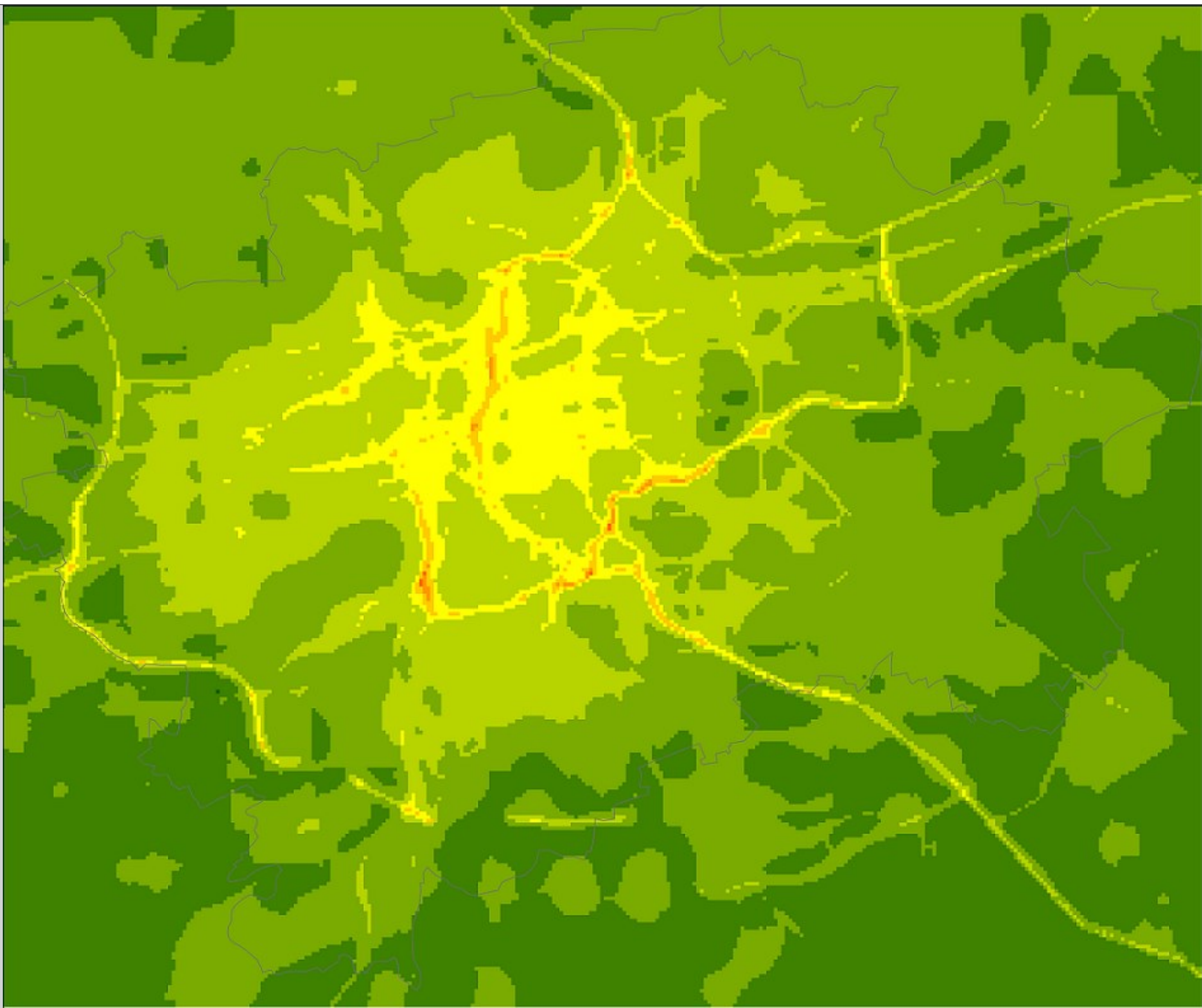
ETC HE Report 2022/24

City-level mapping of air quality at fine spatial resolution – the Prague case study

NO$_2$, PM$_{10}$ and PM$_{2.5}$ maps on a 100 m spatial grid



Authors:
Jan Horálek (CHMI), Daša Damašková (CHMI), Philipp Schneider (NILU), Pavel Kurfürst (CHMI), Markéta Schreiberová (CHMI), Ondřej Vlček (CHMI)

European Environment Agency
European Topic Centre
Human health and the environment

Cover design: EEA
Cover picture: Fine resolution concentration map of NO_2 annual average 2020, Prague. Units: µg/m³. (Map 3.1 of this report.)
Layout: EEA / ETC HE / CHMI

Publication Date: July 2023
ISBN 978-82-93970-27-9

How to cite this report:
Horálek, J., Damašková, D., Schneider, P., Kurfürst, P., Schneiderová, M., Vlček , O. (2023). City-level mapping of air quality at fine spatial resolution – the Prague case study (Eionet Report – ETC HE 2022/24). European Topic Centre on Human Health and the Environment.

The report is available from https://www.eionet.europa.eu/etcs/all-etc-reports and https://zenodo.org/communities/eea-etc/?page=1&size=20.

ETC HE coordinator: NILU - Stiftelsen Norsk institutt for luftforskning (NILU - Norwegian Institute for Air Research)

ETC-HE consortium partners: Federal Environment Agency/Umweltbundesamt (UBA), Aether Limited,
Czech Hydrometeorological Institute (CHMI), Institut National de l'Environnement Industriel et des Risques (INERIS),
Swiss Tropical and Public Health Institute (Swiss TPH), Universitat Autònoma de Barcelona (UAB),
Vlaamse Instelling voor Technologisch Onderzoek (VITO), 4sfera Innova S.L.U., klarFAKTe.U

European Topic Centre on
Human health and the environment (ETC HE)
https://www.eionet.europa.eu/etcs/etc-he

Contents

Contents 3
Acknowledgements 4
Summary 4
1 Introduction 5
2 Methodology and data used 6
2.1 Methodology 6
2.1.1 Statistical downscaling 6
2.1.2 Comparison and validation of the mapping results 6
2.1.3 Population exposure 7
2.2 Data used 7
2.2.1 Original maps in coarse resolution 7
2.2.2 Chemical transport modelling (CTM) output in fine resolution 7
2.2.3 Air quality monitoring data 8
2.2.4 Population density data 9
3 Analysis and map creation 10
3.1 NO_2 annual average 10
3.2 PM_{10} annual average 15
3.3 $PM_{2.5}$ annual average 18
4 Conclusion 23
List of abbreviations 24
References 25

Acknowledgements

The ETC task manager was Núria Blanes (UAB, Spain). The EEA task managers were Eulalia Peris and Artur Gsella. The contributors of the report were Jan Horálek, Daša Damašková, Pavel Kurfürst, Markéta Schreiberová and Ondřej Vlček (all CHMI, Czechia) and Philipp Schneider (NILU, Norway). The reviewer of the report was Vladimíra Volná (CHMI). The chemical transport modelling output in fine resolution was provided by Ondřej Vlček (CHMI, Czechia). The regridding of this model output was executed by Jana Schovánková (CHMI, Czechia).

This work has been partly funded by the Czech national project ARAMIS – Integrated system of research, evaluation and control of air quality (co-funded with the state support of the Technology Agency of the CR under the Environment for Life program).

Summary

This paper examines the creation of fine resolution maps at 100 m x 100 m resolution using statistical downscaling for the area of Prague, as a case study. This Czech city was selected due to the fine resolution proxy data available for this city. The reference downscaling methodology used is the linear regression and the interpolation of its residuals by the area-to-point kriging. Next to this, several other methods of statistical downscaling have been also executed. The results of different downscaling methods have been compared mutually and against the data from the monitoring stations of Prague, separately for urban background and traffic areas. In addition, the population exposure estimates based on the downscaled mapping results have been also calculated.

The downscaled maps in 100 m x 100 m resolution have been constructed for three pollutants, namely for NO_2, PM_{10} and $PM_{2.5}$. In the maps, one can see overall realistic spatial patterns, for example the main roads in Prague are visible through higher air pollution levels. This is distinct especially for NO_2, while for PM_{10} and $PM_{2.5}$ the differences between road increments and urban background are smaller as would be expected. The results of the case study for Prague have proven the usefulness of the statistical downscaling for the air quality mapping, especially for NO_2.

