see Cressie (1993).

Prior to linear regression and interpolation, a logarithmic transformation on measurement and modelling data is applied based on the analysis presented in Guerreiro et al. (2015), as this contributes to a better fitting of the RIMM mapping. After the interpolation, a back-transformation has to be performed.

Separate map layers are created for rural and urban background areas on a grid at resolution of 1x1 km^2. The rural background map layer is based on rural background stations, the urban background map layer on urban and suburban background stations. Subsequently, the separate map layers are merged into one combined final map according to

$$\hat{Z}_F(s_0) = \left(1 - w_U(s_0)\right) \cdot \hat{Z}_R(s_0) + w_U(s_0) \cdot \hat{Z}_{UB}(s_0) \quad (2.3)$$

where $\hat{Z}_F(s_0)$ is the resulting estimated concentration in a grid cell s_o for the final map,
$\hat{Z}_R(s_0)$ and $\hat{Z}_{UB}(s_0)$ is the estimated concentration in a grid cell s_o for the rural and the urban background map layer, respectively
$w_U(s_0)$ is the weight representing the ratio of the urban character of the grid cell s_o.

The supplementary data for Eq. 2.1 have been selected by the stepwise regression, see Section 4.1.

2.2 Pseudo BaP stations

To supplement BaP measurement data, in the mapping procedure we also use data from so-called pseudo BaP stations. These data are the estimates of BaP concentrations at the locations with $PM_{2.5}$ data with no BaP measurement. These estimates are based on $PM_{2.5}$ measurement data (or $PM_{2.5}$ pseudo stations data calculated from PM_{10} stations) and different supplementary data, using exponential regression:

$$\hat{Z}_{BaP}(s) = \exp(c + b.Z_{PM_{2.5}}(s) + a_1X_1(s) + a_2X_2(s) + \cdots + a_nX_n(s)) \quad (2.4)$$

where $\hat{Z}_{BaP}(s)$ is the estimated value of BaP at the station *s*,
$Z_{PM_{2.5}}(s)$ is the measurement (or estimated) value of $PM_{2.5}$ at the station *s*,
c, *b*, a_1,,…, a_n are the parameters of the linear regression model calculated based on the data at the points of stations with both BaP and $PM_{2.5}$ measurements,
$X_1(s)$,…, $X_n(s)$ are the values of other supplementary variables at the station *s*,
n is the number of other supplementary variables used in the linear regression.

When applying this estimation method, all background stations (either classified as rural, urban or suburban) are handled together for estimating BaP values at background pseudo stations. The reason for introducing the exponential regression is the exponential relation between BaP and $PM_{2.5}$.

Equation 2.4 has been applied primarily for the locations with $PM_{2.5}$ measurement with no BaP measurements in areas which lack BaP data (see Section 4.2). In the limited areas with lack of both BaP and $PM_{2.5}$ measurements, the pseudo $PM_{2.5}$ data estimated based on the PM_{10} measurements (see Horálek et al., 2021 and 2022) have been used as an input to Equation 2.4 at the locations with PM_{10} measurements.

2.3 Uncertainty estimates of the concentration maps

The uncertainty estimation of the concentration maps is based primarily on the leave-one-out cross-validation and additionally on the interpolation standard error map, calculated according to the principles of spatial statistics (Cressie, 1993).

The leave-one-out cross-validation method computes the quality of the spatial interpolation for each point of measurement (i.e., monitoring station) from all available information except from the point in question, i.e., it withholds one data point and then makes a prediction at the spatial location of that point. This procedure is repeated for all measurement points in the available set. The technique 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 cross-validation is calculated separately for rural and urban background areas, for individual time steps. The results of the cross-validation are expressed by the statistical indicators *root mean squared error* (RMSE), *relative root mean squared error* (RRMSE) and *mean bias*:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{Z}(s_i) - Z(s_i)\right)^2}\,;\; RRMSE = \frac{RMSE}{\bar{Z}}.100;\; MB = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{Z}(s_i) - Z(s_i)\right) \quad (2.5)$$

where $Z(s_i)$ is the observed air quality indicator value at the i^{th} point,
$\hat{Z}(s_i)$ is the estimated value of the air pollution indicator at the i^{th} point using other information, except the observed indicator value at the i^{th} point,
N is the number of the observational points, for specific area type.

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 the cross-validation, only stations with BaP measurement data are used, not the pseudo stations.

The *standard error map* is calculated based on the theory of spatial statistics, see Cressie (1993). The relative standard error map is calculated by dividing the standard error by the concentration for each grid cell.

3 Input data

3.1 Air quality monitoring data

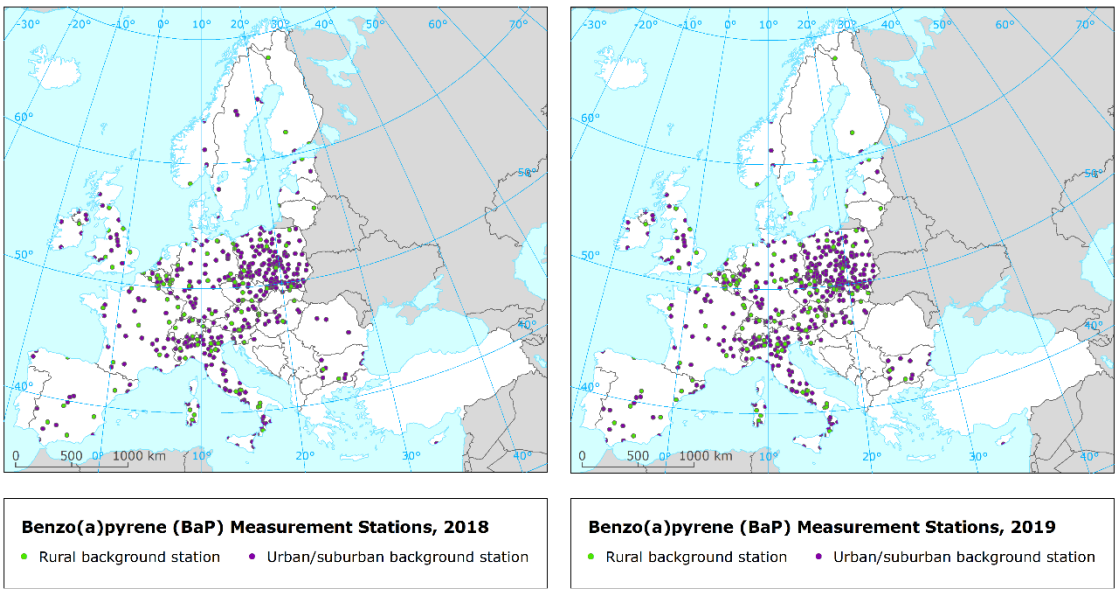
Air quality monitoring data of benzo(a)pyrene (BaP) were extracted from the Air Quality e-Reporting database, EEA (2021). Only data from stations classified as background for the areas rural, suburban and urban are used. Industrial and traffic station types are not considered, as they are deemed to represent local scale concentration levels not applicable at the mapping resolution employed. In agreement with Guerreiro et al. (2016), the following indicators of BaP concentrations in ambient air (with the e-reporting component number, cp_number) were extracted: BaP in PM_{10}, aerosol (cp_number = 5029), BaP in PM_{10}, air+aerosol (cp_number = 5129), BaP, air+aerosol (cp_number = 6015). The parameter applied is

Benzo(a)pyrene – annual average [ng.m^{-3}], years 2018 and 2019

Measurements from stations with data coverage of at least 14 percent valid measurements per year were used in order to maximise the use of the available measurement data, which are already scarce in large areas of Europe. A data coverage of 14% corresponds to the minimum time coverage for indicative measurements laid down in Directive 2004/107/EC (EC, 2004). Be it noted that EC (2004) also requires that the sampling must be spread evenly over the weekdays and the year; this requirement currently is not checked in the AQ e-Reporting database.

We have used 110 rural background and 410 urban or suburban background stations for 2018 and 110 rural background and 439 urban or suburban background stations for 2019 with officially calculated BaP annual average in EEA (2021). For the spatial distribution of the BaP measurement stations in 2018 and 2019, see Map 3.1.

Map 3.1: Spatial distribution of BaP measurement stations, 2018 (left) and 2019 (right)



We have also checked the daily BaP data as extracted from EEA (2021). We have compared this data set with the official EEA´s annual aggregations (EEA, 2021). It should be mentioned that in this daily data set, we have found additional four urban background stations in 2018 (namely PL0191A, PL0239A, PL0646A and PL0671A) and three urban background stations in 2019 (PL0239A, PL0646A and PL0671A), all with annual average values (as aggregated based on the daily data) above the level

of 10 ng.m^{-3}. It appears that these stations were dropped from the official aggregations, however the high values may be realistic. In any case, we did not use these stations in the analysis.

Next to this, we have used $PM_{2.5}$ and PM_{10} stations for the calculation of the BaP pseudo data. For details, see Horálek et al. (2021, 2022).

3.2 Chemical transport modelling data

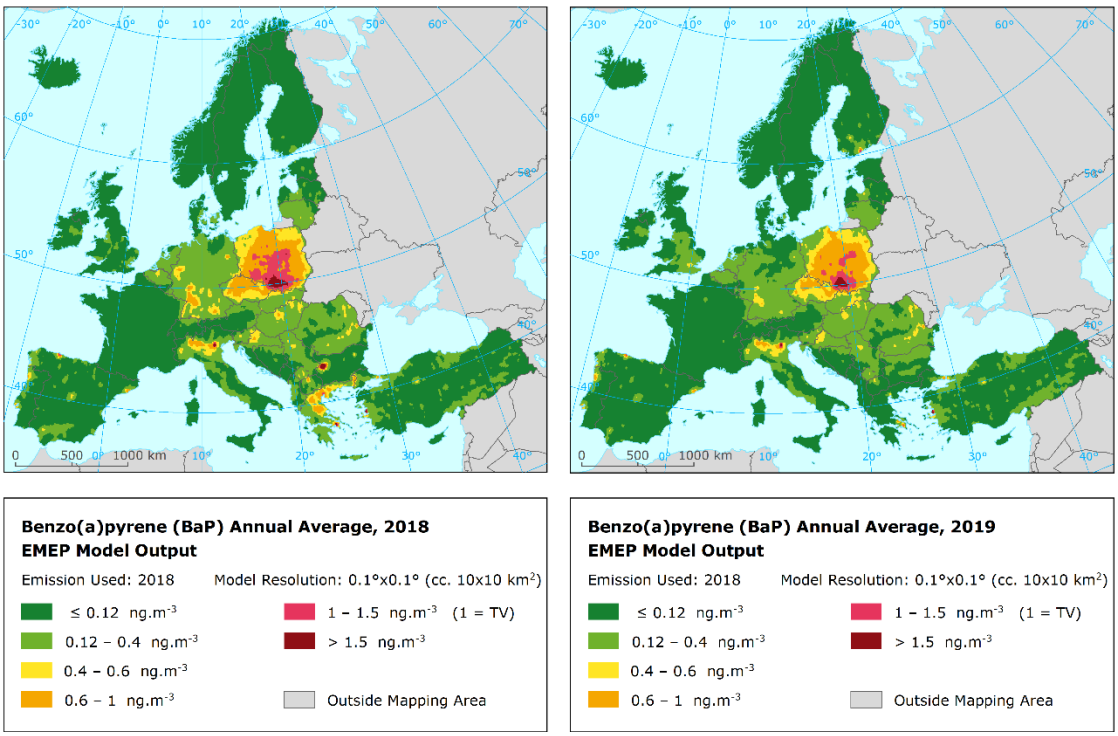
The chemical dispersion model used is the *EMEP MSC-E POP* model, EMEP (2021). It is a three-dimensional Eulerian multi-compartment chemistry transport model (Gusev et al., 2005, 2006). Its resolution is 0.1°x0.1°, i.e., circa 10x10 km^2. The model's output completely covers the mapping domain (i.e., the area of the EEA member and cooperating countries within the map extent Map_2c, EEA, 2018). The parameter used is

Benzo(a)pyrene – annual average [ng.m^{-3}], years 2018 and 2019.

Map 3.2 shows the EMEP model output for 2018 and 2019 years.

For potential regular update of the BaP maps, the availability of the modelling data is crucial. In personal communication with Alexey Gusev (EMEP, MCS-E), we were informed that the EMEP model BaP results of year Y based on both emission and meteorology of year Y are usually ready in the beginning of July of the year Y+2 (although unofficially). They receive official status after the EMEP/WGE Steering Body meeting in the middle of September (Y+2). Next to this, the EMEP MSC-E modelling team runs simulations with the meteo Y and emissions Y-1 in a year Y+1 in late autumn. The EMEP MSC-E modelling team is open for future cooperation.

Map 3.2: Output of EMEP chemical transport model, BaP annual average in 2018 (left) and 2019 (right). Units: ng.m^{-3}



3.3 Other supplementary data

Other supplementary data used are similar as in regular maps creation, Horálek et al. (2021, 2022).

Altitude

We use the altitude data field (in m) of Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010), with an original grid resolution of 15x15 arcseconds. The data comes from the U.S. Geological Survey Earth Resources Observation and Science, see Danielson and Gesch (2011). The data were converted into the EEA reference grids in 1 x 1 km^2 and 10 x 10 km^2 resolutions, as described in Horálek et al. (2021).

Meteorological data

The meteorological parameters used are *wind speed* (annual mean for 2018 and for 2019, in $m.s^{-1}$), *relative humidity* (annual mean for 2018 and for 2019, in percent), *temperature* (annual mean for 2018 and for 2019, in percent), and *surface net solar radiation* (annual mean of daily sum for 2017 and for 2019, in $MWs.m^{-2}$). The ECWMF hourly data in 0.1° x 0.1° resolution extracted from the Climate Data Store (CDS), are used (ECMWF, 2021). For details, see Horálek et al. (2021, 2022).

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 1 x 1 km^2 resolution, in the EEA reference grid.

Land cover

CORINE Land Cover 2018 (CLC2018) – grid 100 x 100 m^2, Version 2020_20 is used (EU, 2020), with additions of World Land Cover (MDA, 2015) and ESA Climate Change Initiative Global Land Cover (ESA, 2019) data resampled to 100m resolution in areas with the lack of the CLC2018 data (i.e., in Andorra, Jan Mayen and some border areas). The 44 CLC classes have been re-grouped into 8 more general classes. In this paper we use four of these general classes, see Table 3.1.

Table 3.1: General land cover classes, based on CLC2012 classes, used in BaP 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^2 grid and into a circle with radius of 5 km around each station. The reason for these two aggregations is this: 1x1 km^2 is directly related to the mapping and calculation resolution, the circle with radius of 5 km corresponds to a buffer of 5 km often used in LUR models (Hoek et al., 2008). The aggregated grid value represents for each general class the percentage of this class in the total area. For details, see Horálek et al. (2021).

4 Analysis

As a first step, we have performed the mapping analysis based on the BaP measurement data, in order to see whether the BaP data coverage is satisfactory for the mapping without the use of the pseudo stations. Then, we have performed the estimation of the pseudo data and prepared the set of the pseudo BaP data for the final mapping. Finally, we have prepared the BaP maps based on both BaP measurements and pseudo BaP data.

4.1 Initial BaP mapping

For both rural and urban background map layers, we have evaluated the selection of the supplementary data. In Guerreiro et al. (2016) and Horálek et al. (2017), the supplementary data used for the mapping based on the BaP measurement data were EMEP model output, altitude and wind speed in rural areas and EMEP model output and temperature in the urban areas. Since then, land cover data have been included in the PM mapping, leading to improved mapping results. This led us to the decision to re-evaluate the supplementary data selection, including the land cover. We have tested and mutually compared the variant without land data, labelled as (N), and the variant including the land cover, labelled (L).

The supplementary data tested for suitability of inclusion in the linear regression model included the following variables: EMEP model output, altitude, meteorological parameters (wind speed, temperature, surface solar radiation and relative humidity), population density and land cover parameters (HDR, LDR, AGR, NAT, in two aggregations: into 1x1 km^2 grid and into the circle with radius of 5 km). For selecting the optimal set of the supplementary data (in both variants with and without the land cover data), the stepwise regression and backwards elimination was applied, following Guerreiro et al. (2015) and Horálek et al. (2019).

In the rural areas, the selected variables are EMEP model output, altitude, wind speed, and temperature for the variant without the land cover (N) and EMEP model output, altitude, wind speed, temperature and the land cover parameter NAT_1km for the variant including the land cover (L). In the urban background areas, the selected variables are EMEP model output and temperature for the variant without the land cover (N) and EMEP model output, temperature, and the land cover parameters LC_LDR_1K, LC_LDR_1K and LC_NAT_5R1K for the variant including the land cover (L).

Table 4.1 presents the supplementary variables selected and applied for both rural and urban background areas, for both years 2018 and 2019. It shows also the estimated parameters of the multiple linear regression (c, a_1, a_2, …) and of the residual kriging (nugget, sill, range) and includes the statistical indicators of the regression part of the mapping.