1 Introduction

The European-wide air quality maps (Horálek et al., 2022b and references therein) as routinely produced using the Regression – Interpolation – Merging Mapping (RIMM) methodology under the ETC HE consortium (and its predecessors) are constructed at 1 km by 1 km resolution for health-related pollutants and their indicators. However, a need of a finer spatial resolution compared to the current grid has recently emerged. More specifically, for the purposes of the integrated assessment of noise and air quality in Europe, consistent air quality maps for European cities on a 100 metres by 100 metres grid (i.e., in the same resolution as the noise maps) are required. In Horálek et al. (2022a), several possibilities of future development towards European-wide city level mapping at a fine spatial resolution have been suggested.

This paper examines the creation of fine resolution maps at 100 m x 100 m resolution using the method of statistical downscaling for the area of Prague, as a case study. This Czech city was selected due to the fine resolution proxy data available for this city. In statistical downscaling, the spatial resolution of a coarse-resolution source dataset is increased to finer spatial resolution using additional information from a spatial proxy dataset available at a fine spatial resolution, which is correlated with the original coarse-resolution source dataset. As the input dataset at coarse spatial resolution, we used here the regular RIMM maps in 1 km x 1 km resolution (Horálek et al., 2022b). As the proxy dataset in fine spatial resolution, chemical transport modelling (CTM) output for Prague in 100 m x 100 m resolution has been used, namely the combination of the Eulerian model CAMx (ENVIRON, 2011) and the Gaussian model SYMOS (CHMI, 2016) outputs.

As a reference downscaling methodology, the combination of linear regression and geostatistical area-to-point kriging as used in Stebel et al. (2021) has been applied, in agreement with the proposal of Horálek et al. (2022a). Additionally, several other simpler approaches has been examined. The results have been compared mutually and with the routine maps in 1 km x 1 km resolution based on the air quality measurement data from monitoring stations located in Prague, separately for the urban background and urban traffic areas. The analysis has been performed for the 2020 annual averages of NO_2, PM_{10}, and $PM_{2.5}$.

Next to the mapping results in different variants, the population exposure estimates based on these downscaled mapping results have been calculated and mutually compared.

Chapter 2 describes the methodology and the data used. Chapter 3 shows the analysis and the final fine resolution maps for Prague. Chapter 4 presents conclusions.

2 Methodology and data used

2.1 Methodology

2.1.1 Statistical downscaling

For preparing the fine resolution maps, statistical downscaling using a linear regression model and interpolation of its residuals has been applied, similarly as in Stebel et al. (2021) and Park (2013). This methodology increases the spatial resolution of a coarse source dataset using spatial proxy or auxiliary datasets that are available at a fine spatial resolution and that are to some extent correlated with the source dataset. The parameters of the linear regression model (LRM) are estimated at the coarse resolution: the source dataset in this resolution is a dependent variable, while the proxy datasets spatially aggregated to the same coarse resolution are the independent variables. As the reference downscaling method for the interpolation of the LRM residuals, area-to point kriging (Kyriakidis, 2004) is used. An alternative simpler method for the interpolation of the LRM residuals is bilinear interpolation (ESRI, 2021). Bilinear interpolation calculates the fine grid value based on the values of four closest coarse grid cell centres using the weighted average, applying weights based on the distance of the grid cell centres.

The statistical downscaling is performed according to the equation

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

where $\hat{Z}(s_0)$ is the estimated concentration at a point s_o of the fine grid,
$X_1(s_0),\ldots, X_n(s_0)$ are n proxy variables at a point s_o of the fine grid,
$c, a_1, a_2,,\ldots, a_n$ are the $n+1$ parameters of the linear regression model calculated at the coarse resolution (using source dataset and aggregated proxy datasets) ,
$\hat{\eta}(s_0)$ is the downscaling interpolation of the residuals of the linear regression model (as calculated at the coarse resolution) at a point s_o of the fine grid, either using the area-to point kriging or the bilinear interpolation.

One can see that Equation 2.1 consists of the regression and the interpolation part. Next to the two downscaling methods described by the whole Equation 2.1, i.e. LRM and area-to-point kriging on its residuals (LR.a2p) and LRM and bilinear interpolation on its residuals (LR.bl), three additional methods are also used that are decribed only by either regression or interpolation part of this equation. These methods are LRM without further interpolation (LR) and the interpolation without preciding LRM using either area-to-point kriging (a2p) or bilinear interpolation (bl).

In this report, the operational mapping result in 1 km x 1 km resolution (see Section 2.2.1) is used as a coarse source dataset, while the chemical transport modelling (CTM) output in 100 m x 100 m resolution (see Section 2.2.2) is used as the only fine resolution proxy dataset.

2.1.2 Comparison and validation of the mapping results

The evaluation of the maps and their mutual comparison is executed against the measurement data. The statistical indicators used are root mean square error (RMSE), relative root mean square error (RRMSE) and bias (mean prediction error, MPE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{Z}(s_i) - Z(s_i)\right)^2} \qquad (2.2)$$

$$RRMSE = \frac{RMSE}{\bar{Z}} . 100 \qquad (2.3)$$

$$bias(MPE) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{Z}(s_i) - Z(s_i)\right) \qquad (2.4)$$

where $Z(s_i)$ is the measurement value at the i^{th} point, $i = 1, \ldots, N$,

$\hat{Z}(s_i)$ is the estimated value at the i^{th} point, $i = 1, \ldots, N$,,

$\bar{Z}$ is the mean of the values $Z(s_1), \ldots, Z(s_N)$ as measured at points $i = 1, \ldots, N$,

N is the number of the measurement points.

RMSE and RRMSE should be as small as possible, bias (MPE) should be as close to zero as possible.

2.1.3 Population exposure

Population exposure (i.e. percentage population for several concentration classes) has been calculated for different mapping results in the 100 m x 100 m resolution (see Section 2.1.1). These results in several variants have been compared with the commonly calculated population exposure based on the separate mapping layers (i.e. using the background and the urban traffic map layers at 1 km x 1 km resolution and weighting them based on buffers around the roads), as routinely applied in the regular mapping (Horálek et al., 2022b). In addition, the population exposure based on the coarse resolution RIMM maps (without taking into account the separate mapping layers) have been calculated. In all cases, the population density data in 1 km x 1 km resolution (see Section 2.2.4) have been used.

Next to this, the population-weighted concentration has been also calulated for different mapping variants, according to the equation:

$$\hat{c} = \frac{\sum_{i=1}^{N} c_i p_i}{\sum_{i=1}^{N} p_i} \qquad (2.5)$$

where $\hat{c}$ is the population-weighted average concentration in the mapping area,

p_i is the population in the i^{th} grid cell,

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

N is the number of grid cells in the mapping area.

2.2 Data used

2.2.1 Original maps in coarse resolution

As the input maps in coarse resolution, the regular RIMM maps in 1x1 km resolution for 2020 as routinely produced under ETC HE (Horálek et al., 2022b) have been applied.

The maps of following pollutants and aggregations have been used:

NO_2 – annual average [µg/m^3], year 2020,
PM_{10} – annual average [µg/m^3], year 2020,
$PM_{2.5}$ – annual average [µg/m^3], year 2020.

2.2.2 Chemical transport modelling (CTM) output in fine resolution

As the proxy data in fine resolution, a combination of two air quality dispersion model outputs in 100 m x 100 m resolution as earlier prepared for the Czech national project ARAMIS have been used. In this combination, the fine-scale output of the Gaussian model SYMOS (CHMI, 2016) was combined with the 2325 m × 2325 m resolution output of the Eulerian chemical transport model (CTM) CAMx (ENVIRON, 2011). The SYMOS model provided annual average concentrations (contributions) of primary pollutants resulting from the road traffic (only roads covered by the traffic census were included), while in the CAMx model, all known sources were included and outputs for all hours were calculated and subsequently aggregated to the annual averages, which were further processed.