Table 4.2 presents the mapping results of all variants for BaP annual average, with their different sets of supplementary data that are mutually compared by means of the 'leave one out' cross-validation (Section 2.3), separately for the rural background and urban background areas. The table presents the statistics of each combination of variant and type of area and provides the comparison of the variants´ performance: the green marking means the better performance. For the green highlighting, the ad hoc criterion of more than ca. 5% difference for result distinguishing has been applied.

Looking at the results, one can see quite high uncertainty in both years for both variants, especially in the rural areas. The uncertainty expressed as the relative RMSE is at the level of about 140% (in 2018) and almost 110% (in 2019) in the rural areas and about 70% (in both years) in the urban background areas. This is approximately at the level of the BaP maps presented in Guerreiro et al. (2016) and Horálek et al. (2017), however at a higher level (namely in the rural areas) compared to the 60% value defined as the data quality objective for the modelling uncertainty in the European directive (EU, 2004).

Table 4.1: Parameters of linear regression and spatial interpolation (ordinary kriging) in RIMM mapping of BaP annual average for 2018 (left) and 2019 (right) in rural and urban background areas using BaP stations and proxy data for different mapping variants

Linear Regression Model and Ordinary Kriging of residuals	2018				2019			
	(N) Without LC		(L) Using LC		(N) Without LC		(L) Using LC	
	Rural	Urb. B.	Rural	Urb. B.	Rural	Urb. B.	Rural	Urb. B.
c (constant)	3.25	2.34	3.30	0.67	1.20	2.18	1.49	0.32
a1 (log. EMEP model)	0.649	0.655	0.637	0.597	0.736	0.727	0.471	0.679
a2 (altitude)	-0.00222		-0.00177		-0.00123		-0.00072	
a3 (wind speed)	-0.508		-0.447		-0.420		-0.222	
a4 (temperature)	-0.154	-0.171	-0.162	-0.133	*n. sign.*	-0.156	-0.089	-0.116
a5 (land cover NAT_1km)			-0.0084				-0.0085	
a5 (land cover LDR_1km)				0.0055				0.0057
a5 (land cover AGR_5km_r)				0.0180				0.0197
a5 (land cover NAT_5km_r)				0.0118				0.0154
Adjusted R^2	**0.49**	**0.54**	**0.52**	**0.64**	**0.58**	**0.50**	**0.42**	**0.64**
St. Err. [$ng.m^{-3}$]	**1.10**	**0.97**	**1.07**	**0.86**	**0.94**	**0.97**	**0.948**	**0.832**
Nugget	0.300	0.100	0.333	0.120	0.288	0.192	0.465	0.230
Sill	0.843	0.840	0.795	0.745	0.780	0.777	0.763	0.617
Range [km]	455	710	462	720	339	720	720	730

Note: Grey empty cells indicate variables not used in the variant of the linear regression model.

Table 4.2: Comparison of mapping variants showing RMSE, RRMSE, bias, R^2 and regression equation from cross-validation scatter plots in rural (top) and urban background (bottom) areas for BaP annual average for 2018 (left) and 2019 (right). Units: $ng.m^{-3}$ except RRMSE and R^2

Mapping variant		2018 Rural background areas					2019 Rural background areas				
		RMSE	RRMSE	Bias	R^2	Regr. eq.	RMSE	RRMSE	Bias	R^2	Regr. eq.
(N)	Without LC	0.57	136.2%	0.10	0.436	y = 0.673x + 0.23	0.37	108.5%	0.08	0.541	y = 0.757x + 0.16
(L)	Using LC	0.61	146.1%	0.11	0.423	y = 0.728x + 0.22	0.36	107.2%	0.06	0.537	y = 0.749x + 0.14

Mapping variant		2018 Urban background areas					2019 Urban background areas				
		RMSE	RRMSE	Bias	R^2	Regr. eq.	RMSE	RRMSE	Bias	R^2	Regr. eq.
(A)	Without LC	1.08	69.2%	0.03	0.726	y = 0.790x + 0.36	0.96	71.7%	0.05	0.697	y = 0.766x + 0.36
(H)	Using LC	1.10	70.0%	0.11	0.741	y = 0.873x + 0.31	0.94	69.8%	0.10	0.729	y = 0.848x + 0.31

Note: The green marking means the better performance of the given mapping variant.

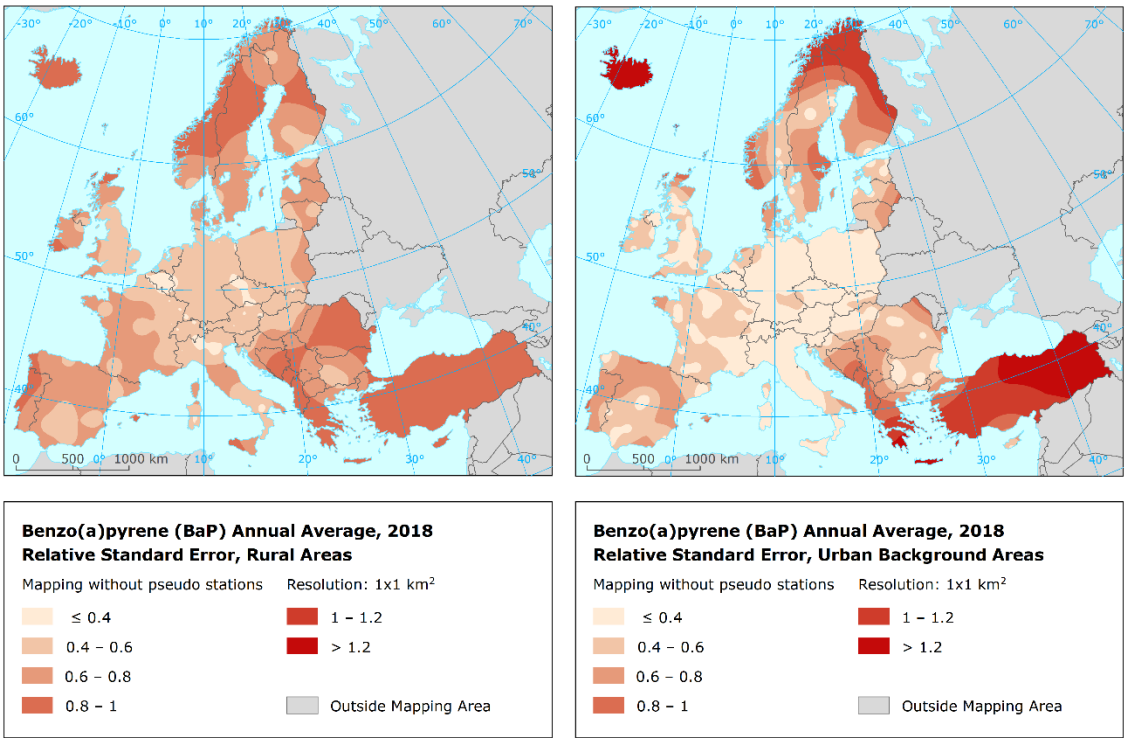
The high uncertainty in the rural areas is probably highly affected by the fact that stations classified as "rural background" comprise both regional stations with low BaP values and stations located in villages, which are often highly influenced by the local heating leading to high BaP concentrations.

Comparing the results for the variants without and with the inclusion of the land cover, both variants give quite similar results in general. In this situation, we have decided to apply the (L) variant (i.e., using the land cover variable NAT_1km) in the rural areas and to apply the (N) variant (i.e., not to add the land cover variables) in the urban background areas, for consistency with $PM_{2.5}$ and PM_{10} mapping. Thus, further in the report we use these selected variants.