For the calculation by the SYMOS model, all roads (line sources) were divided into segments of maximum length of 20 m. For each point of the SYMOS's grid, contributions of all the road segments within 2-km distance were summed up. The combination of CAMx and SYMOS models was performed in five steps. At first, the SYMOS model was calculated on the 155 m × 155 m grid, in three levels of 2, 25 and 48 metres above ground, in order to cover homogenously the CAMx three-dimensional grid of 2325 m × 2325 m horizontally and cca 50 m vertically. Then, the aggregation of this SYMOS model output to the three-dimensional CAMx grid was performed. In the second step, this aggregated SYMOS model output was subtracted from the CAMx model output, in order to avoid a double counting of traffic emissions. In the third step, this adjusted CAMx output was resampled to the final 100 m × 100 m grid (each grid point was assigned value from the corresponding CAMx grid cell). In the fourth step, the SYMOS model was calculated on the same final 100 m × 100 m grid. And in the last fifth step, adjusted CAMx model and SYMOS model (both on the final 100 m × 100 m grid) were summed up.

The modelling outputs of following pollutants have been used:
NO_x – annual average [µg/m^3], year 2020,
PM_{10} – annual average [µg/m^3], year 2020.

The reason for calculation of NO_x (not NO_2) was the fact that in the Gaussian dispersion model SYMOS, only simple empirical parametrization of conversion of NO to NO_2 is taken into account.

The NO_x modelling output has been further used as a fine resolution proxy data in the downscaling for NO_2 map, while the PM_{10} modelling output as a proxy in the downscaling of both PM_{10} and $PM_{2.5}$ maps.

2.2.3 Air quality monitoring data

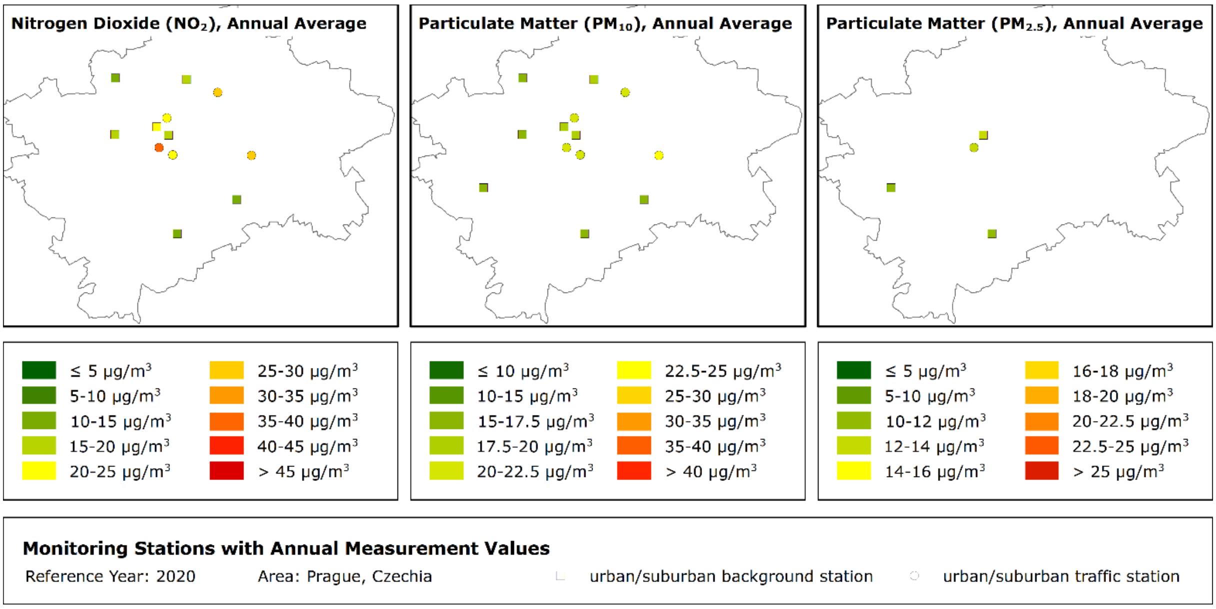
For the validation and mutual comparison of the downscaling results, we have used the air quality station monitoring data for 2020 coming from the Air Quality e-Reporting database (EEA, 2022). Data from stations located in Prague and classified as background or traffic (for the two relevant types of area, i.e. urban and suburban) have been considered. Only stations with annual data coverage of at least 75 percent have been used. Table 2.1 shows the number of the measurement stations selected for the individual pollutants.

Table 2.1: Number of stations used in validation for each pollutant and area type, 2020, Prague

Station type	NO_2	PM_{10}	$PM_{2.5}$
	Annual average	Annual average	Annual average
Urban and suburban background	7	8	3
Urban and suburban traffic	5	5	1

Figure 2.1 presents the monitoring stations used in validation for different pollutants, including their annual average values for 2020.

Figure 2.1: Spatial distribution of NO_2 (left), PM_{10} (middle) and $PM_{2.5}$ (right) monitoring stations used in validation, including the measured annual average concentration, 2020, Prague



2.2.4 Population density data

Population density (in inhbs/km^2, census 2011) is based on the Geostat 2011 grid dataset, Eurostat (2014). Just as in Horálek et al. (2022b), these data have been scaled to 2020 data using the national population totals for 2020 (Eurostat, 2022).

3 Analysis and map creation

In this chapter, the results of the different downscaling techniques are compared mutually and with the original regular maps in coarse 1 km x 1 km resolution. For each pollutant, the following mapping results are compared:

- original regular RIMM map (O), coarse resolution 1 km x 1 km,
- area-to-point kriging on original map (a2p), resolution 100 m x 100 m,
- bilinear interpolation on original map (bl), resolution 100 m x 100 m,
- linear regression (LR), resolution 100 m x 100 m,
- linear regression and area-to-point kriging on its residuals (LR.a2p), resolution 100 m x 100 m,
- linear regression and bilinear interpolation on its residuals (LR.bl), resolution 100 m x 100 m.

The reference downscaling methodology used is the linear regression and the interpolation of its residuals by the area-to-point kriging (LR.a2p). As an alternative, simple bilinear interpolation is used for the interpolation of the linear regression´s residuals instead of the geostatistical area-to-point kriging. The advantage of the bilinear interpolation is that it is not so computationally demanding compared to the area-to-point kriging. Next to these two downscaling methods (LR.a2p) and (LR.bl) that comprise both linear regression and interpolation, downscaling methods using only interpolation and linear regression only are also examined in addition, i.e. methods (a2p), (bl) and (LR).

Apart from a statistical analysis, the final fine resolution maps are presented for each pollutant. For consistency, the same downscaling method have been used for production of these maps, namely the linear regression and area-to-point kriging on its residuals (LR.a2p), which is the reference downscaling method in this paper.

Next to the downscaled mapping results, population exposure based on these results have been also estimated. The population exposure calculated based on the different downscaled mapping results in the 100 m x 100 m resolution have been compared with the commonly calculated population exposure based on the separate urban background and traffic map layers (using buffers around the roads), as routinely applied in the regular mapping (Horálek et al., 2022b), labelled (SO_bf). In addition, the population exposure based on the coarse resolution RIMM maps (without taking into account the separate mapping layers) have been calculated, labelled (O_1k).

Section 3.1 presents the results for NO_2, Section 3.2 for PM_{10} and Section 3.3 for $PM_{2.5}$.

3.1 NO_2 annual average

As a first step, the linear regression parameters have been estimated at the coarse resolution (when the dependent variable is the original regular RIMM map and the independent variable is the fine-resolution CTM output aggregated to the same coarse resolution). The estimated intercept is c = 4.41, the estimated slope is a = 0.683, the adjusted R^2 is 0.639 and the standard error is 2.20 $\mu g/m^3$.

Table 3.1 presents the comparison of different mapping results in the fine 100 m x 100 m resolution and the original RIMM map in the coarse 1 km x 1 km resolution for NO_2 annual average 2020 in Prague. The table rows highlighted by green and light green show the statistics that provide the best and the second best performances.

In addition, the table presents also the scores for the modelling proxy data in the fine resolution (highligted by light blue), although the modelling results show NO_x values, not NO_2. These scores (i.e. statistics of NO_x modelling results against the NO_2 measurement data) are presented for illustration only and are not further used in the comparison.

Table 3.1: Comparison of different mapping results showing RMSE, RRMSE and bias against measurement data of urban and suburban background (left) and traffic (right) stations for NO_2 annual average 2020 in Prague. Units: µg/m^3 except RRMSE

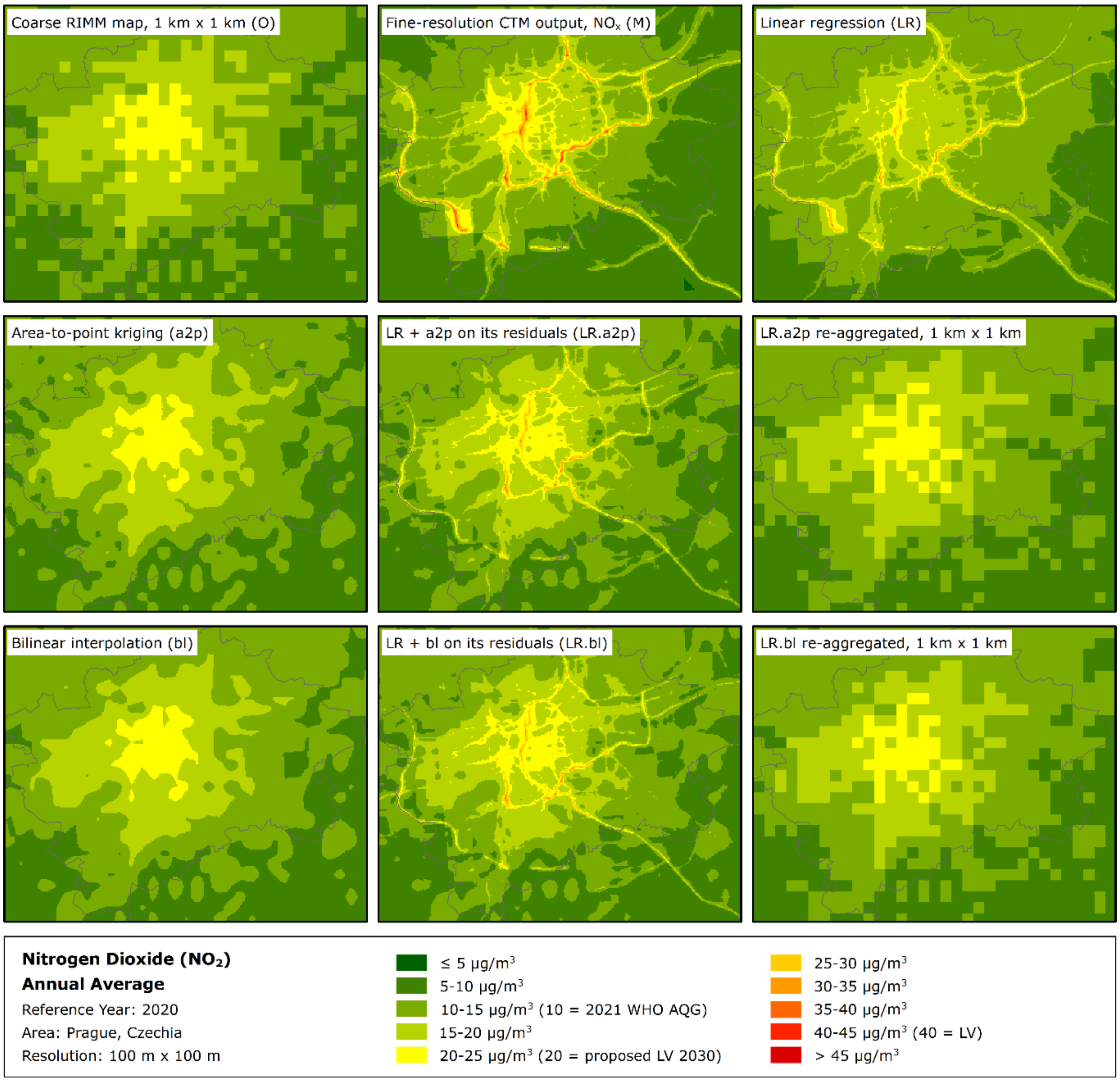
Mapping Variant		NO_2 Annual Average					
		Urban Background Areas			Urban Traffic Areas		
		RMSE	RRMSE	Bias	RMSE	RRMSE	Bias
M	Fine-resolution CTM model, NO_x, 100 m x 100 m	1.6	9.3	-1.0	7.4	26.3	-5.3
O	Original regular RIMM map, 1 km x 1 km	2.6	14.8	0.7	10.8	38.3	-9.1
a2p	Area-to point kriging on original map, 100 m x 100 m	2.5	14.3	1.1	11.1	39.3	-9.4
bi	Bilinear interpolation on original map, 100 m x 100 m	2.3	12.8	0.9	10.9	38.6	-9.2
LR	Linear regression, 100 m x 100 m	2.3	13.1	-1.9	9.4	33.2	-8.2
LR.a2p	LR + a2p kriging on its residuals, 100 m x 100 m	1.8	10.0	0.0	8.5	29.9	-7.0
LR.bi	LR + bl interpolation on its residuals, 100 m x 100 m	1.9	10.8	-0.1	8.6	30.3	-7.0

Looking at Table 3.1, one can see that the best results are given by the downscaling methods (LR.a2p) and (LR.bl), with slightly better performance provided by (LR.a2p). Both of these methods improve the results given by the original map in the coarse 1 km x 1 km resolution in terms of all statistics, i.e. RMSE, RRMSE and bias. While the original coarse resolution results are moderately overestimated in the urban background areas, the downscaled results show no or almost no bias in these areas. The traffic areas are still somewhat underestimated in the fine 100 m x 100 m resolution, however the mean level of this bias is reduced from -9 µg/m^3 to -7 µg/m^3 for the area of Prague.

Figure 3.1 shows the coarse-resolution input data (i.e., original RIMM map in 1 km x 1 km resolution), the fine-resolution modelled proxy data and the different mapping results in the fine 100 m x 100 m resolution. In addition, re-aggregation of two downsclaing results (LR.a2p) and (LR.bl) to the coarse resolution 1 km x 1 km is also shown.

Looking at the maps, one can see the main roads with elevated NO_x and NO_2 concentrations both in the fine-resolution CTM proxy data and in the downscaled mapping results of methods using linear regression, i.e. (LR), (LR.a2p) and (LR.bl). It can be seen that the downscalling methods (LR.a2p) and (LR.bl) give almost the same results.

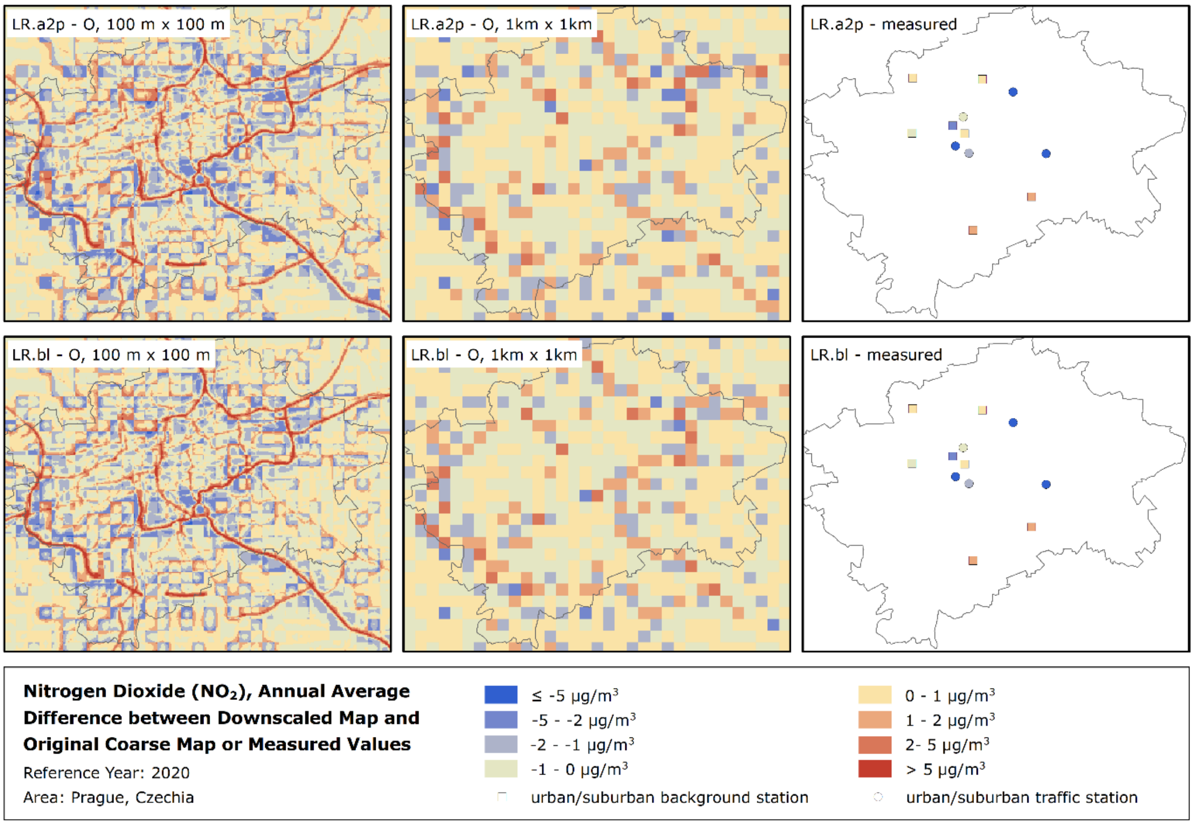
Figure 3.1: Coarse-resolution input data (top left), fine-resolution modelled proxy data (top middle) and the fine resolution mapping results of the downscaling methods LR (top right), a2p (centre left), LR.a2p (centre middle), bl (bottom left) and Lr.bl (bottom middle) and the re-aggregated mapping results of the downscaling methods LR.a2p (centre right) and LR.bl (bottom right) to the coarse resolution for NO_2 annual average 2020, Prague