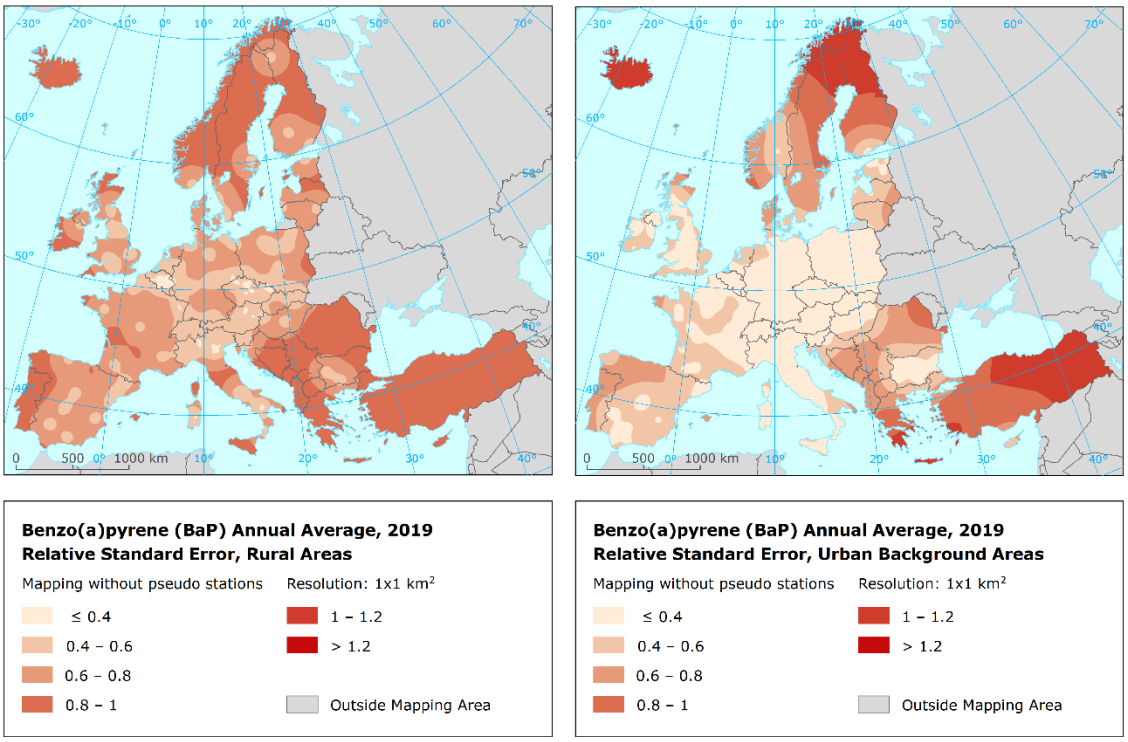
Apart from the uncertainty based on the cross-validation, we have also prepared uncertainty maps (i.e., the interpolation relative standard error maps), as described in Section 2.3. Using these maps, one can evaluate the spatial distribution of the uncertainty across the mapping domain.

Maps 4.1 and 4.2 show the uncertainty maps for 2018 and 2019, separately for rural and urban background map areas. Note that the map layers are applicable for the relevant type of area only.

Map 4.1: Uncertainty map showing interpolation relative standard error of rural (left) and urban background (right) map layers of BaP annual average 2018 as created without the use of the pseudo stations. The map layers are applicable for relevant (rural, urban) areas only



Map 4.2: Uncertainty map showing interpolation relative standard error of rural (left) and urban background (right) map layers of BaP annual average 2019 as created without the use of the pseudo stations. The map layers are applicable for relevant (rural, urban) areas only



Both 2018 and 2019 uncertainty maps show large areas with high interpolation uncertainty, mainly in rural, but also in the urban areas. One of the main reasons is the poor spatial coverage of the measurement stations. Trying to overcome this limitation, similarly as in Horálek et al. (2017), we introduce so-called pseudo BaP stations, i.e., the estimated BaP values in the points of stations with $PM_{2.5}$ data (see Section 2.2).

4.2 Pseudo BaP data

As described in Section 2.2, a nonlinear regression model (NLRM), i.e., the exponential regression of the observed BaP concentrations with the observed $PM_{2.5}$ concentrations and other supplementary data has been used to produce pseudo BaP stations, and to thus increase the station coverage for the mapping. At first, suitable supplementary data were selected. The supplementary data tested included latitude, longitude, altitude, wind speed, temperature, surface net solar radiation, relative humidity, population and land cover parameters. For both years, the supplementary data selected as most optimal are latitude, longitude and land cover parameters NAT_1km and NAT_5km_r (i.e., two variants of the natural general class). In the regression relation, the positive dependence on $PM_{2.5}$ is quite interestingly supplemented with a negative dependence on NAT_1km (i.e., natural areas at the 1x1 km^2 grid) and a positive dependence on NAT_5km_r (i.e., the floating average of the natural areas at the 5 km radius around 1x1 km^2 grid cells). On top of that, the dependence on NAT_5km_r is slightly stronger. This can be interpreted such that the correlation of BaP and $PM_{2.5}$ is the strongest in small settlements.

Table 4.3 presents the regression coefficients determined for pseudo BaP stations data estimation, based on the 262 (for 2018) and 285 (for 2019) rural and urban/suburban background that have both BaP and $PM_{2.5}$ and measurements available (see Section 2.2).

Table 4.3: Coefficients and statistics of nonlinear regression model for generation of pseudo BaP data, for BaP annual average 2018 (left) and 2019 (right)

BaP Annual Average		2018	2019
Nonlinear regression model (NLRM, Eq. 2.4)	c (constant)	-6.822	-7.599
	a1 ($PM_{2.5}$ annual average)	0.119	0.153
	a2 (latitude)	0.0821	0.0914
	a3 (longitude)	0.0357	0.0325
	a4 (land cover NAT_1km)	-0.0089	-0.0057
	a5 (land cover NAT_5km_r)	0.0116	0.0104
	Adjusted R^2	**0.81**	**0.74**

Looking at the parameters of the regression, one can note that the R^2 of 0.81 and 0.74 is a relatively poor correlation. Based on this, it can be recommended to use the pseudo stations only in areas with a significant lack of the BaP measurements.

Keeping in mind this recommendation and based on Maps 4.1 and 4.2, we have selected areas for using the pseudo BaP stations. As a rough criterion, we chose a value of the relative uncertainty of more than 0.6 for the given area (in average in both years). We have selected the countries and areas for use the pseudo stations, as follows.

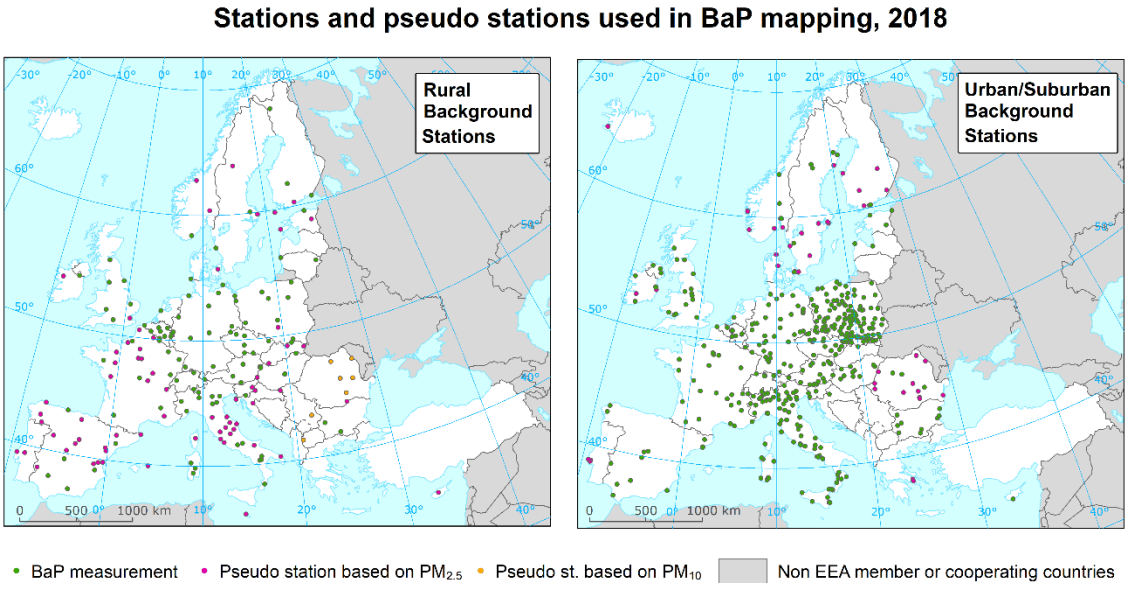
For the rural areas: All the mapping area, apart from Austria, Benelux, Czechia, Germany, Poland, Slovakia and Switzerland. For the urban background areas: Iceland, Ireland, Portugal, Scandinavia (including Denmark), Italy south of 45 degrees latitude, and most of the Balkan countries (namely, Albania, Bosnia and Herzegovina, Greece, Montenegro, Northern Macedonia, Romania, and Serbia including Kosovo). We have decided to exclude Turkey from the mapping area, due to the fact that there are no BaP data from Turkey in the AQ e-reporting database and that a $PM_{2.5}$ map itself cannot be constructed for the area of Turkey, due to non-sufficient $PM_{2.5}$ data (Horálek et al., 2021).

After the addition of the pseudo stations calculated based on the $PM_{2.5}$ measurements, there is still a lack of rural stations in the area of Bosnia and Herzegovina, Northern Macedonia, Romania and Serbia. For these countries, we have added the pseudo BaP stations calculated based on the pseudo $PM_{2.5}$ data (as estimated based on the PM_{10} measurements, see Horálek et al., 2021, 2022).