In order further to examine the downscaled mapping results, the difference maps of downscaled vs. coarse-resolution mapping data and re-aggregated downscaled vs. coarse-resolution mapping data have been prepared, as well as the maps showing the differences of the downscaled mapping results vs. measurement data in the points of the measurement stations (separatelly for the urban/suburban background and the urban/suburban traffic stations). These maps have been prepared for the downscaling methods (LR.a2p) and (LR.bl), see Figure 3.2.

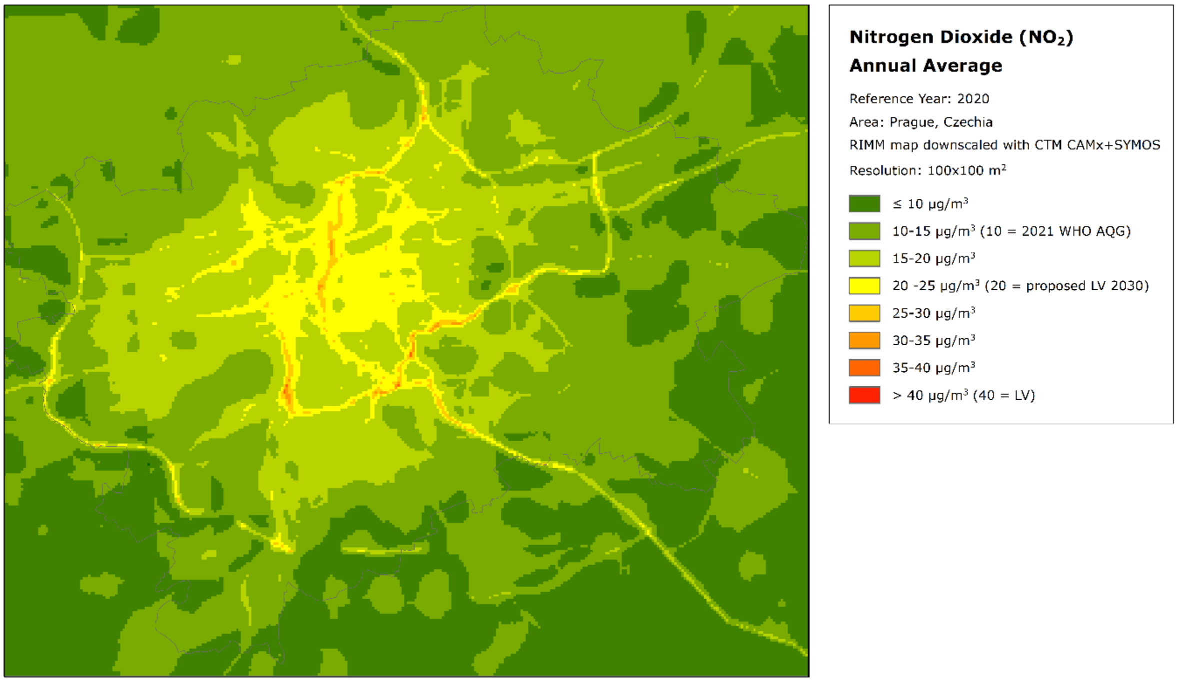
Looking at the difference maps of fine-resolution downscaled maps vs. coarse-resolution input data, one can see the effect of the downscalling. In these difference maps, the main roads with higher NO_2 concentrations estimated by the downscaled maps can be seen. The difference maps of re-aggregated downscaled maps vs. coarse-resolution map input data show into which level the downscaling method is mass-conservative, i.e. how it changes the values of the coarse-resolution grid cells.

Figure 3.2: Difference between downscaled mapping results and coarse-resolution input data (left), re-aggregated downscaled mapping results and coarse-resolution input data (middle) and downscaled mapping results and measurement data (right) based on the downscaling methods LR.a2p (centre right) and LR.bl (bottom right) for NO_2 annual average 2020, Prague



Map 3.1 presents the final fine resolution map of NO_2 annual average 2020 for Prague, as created by the reference downscaling method (LR.a2p), i.e. the linear regression and area-to-point kriging on its residuals. In the map, one can see the main roads in Prague with elevated NO_2 concentrations.

Map 3.1: Fine resolution concentration map of NO_2 annual average 2020, Prague



Next to the downscaled mapping results, population exposure based on these results have been estimated and compared with the commonly calculated population exposure based on the separate urban background and traffic map layers, as routinely applied in the regular mapping (Horálek et al., 2022b). In addition, the population exposure based on the coarse resolution RIMM maps (without taking into account the separate mapping layers) have been also calculated.

Table 3.2 presents the population frequency distribution for a limited number of exposure classes to NO_2 concentrations and the population-weighted concentration as calculated based on the different mapping results.

Table 3.2: Population exposure and population-weighted concentration for NO_2 annual average 2020 in Prague calculated based on different mapping results

Mapping Variant		Population [inhbs]	NO_2 Annual Average						
			Exposed Population						Pop.-w. Conc. [µg/m³]
			< 10 µg/m³	10-20 µg/m³	20-30 µg/m³	30-40 µg/m³	40-45 µg/m³	> 45 µg/m³	
SO_bf	Separate RIMM map layers + buffers	1 421 394	20 716	1 144 980	255 698	0	0	0	17.5
O_1k	Coarse RIMM map, 1 km x 1 km	1 421 394	19 672	1 075 290	326 432	0	0	0	17.5
a2p	Area-to point kriging, 100 m x 100 m	1 421 394	22 550	1 083 990	314 854	0	0	0	17.2
bl	Bilinear interpolation, 100 m x 100 m	1 421 394	22 403	1 090 692	308 299	0	0	0	17.1
LR	Linear regression, 100 m x 100 m	1 421 394	57 865	1 254 947	107 251	1 330	0	0	15.3
LR.a2p	LR + a2p on its residuals, 100 m x 100 m	1 421 394	30 553	1 078 150	310 959	1 732	0	0	17.1
LR.bl	LR + bl on its residuals, 100 m x 100 m	1 421 394	28 586	1 081 932	309 144	1 732	0	0	17.1

Looking at the results, one can see that the method (LR) gives noticebly smaller estimate of the population-weighted concentration, compared to the other methods.

3.2 PM_{10} annual average

At first, the linear regression parameters have been estimated at the coarse resolution (when the dependent variable is the original regular RIMM map and the independent variable is the fine-resolution CTM output aggregated to the same coarse resolution). The estimated intercept is c = 12.03, the estimated slope is a = 0.328, the adjusted R^2 is 0.369 and the standard error is 1.07 µg/m^3. One can see only quite a weak correlation between the coarse-resolution data and the proxy data.