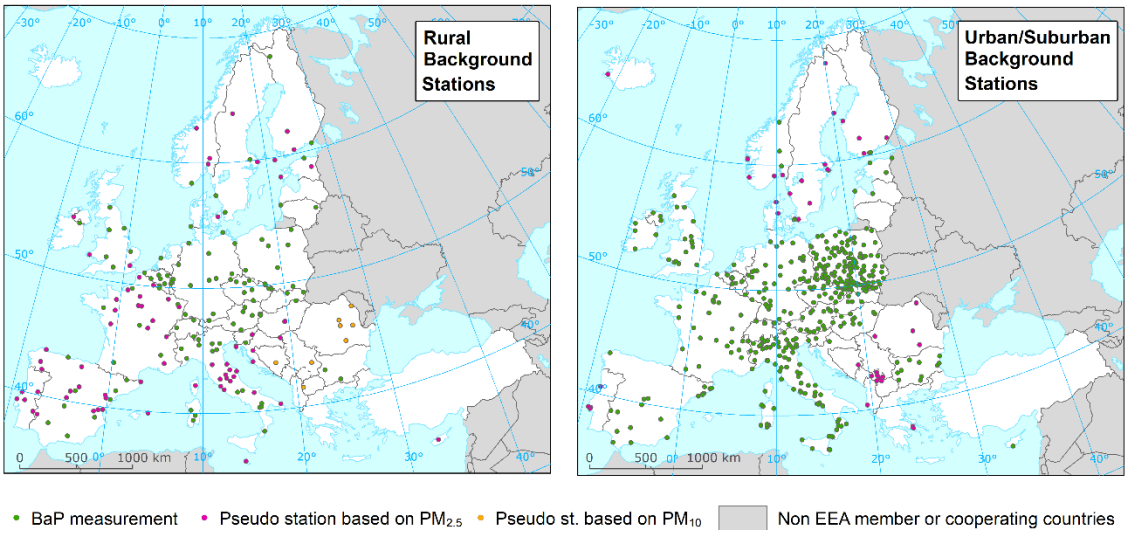
In total, for mapping of the year 2018, we have used 70 rural and 48 urban background pseudo BaP stations calculated based on the $PM_{2.5}$ measurements and 7 rural pseudo BaP stations calculated based on the pseudo $PM_{2.5}$ data (as estimated based on the PM_{10} measurements). For mapping the year 2019, we have used 71 rural and 48 urban background pseudo BaP stations calculated based on the $PM_{2.5}$ measurements and 8 rural pseudo BaP stations calculated based on the pseudo $PM_{2.5}$ data.

Maps 4.3 and 4.4 show the spatial distribution of the rural and urban background stations and pseudo stations as used in the mapping of BaP annual averages 2018 and 2019.

Map 4.3: Spatial distribution of BaP stations and pseudo stations rural (left) and urban/suburban (right) background stations used in mapping, 2018



Map 4.4: Spatial distribution of BaP stations and pseudo stations rural (left) and urban/suburban (right) background stations used in mapping, 2019



4.3 Final BaP mapping

Based on the data of both BaP stations and pseudo stations (see Section 4.2), final BaP maps have been prepared using the RIMM methodology.

Table 4.4 presents for both rural and urban background areas the estimated parameters of the linear regression models (c, a_1, a_2,…) and of the residual kriging (nugget, sill, range) and includes the statistical indicators of the linear regression.

Table 4.4: Parameters of linear regression and spatial interpolation (ordinary kriging) in RIMM mapping of BaP annual average for 2018 (left) and 2019 (right) in rural and urban background areas using BaP stations and pseudo stations, together with proxy data

BaP Annual Average		2018		2019	
		Rural Areas	Urban B. Areas	Rural Areas	Urban B. Areas
Linear regresion model (LRM, Eq. 2.1)	c (constant)	1.93	2.50	1.49	2.38
	a1 (log. EMEP model)	0.394	0.620	0.471	0.629
	a2 (altitude GMTED)	-0.00124		-0.00072	
	a3 (wind speed)	-0.276		-0.222	
	a4 (temperature)	-0.105	-0.189	-0.089	-0.181
	a5 (land cover NAT_1km)	-0.0086		-0.009	
	Adjusted R^2	**0.39**	**0.52**	**0.42**	**0.46**
	Standard Error [µg.m^{-3}]	**1.05**	**0.95**	**0.95**	**0.99**
Ordinary kriging (OK) of LRM residuals (Eq. 2.2)	nugget	0.55	0.12	0.46	0.18
	sill	0.82	0.82	0.76	0.80
	range [km]	720	720	720	720

Note: Grey empty cells indicate variables not used in the variant of the linear regression model.

Table 4.5 shows the cross-validation mapping results of BaP annual average 2018 and 2019, indicated separately for the rural background and urban background areas.

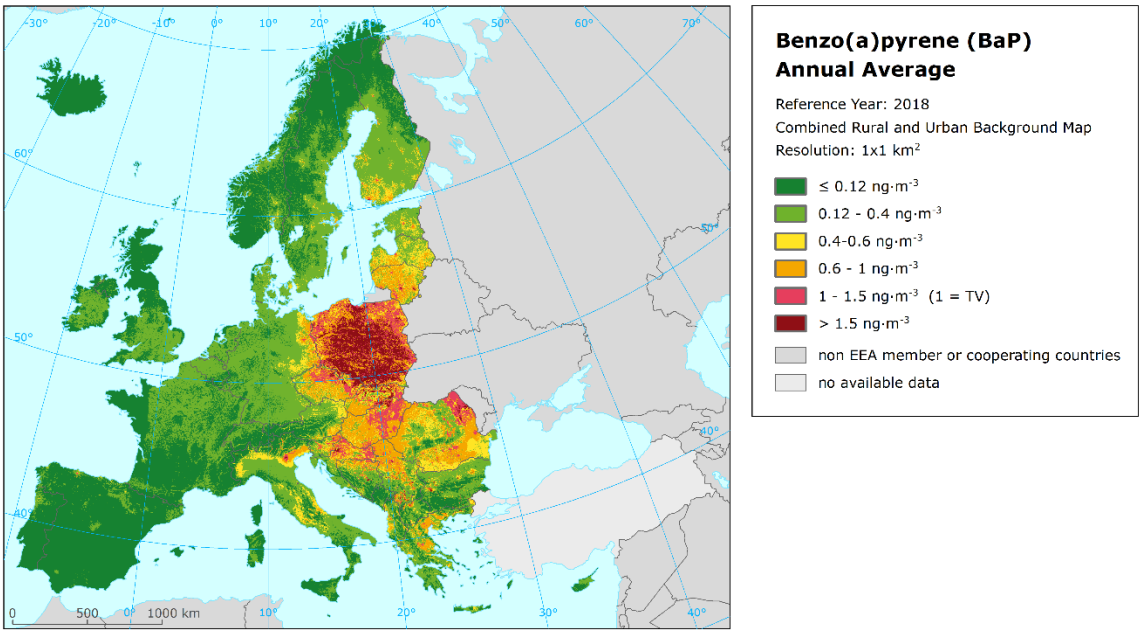
Table 4.5: Comparison of mapping variants showing RMSE, RRMSE, bias, R^2 and regression equation from cross-validation scatter plots in rural (top) and urban background (bottom) areas for BaP annual average for 2018 (left) and 2019 (right). Units: ng.m^{-3} except RRMSE and R^2

Area Type	2018					2019				
	RMSE	RRMSE	Bias	R^2	Regr. eq.	RMSE	RRMSE	Bias	R^2	Regr. eq.
Rural Areas	0.52	123.7%	0.11	0.508	y = 0.693x + 0.19	0.34	100.9%	0.06	0.536	y = 0.628x + 0.19
Urban B. Areas	1.07	68.1%	0.04	0.734	y = 0.791x + 0.37	0.96	71.4%	0.04	0.698	y = 0.759x + 0.36

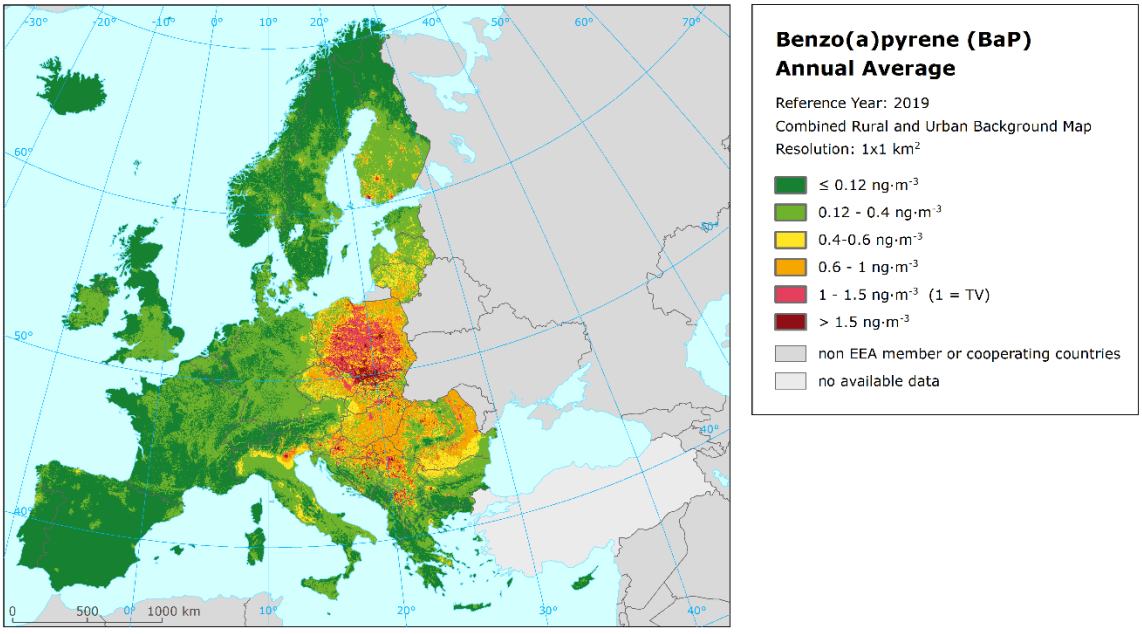
Looking at the cross-validation results, one can see slightly better fit for the rural areas and the similar fit for the urban background areas, compared with the mapping performed without the use of the pseudo stations (see Table 4.2). The cross-validation relative uncertainty RRMSE is still at the higher level (namely in the rural areas) compared to the 60%, being the data quality objective for the modelling uncertainty in the European directive (EC, 2004).

Maps 4.5 and 4.6 present the final maps of the BaP annual average for 2018 and 2019 years, where red and purple areas indicate exceedances of 1.0 ng·m^{-3} (EC, 2004). The highest BaP concentrations are shown in Poland, north-eastern Czechia and some populated locations in the central, and eastern Europe and the eastern Po Valley in northern Italy. Contrary to that, western Europe show low BaP values in both years. In Poland, eastern Hungary and north-eastern Romania, a decrease from 2018 to 2019 is shown. In the maps, relative "noise" is caused by the use of the land cover parameter NAT_1km and indicates lower BaP values in the natural areas.

Map 4.5: Concentration map of BaP annual average, rural and urban background merged map using pseudo BaP stations in areas with a poor BaP data coverage, 2018. Units: ng·m^{-3}

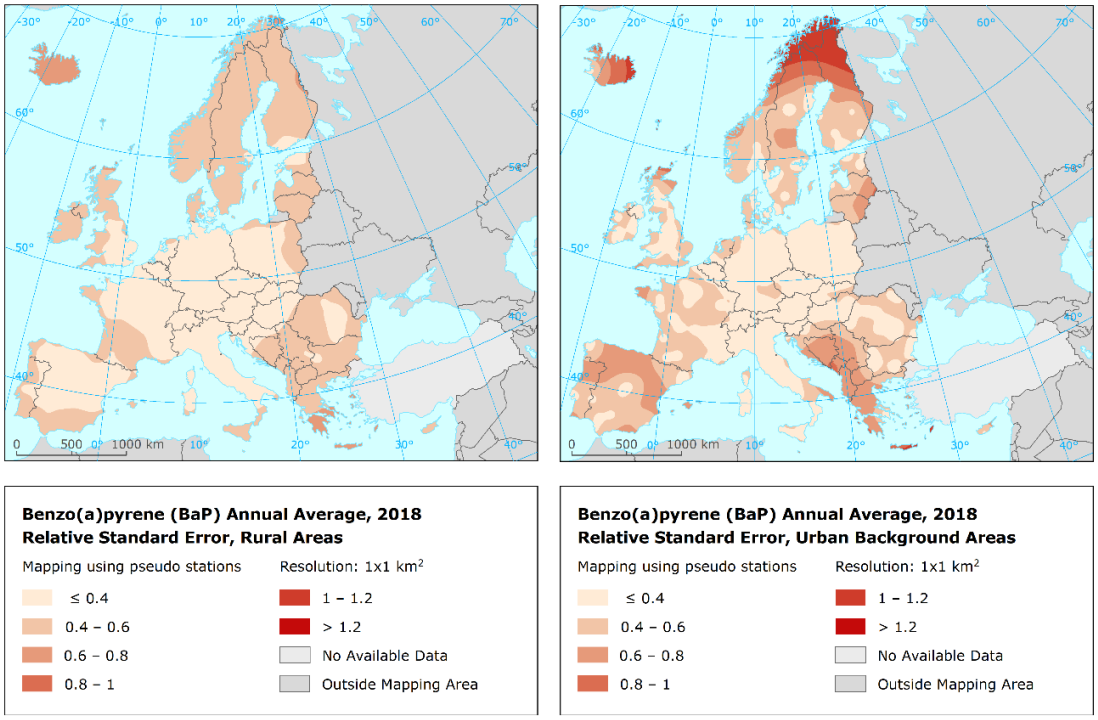


Map 4.6: Concentration map of BaP annual average, rural and urban background merged map using pseudo BaP stations in areas with a poor BaP data coverage, 2019. Units: ng.m^{-3}

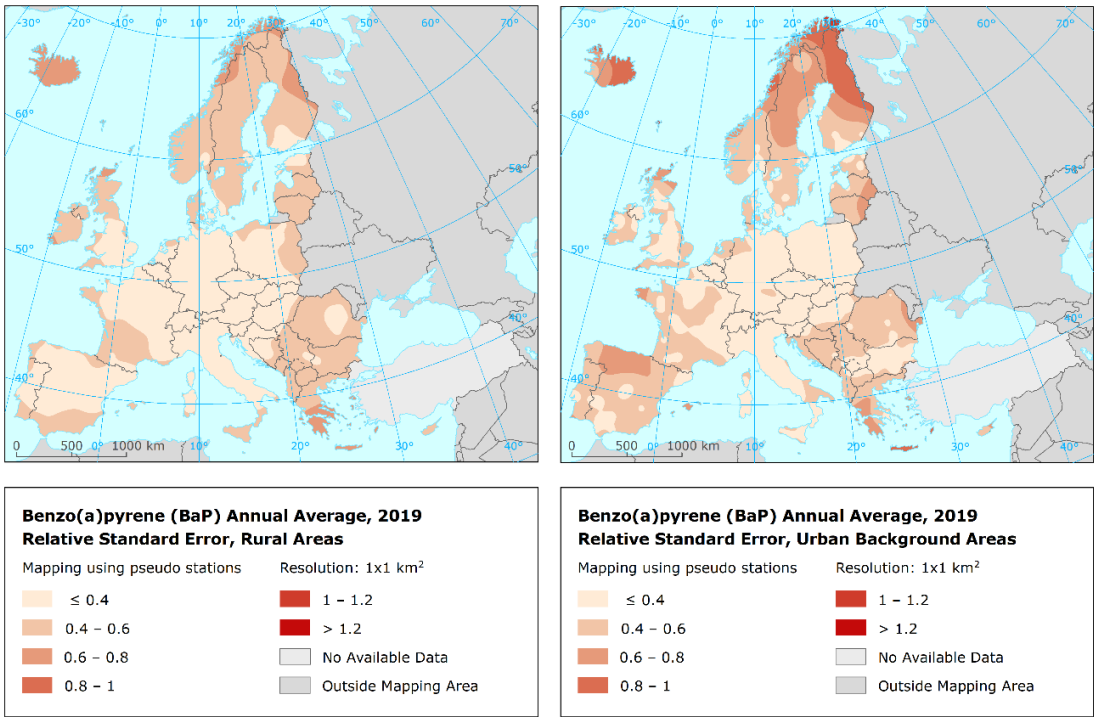


Maps 4.7 and 4.8 present the uncertainty maps (i.e., the interpolation relative standard error maps as described in Section 2.3) for 2018 and 2019, separately for rural and urban background map areas. Comparing these maps with Maps 4.1 and 4.2 showing the interpolation uncertainty for the maps created without the use of the pseudo stations, one can see an improvement in terms of the interpolation uncertainty for maps constructed with inclusion of the pseudo stations. This confirms the applicability of the pseudo stations for the BaP mapping.