Table 3.3 presents the comparison of different mapping results in the fine 100 m x 100 m resolution and the original RIMM map in the coarse 100 m x 100 m resolution for PM_{10} annual average 2020 in Prague. Again, the rows highlighted by green and light green show the statistics that provide the best and the second best downscaling performances.

Table 3.3: Comparison of different mapping results showing RMSE, RRMSE and bias against measurement data of urban and suburban background (left) and traffic (right) stations for PM_{10} annual average 2020 in Prague. Units: µg/m^3 except RRMSE

Mapping Variant		PM_{10} Annual Average					
		Urban Background Areas			Urban Traffic Areas		
		RMSE	RRMSE	Bias	RMSE	RRMSE	Bias
M	Fine-resolution CTM model, 100 m x 100 m	1.5	9.2	-1.1	2.9	13.7	-2.3
O	Original regular RIMM map, 1 km x 1 km	1.5	8.8	1.1	3.2	15.1	-2.9
a2p	Area-to point kriging on original map, 100 m x 100 m	1.5	9.3	1.1	3.3	15.7	-3.1
bl	Bilinear interpolation on original map, 100 m x 100 m	1.5	9.2	1.1	3.3	15.3	-3.0
LR	Linear regression, 100 m x 100 m	1.0	6.1	0.5	3.2	15.0	-3.0
LR.a2p	LR + a2p kriging on its residuals, 100 m x100 m	1.4	8.1	0.9	2.8	13.3	-2.6
LR.bl	LR + bl interpolation on its residuals, 100 m x 100 m	1.3	8.0	0.9	2.8	13.3	-2.6

The results presented in Table 3.3 for PM_{10} are not as straigtforward as those shown for NO_2 in Table 3.1. The downscaling methods (LR.a2p) and (LR.bl) are the best ones for the traffic areas, however only the second best ones for the urban background areas, for which the best results are given by the method (LR), i.e. by simple linear regression without further interpolation of its residuals. Compared to the original map in coarse resolution, these downscaled fine resolution maps give only slightly better results. The results are driven by three stations (and mainly by two of them) in the suburban areas with coarse results overestimated. Thus, the pure linear regression gives better results for them than the variants in which the residuals bring the values closer to the coarse map.

Figure 3.3 shows the coarse-resolution input data (i.e., original RIMM map in 1 km x 1 km resolution), the fine-resolution modelled proxy data and the different mapping results in the fine 100 m x 100 m resolution. In addition, re-aggregation of two downscaling results (LR.a2p) and (LR.bl) to the coarse resolution 1 km x 1 km is also shown.

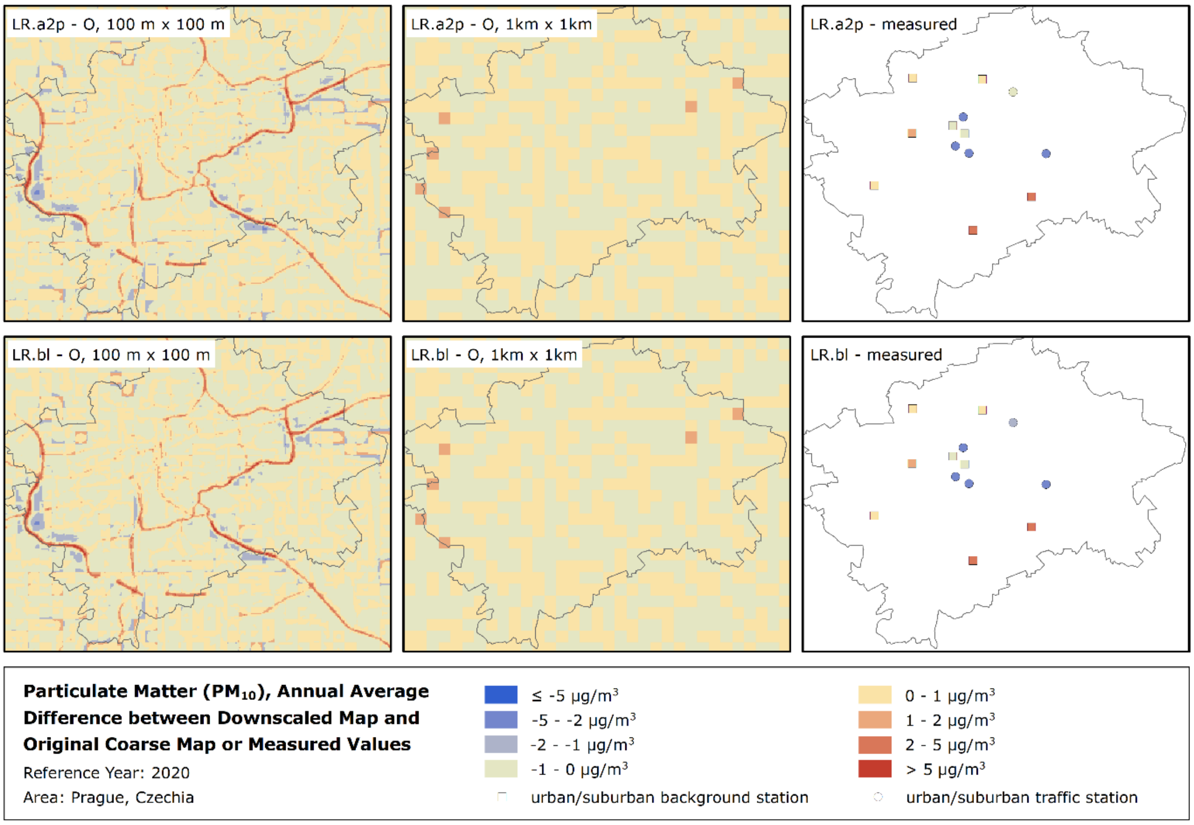
Compared to the downscaled NO_2 maps (see Figure 3.1), the main roads are less visible in the downscaled mapping results. As in the case of NO_2, the downscalling methods (LR.a2p) and (LR.bl) give almost the same results.

Figure 3.3: Coarse-resolution input data (top left), fine-resolution modelled proxy data (top middle) and the fine resolution mapping results of the downscaling methods LR (top right), a2p (centre left), LR.a2p (centre middle), bl (bottom left) and Lr.bl (bottom middle) and the re-aggregated mapping results of the downscaling methods LR.a2p (centre right) and LR.bl (bottom right) to the coarse resolution for PM_{10} annual average 2020, Prague

Just as for NO_2, the difference maps of downscaled vs. coarse-resolution mapping data and re-aggregated downscaled vs. coarse-resolution mapping data have been prepared, as well as the maps showing the differences of the downscaled mapping results vs. measurement data in the points of the measurement stations, see Figure 3.4. Again, these maps have been prepared for the downscaling methods (LR.a2p) and (LR.bl).

One can see that the effect of the downscaling is smaller for PM_{10} compared to the NO_2 (see Figure 3.2). In the difference maps of fine-resolution downscaled maps vs. coarse-resolution input data, one can see several main roads with higher PM_{10} concentrations estimated by the downscaled maps. Looking at the difference maps of re-aggregated downscaled maps vs. coarse-resolution map input data, one can state that the re-aggregated downscaled maps give almost the same results as the original coarse map.

Figure 3.4: Difference between downscaled mapping results and coarse-resolution input data (left), re-aggregated downscaled mapping results and coarse-resolution input data (middle) and downscaled mapping results and measurement data (right) based on the downscaling methods LR.a2p (centre right) and LR.bl (bottom right) for PM_{10} annual average 2020, Prague



Map 3.2 shows the final fine resolution map of PM_{10} annual average 2020 for Prague, as created by the reference downscaling method (LR.a2p), i.e. the linear regression and area-to-point kriging on its residuals.

Looking at Map 3.2, one can note that the roads are not as visible as in the case of NO_2 (see Map 3.1). The reason probably is that the differences between the traffic and the background air pollution are not so profound for PM_{10} as for NO_2.