Map 4.7: Uncertainty map showing interpolation relative standard error of rural (left) and urban background (right) map layers of BaP annual average 2018 as created using pseudo stations in areas with a poor BaP data coverage. The map layers are applicable for relevant (rural, urban) areas only



Map 4.8: Uncertainty map showing interpolation relative standard error of rural (left) and urban background (right) map layers of BaP annual average 2019 as created using pseudo stations in areas with a poor BaP data coverage. The map layers are applicable for relevant areas only



In addition, it can be stated that for the majority of the mapping area, the interpolation uncertainty is lower than the data quality objective of 60% for the modelling uncertainty in the directive EC (2004). In limited areas, the interpolation uncertainty is higher that 60%, specifically, in Iceland, in the urban areas in the northern Scandinavia and in the urban areas in the north of Spain. For Iceland and the northern Scandinavia, the level of the relative uncertainty is highly influenced by low absolute BaP values, so we do not consider this elevated uncertainty should prevent the BaP mapping in these areas. Concerning the north of Spain, for the future we recommend the use of the pseudo stations in this area, in order to decrease the interpolation uncertainty.

Bearing in mind both an improvement of the interpolation uncertainty compared to the BaP maps presented in Guerreiro et al. (2016) and Horálek et al. (2017) and a still rather poor cross-validation uncertainty (namely in the rural areas), we **recommend the regular production of the BaP maps** (if all data needed is available), **however, as experimental only**. On the one hand, we consider that the regular provision of a spatial perspective of BaP at the European scale is useful, even despite relatively high uncertainties, because they show spatial patterns, not so influenced by the high uncertainty; the same is true also for the inter-annual differences. Next to this, a lower level of the relative uncertainty in the urban areas enables to estimate a general population exposure. On the other hand, the maps should be considered as influenced by a higher level of uncertainty in the rural areas. In the year-to-year mode, the level of their uncertainty should be regularly evaluated.

When thinking about a way forward, we have briefly checked the potential usability of the urban and suburban traffic stations in the BaP mapping. As traffic (contrary to the domestic heating) is not a large source of the BaP emission in most European countries, we thought about using the traffic stations in the urban map layer creation. In order to examine this, we have briefly compared the BaP measurement data from the urban and suburban traffic stations with the underlying grid values of the urban background layer. By this comparison, it was verified that the urban/suburban traffic stations do not show on average higher values than the underlying background air quality. This confirms the results of studies from Finland (Hellén et al., 2017) and from Czechia (Schreiberová et al., 2020) indicating that local traffic has only a minor effect on BaP concentrations, compared with the corresponding effect of domestic heating. (E.g., in street canyons of Helsinki, the measured concentrations of BaP were at the same level as those in the urban background, clearly lower than those in suburban detached-house areas.) Thus, we recommend to test the use of the urban and suburban traffic stations together with the urban and suburban background stations in the creation of the urban map layer.

5 Conclusions and recommendations

The paper examines the potential regular production of BaP maps at the European scale. The analysis has been executed based on 2018 and 2019 data.

Potential mapping based on the BaP measurement data only (together with the supplementary data) has been examined at first. However, such maps show large areas with high interpolation uncertainty, mainly in rural, but also in the urban areas. One of the main reasons is the poor spatial coverage of the measurement stations. Trying to overcome this limitation, we have introduced so-called pseudo BaP stations, similarly as in Horálek et al. (2017).

Pseudo BaP data in locations with $PM_{2.5}$ measurements (or with pseudo $PM_{2.5}$ data based on PM_{10} measurements) and with no BaP measurements have been estimated, based on the exponential regression of the observed BaP concentrations with the $PM_{2.5}$ data, geographical coordinates and the land cover (specifically, the general class natural areas in two resolutions). Due to quite high uncertainty of the pseudo data estimates, we have further used the pseudo data only in areas with a lack of BaP measurements. Due to the serious lack of Turkish data, we have decided not to inlude Turkey in the mapping area.

Based on both actual BaP measurements and pseudo data (in areas with a lack of BaP stations), we have prepared the final BaP maps. The uncertainty of the final maps were evaluated by the leave-one-out cross-validation (showing the total uncertainty of the map) and by the relative standard uncertainty of the interpolation (showing the spatial uncertainty). The cross-validation uncertainty (expressed as the relative RMSE) has been estimated at a level of about 120% (for 2018) and 100% (for 2019) in the rural areas and about 70% (in both years) in the urban areas. This is at somewhat lower level in rural areas and a similar level in the urban areas, compared to BaP maps presented in Guerreiro et al. (2015) and Horálek et al. (2017). However, it is still at a higher level (namely in the rural areas) compared to the threshold of 60%, which is defined as the data quality objective for modelling uncertainty in the European directive (EU, 2004). The high uncertainty in the rural areas is probably highly affected (besides the low density of the rural stations) by the fact that stations classified as “rural background” comprise both regional stations with low BaP values and stations located in villages, which are often highly influenced by the local heating leading to high BaP concentrations. Next to the cross-validation, the interpolation relative standard uncertainty of the final maps was also estimated. The final BaP maps based on both the true and the pseudo stations show satisfactory (i.e., smaller than 60%) interpolation uncertainty in the most of the mapping area. This improves the results of both Guerreiro et al. (2016) and Horálek et al. (2017).

Based on the analysis results, we recommend the regular production of the BaP maps (if all data needed is available). However, due to the relatively high cross-validation uncertainty (particularly in the rural areas), we recommend to label them as experimental maps to indicate that they do not yet meet the same accuracy standards as the regularly produced maps of other pollutants. (Doing this, the uncertainty of these maps should be regularly evaluated.) Based on the information of the EMEP MSC-E modelling team, the availability of the EMEP modelling data can be anticipated. Thus, we suppose it will be possible to construct the BaP maps regularly. We consider the regular BaP mapping as useful despite higher uncertainties, in order to show spatial patterns of this pollutant.

In order to increase the quality of the urban map layer, the BaP measurement data from the traffic stations might be utilized. We recommend to test the use of the urban and suburban traffic stations together with the urban and suburban background stations in the creation of the urban map layer.

6 References

Cressie, N., 1993, *Statistics for spatial data,* Wiley series, New York.

Danielson, J. J. and Gesch, D. B., 2011, *Global multi-resolution terrain elevation data 2010 (GMTED2010)*, U.S. Geological Survey Open-File Report, pp. 2011-1073 (https://pubs.er.usgs.gov/publication/ofr20111073) accessed 19 November 2020.

Denby, B., et al., 2011, *Mapping annual mean $PM_{2.5}$ concentrations in Europe: application of pseudo $PM_{2.5}$ station data*, ETC/ACM Technical Paper 2011/5 (http://www.eionet.europa.eu/etcs/etc-atni/products/etc-atni-reports/etcacm_tp_2011_5_spatialpm2-5mapping) accessed 26 August 2020.

EC, 2004, 'Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 relating to arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air', OJ L 23, 26.1.2005, pp. 3–16.

ECMWF, 2021, Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/home). Data extracted in March 2021.

EEA, 2018, *Guide for EEA map layout. EEA operational guidelines*, January 2015, version 5 (https://www.eionet.europa.eu/gis/docs/GISguide_v5_EEA_Layout_for_map_production.pdf) accessed 26 August 2020.

EEA, 2020, *Air quality in Europe – 2020 Report*, EEA Report 9/2020 (https://www.eea.europa.eu/publications/air-quality-in-europe-2020-report) accessed 8 December 2020.

EEA, 2021, *Air Quality e-Reporting. Air quality database (*https://www.eea.europa.eu/data-and-maps/data/aqereporting-8). Data extracted in March and June 2021.

EMEP, 2021, Data of HMs and POPs for the EMEP region, Meteorological Synthetizing Centre – East (MCS-E) (https://www.msceast.org/index.php/pollution-assessment/emep-domain-menu/data-hm-pop-menu) accessed 30 August 2021.

ESA, 2019, *Land cover classification gridded maps from 1992 to present derived from satellite observations*, (https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover) accessed 17 February 2021.

EU, 2020, *Corine land cover 2018 (CLC2018) raster data*, 100x100m^2 gridded version 2020_20 (https://land.copernicus.eu/pan-european/corine-land-cover/clc2018) accessed 19 November 2020.

Eurostat, 2014, *GEOSTAT 2011 grid dataset*, Population distribution dataset (http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography) accessed 26 August 2020.

Guerreiro, C., et. al., 2015, *Mapping ambient concentrations of benzo(a)pyrene in Europe*, ETC/ACM Technical paper 2014/6 (https://www.eionet.europa.eu/etcs/etc-atni/products/etc-atni-reports/etcacm_tp_2014_6_bap_hia) accessed 2 December 2021.