Map 3.2: Fine resolution concentration map of PM_{10} annual average 2020, Prague

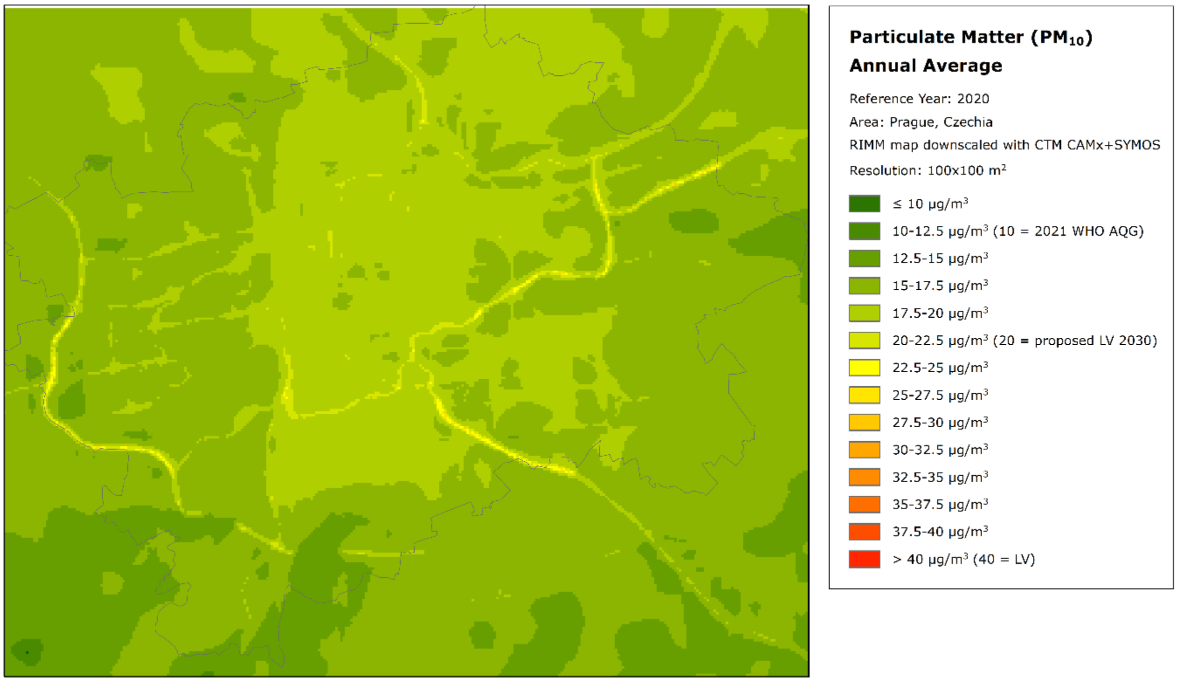


Table 3.4 presents the population frequency distribution for a limited number of exposure classes to PM_{10} concentrations and the population-weighted concentration as calculated based on the different mapping results.

Table 3.4: Population exposure and population-weighted concentration for PM_{10} annual average 2020 in Prague calculated based on different mapping results

Mapping Variant		Population [inhbs]	PM_{10} Annual Average						
			Exposed Population						Pop.-w. Conc. [µg/m³]
			< 10 µg/m³	10-20 µg/m³	20-30 µg/m³	30-40 µg/m³	40-50 µg/m³	> 50 µg/m³	
SO_bf	Separate RIMM map layers + buffers	1 421 394	8 465	1 318 216	94 713	0	0	0	17.8
O_1k	Coarse RIMM map, 1 km x 1 km	1 421 394	7 128	1 414 266	0	0	0	0	17.8
a2p	Area-to point kriging, 100 m x 100 m	1 421 394	12 630	1 408 763	0	0	0	0	17.7
bl	Bilinear interpolation, 100 m x 100 m	1 421 394	12 510	1 408 883	0	0	0	0	17.7
LR	Linear regression, 100 m x 100 m	1 421 394	0	1 405 207	16 187	0	0	0	17.2
LR.a2p	LR + a2p on its residuals, 100 m x 100 m	1 421 394	17 929	1 386 612	16 853	0	0	0	17.7
LR.bl	LR + bl on its residuals, 100 m x 100 m	1 421 394	16 415	1 387 960	17 019	0	0	0	17.7

3.3 $PM_{2.5}$ annual average

At first, the linear regression parameters have been estimated at the coarse resolution (when the dependent variable is the original regular RIMM map and the independent variable is the fine-resolution CTM output aggregated to the same coarse resolution). The estimated intercept is c = 8.88, the estimated slope is a = 0.213, the adjusted R^2 is 0.358 and the standard error is 0.71 µg/m³. Similarly as in the case of PM_{10}, there is only quite a weak correlation between the coarse-resolution data and the proxy data.

Table 3.5 gives the comparison of different mapping results in the fine 100 m x 100 m resolution and the original RIMM map in the coarse 100 m x 100 m resolution for $PM_{2.5}$ annual average 2020 in Prague.

Again, the rows highlighted by green and light green show the statistics that provide the best and the second best performances. It should be noted that $PM_{2.5}$ observations from only three background stations and one traffic station have been available for evaluation (EEA, 2022), which influences the robustness of the results.

Table 3.5: Comparison of different mapping results showing RMSE, RRMSE and bias against measurement data of urban and suburban background (left) and traffic (right) stations for $PM_{2.5}$ annual average 2020 in Prague. Units: µg/m³ except RRMSE

Mapping Variant		$PM_{2.5}$ Annual Average					
		Urban Background Areas			Urban Traffic Areas		
		RMSE	RRMSE	Bias	RMSE	RRMSE	Bias
M	Fine-resolution CTM model, PM_{10}, 100 m x 100 m	3.24	27.4	3.18	8.52	63.7	8.52
O	Original regular RIMM map, 1 km x 1 km	0.82	6.9	0.65	0.32	2.4	-0.32
a2p	Area-to point kriging on original map, 100 m x 100 m	0.93	7.9	0.81	0.40	3.0	-0.40
bi	Bilinear interpolation on original map, 100 m x 100 m	0.88	7.5	0.77	0.40	3.0	-0.40
LR	Linear regression, 100 m x 100 m	0.50	4.2	0.27	0.18	1.3	0.18
LR.a2p	LR + a2p kriging on its residuals, 100 m x100 m	0.87	7.4	0.63	0.23	1.7	0.23
LR.bi	LR + bl interpolation on its residuals, 100 m x 100 m	0.84	7.1	0.61	0.25	1.8	0.25

Looking at Table 3.5, one can see that for the urban background areas, the best results are given by the simple linear regression (LR), while for the traffic areas, the best results are shown by methods (LR), (LR.a2p) and (LR.bl). The results of these downscaled maps are only slightly better compared to the results of the original coarse resolution map. This is probably highly influenced by the weak correlation between coarse-resolution data and the proxy data. In any case, due to the very limited number of the monitoring stations, one cannot make strong conclusions.

Figure 3.5 shows the coarse-resolution input data (i.e., original RIMM map in 1 km x 1 km resolution), the fine-resolution modelled proxy data and the different mapping results in the fine 100 m x 100 m resolution. In addition, re-aggregation of two downsclaing results (LR.a2p) and (LR.bl) to the coarse resolution 1 km x 1 km is also shown.

Again, the downscalling methods (LR.a2p) and (LR.bl) give almost the same results. Like for PM_{10}, the main roads are less visible in the downscaled mapping results compared to the downscaled NO_2 maps (see Figure 3.1).