Guerreiro, C., et al., 2016, 'Benzo(a)pyrene in Europe: Ambient air concentrations, population exposure and health effects, *Environmental Pollution* 214, 657–667 (https://doi.org/10.1016/j.envpol.2016.04.081) accessed 9 December 2021.

Gusev, A., et al., 2005, *Regional Multicompartment Model MSCE-POP*, EMEP/MSC-E Technical Report 5/2005 (http://www.msceast.org/reports/5_2005.pdf) accessed 9 December 2021.

Gusev, A., et al., 2006, *Progress in further development of MSCE-HM and MSCE-POP models (implementation of the model review recommendations*, EMEP/MSC-E Technical Report 4/2006 (http://www.msceast.org/reports/4_2006.pdf) accessed 10 December 2021.

Hellén, H., et al., 2017, 'Evaluation of the impact of wood combustion on benzo[a]pyrene (BaP) concentrations; ambient measurements and dispersion modeling in Helsinki, Finland', *Atmospheric Chemistry and Physics* 17, pp. 3475–3487 (https://doi.org/10.5194/acp-17-3475-2017) accessed 10 December 2021.

Hoek, G., et al., 2008, 'A review of land-use regression models to assess spatial variation of outdoor air pollution', *Atmospheric Environment* 42, pp. 7561–7578 (https://doi.org/10.1016/j.atmosenv.2008.05.057) accessed 9 December 2021.

Horálek, J., et. al., 2017, *Potential improvements on benzo(a)pyrene (BaP) mapping*, ETC/ACM Technical Paper 2016/3 (http://www.eionet.europa.eu/etcs/etc-atni/products/etc-atni-reports/etcacm_tp_2016_3_bap_improved_mapping) accessed 27 May 2021.

Horálek, J., et al., 2019, *Land cover and traffic data inclusion in PM mapping*, Eionet Report ETC/ACM 2018/18 (http://www.eionet.europa.eu/etcs/etc-atni/products/etc-atni-reports/etc-acm-report-18-2018-land-cover-and-traffic-data-inclusion-in-pm-mapping) accessed 27 May 2021.

Horálek, J., et. al., 2021, *European air quality maps for 2018*, Eionet Report ETC/ATNI 2020/10 (https://doi.org/10.5281/zenodo.4638651) accessed 27 May 2021.

Horálek, J., et. al., 2022, *European air quality maps for 2019*, Eionet Report ETC/ATNI 2021/1, *in preparation*.

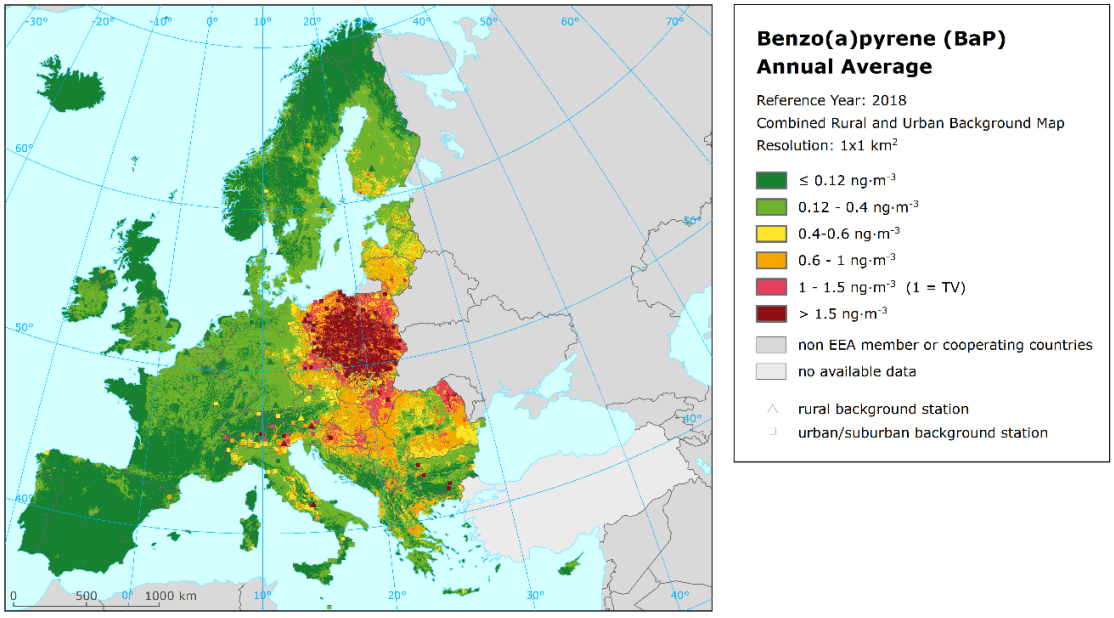
MDA, 2015, *World Land Cover at 30m resolution from MDAUS BaseVue 2013* (https://www.arcgis.com/home/item.html?id=1770449f11df418db482a14df4ac26eb) accessed 17 Fegruary 2021.

Schreiberová, M., et al., 2020, 'Benzo[a]pyrene in the Ambient Air in the Czech Republic: Emission Sources, Current and Long-Term Monitoring Analysis and Human Exposure', *Atmosphere* 11, 955. (https://doi.org/10.3390/atmos11090955) accessed 10 December 2021.

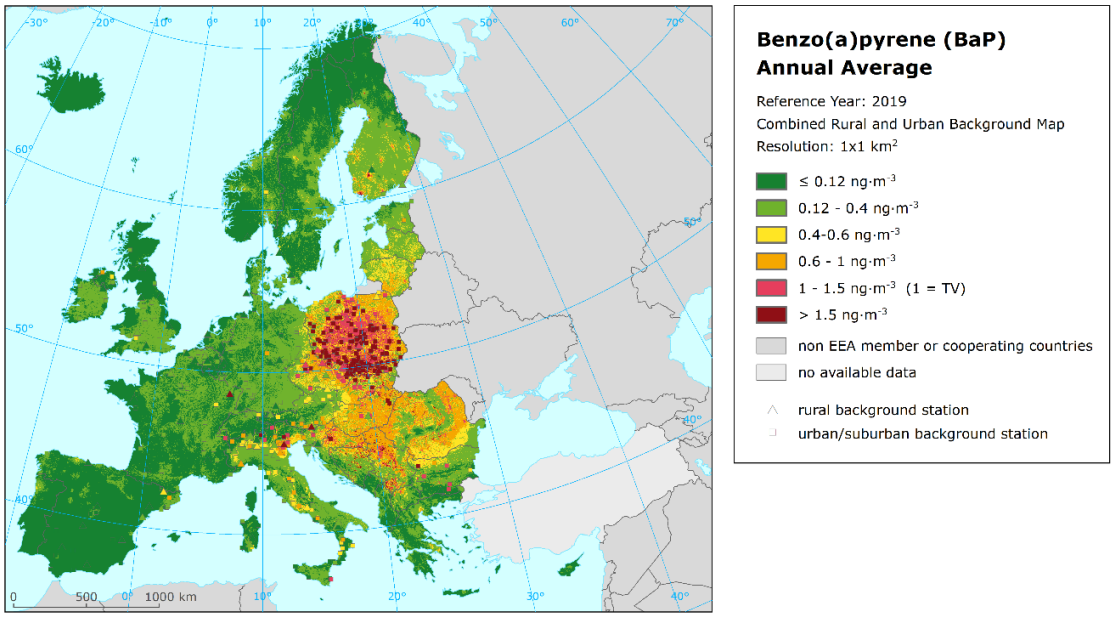
Annex
Concentration maps including station points

In the main report, the BaP maps presented do not include station points. The reason is to better visualise the maps. However, the interpolation somewhat smooths the concentration field. Thus, it is valuable to present in this Annex the maps *including* the BaP values resulting from the measurement data at the station points (without including the pseudo stations). Maps A.1 and A.2 present BaP annual average 2018 and 2019 and include the stations points used in the mapping. They correspond to Maps 4.5 and 4.6 of the main report.

Map A.1: Concentration map of BaP annual average including station points, 2018. Units: ng·m^{-3}



Map A.2: Concentration map of BaP annual average including station points, 2019. Units: ng·m^{-3}



European Topic Centre on Air pollution,
transport, noise and industrial pollution
c/o NILU – Norwegian Institute for Air Research
P.O. Box 100, NO-2027 Kjeller, Norway
Tel.: +47 63 89 80 00
Email: etc.atni@nilu.no
Web : https://www.eionet.europa.eu/etcs/etc-atni

The European Topic Centre on Air pollution, transport, noise and industrial pollution (ETC/ATNI) is a consortium of European institutes under a framework partnership contract to the European Environment Agency.

European Environment Agency
European Topic Centre on Air pollution,
transport, noise and industrial pollution