Figure 3.5: Coarse-resolution input data (top left), fine-resolution modelled proxy data (top middle) and the fine resolution mapping results of the downscaling methods LR (top right), a2p (centre left), LR.a2p (centre middle), bl (bottom left) and Lr.bl (bottom middle) and the re-aggregated mapping results of the downscaling methods LR.a2p (centre right) and LR.bl (bottom right) to the coarse resolution for $PM_{2.5}$ annual average 2020, Prague

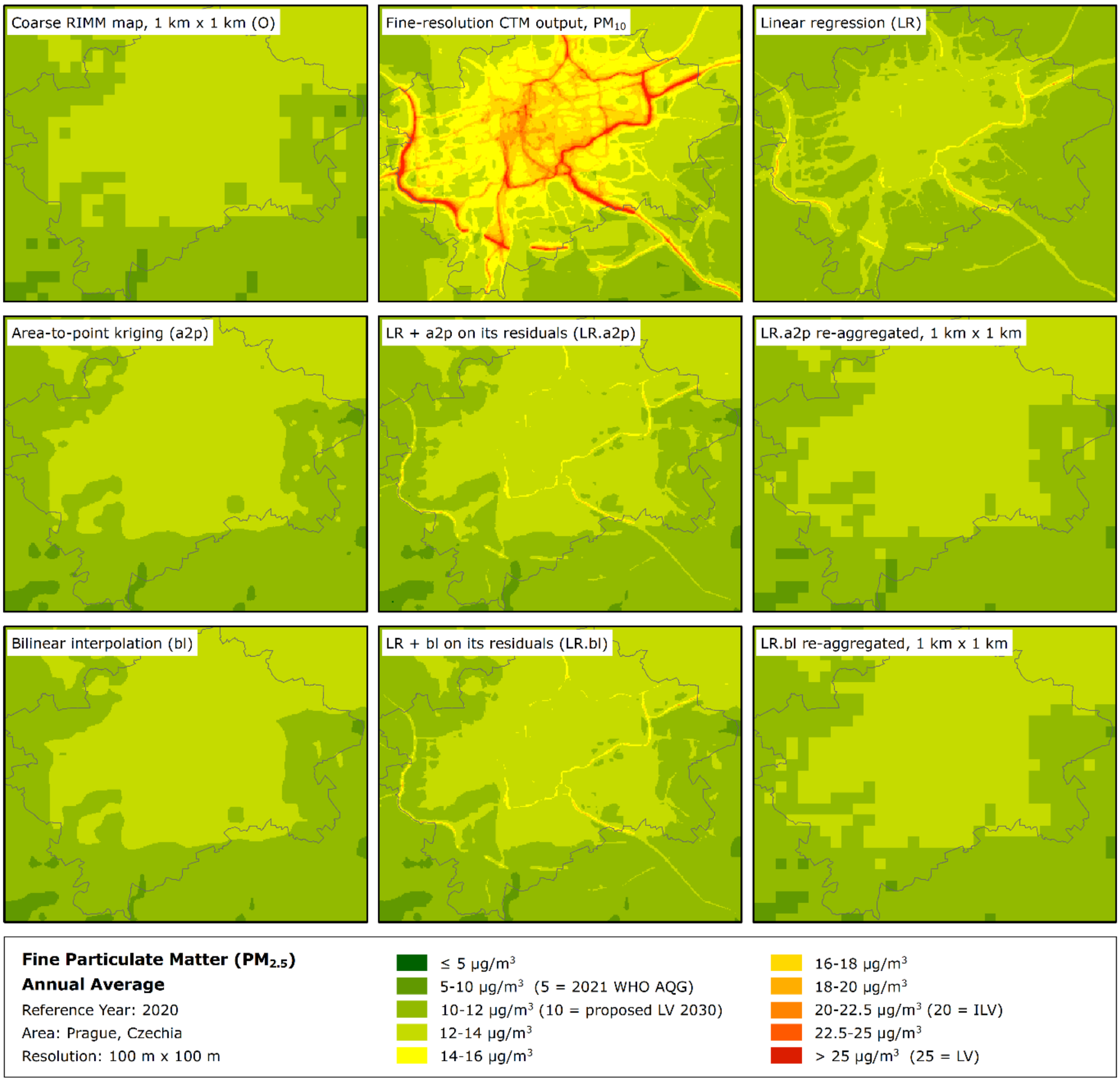
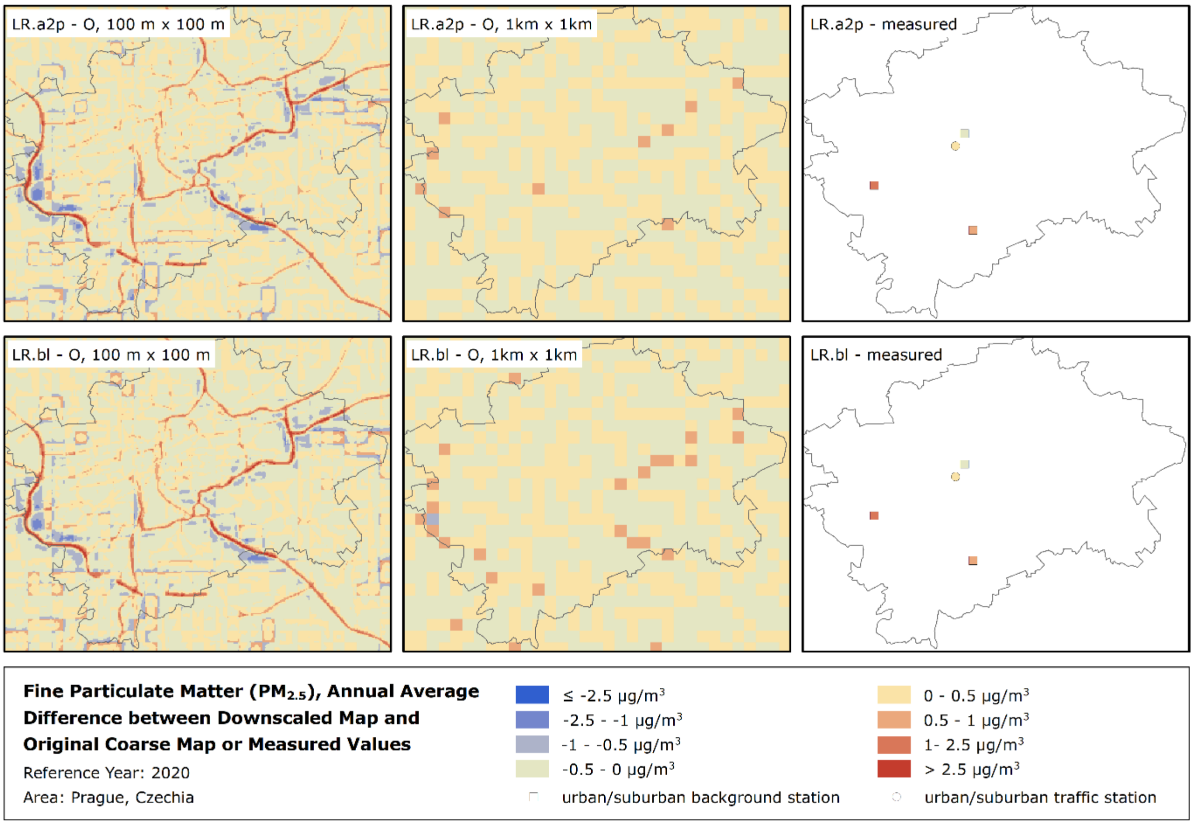


Figure 3.6 gives the difference maps of downscaled vs. coarse-resolution mapping data and re-aggregated downscaled vs. coarse-resolution mapping data, as well as the maps showing the differences of the downscaled mapping results vs. measurement data in the points of the measurement stations. Again, these maps have been prepared for the downscaling methods (LR.a2p) and (LR.bl).

Similarly as for PM_{10}, the effect of the downscaling is smaller for $PM_{2.5}$ compared to the NO_2 (see Figure 3.2). Again, one can see several main roads with higher $PM_{2.5}$ concentrations estimated by the downscaled maps.

Figure 3.6: Difference between downscaled mapping results and coarse-resolution input data (left), re-aggregated downscaled mapping results and coarse-resolution input data (middle) and downscaled mapping results and measurement data (right) based on the downscaling methods LR.a2p (centre right) and LR.bl (bottom right) for $PM_{2.5}$ annual average 2020, Prague



Map 3.3 presents the fine resolution map of $PM_{2.5}$ annual average 2020 for Prague, as created by the reference downscaling method (LR.a2p), i.e. the linear regression and area-to-point kriging on its residuals.

Looking at Map 3.3, one can see only slightly higher $PM_{2.5}$ concentrations in the main roads. The reason probably is that the traffic increment is not so large in the case of $PM_{2.5}$.

Map 3.3: Fine resolution concentration map of $PM_{2.5}$ annual average 2020, Prague

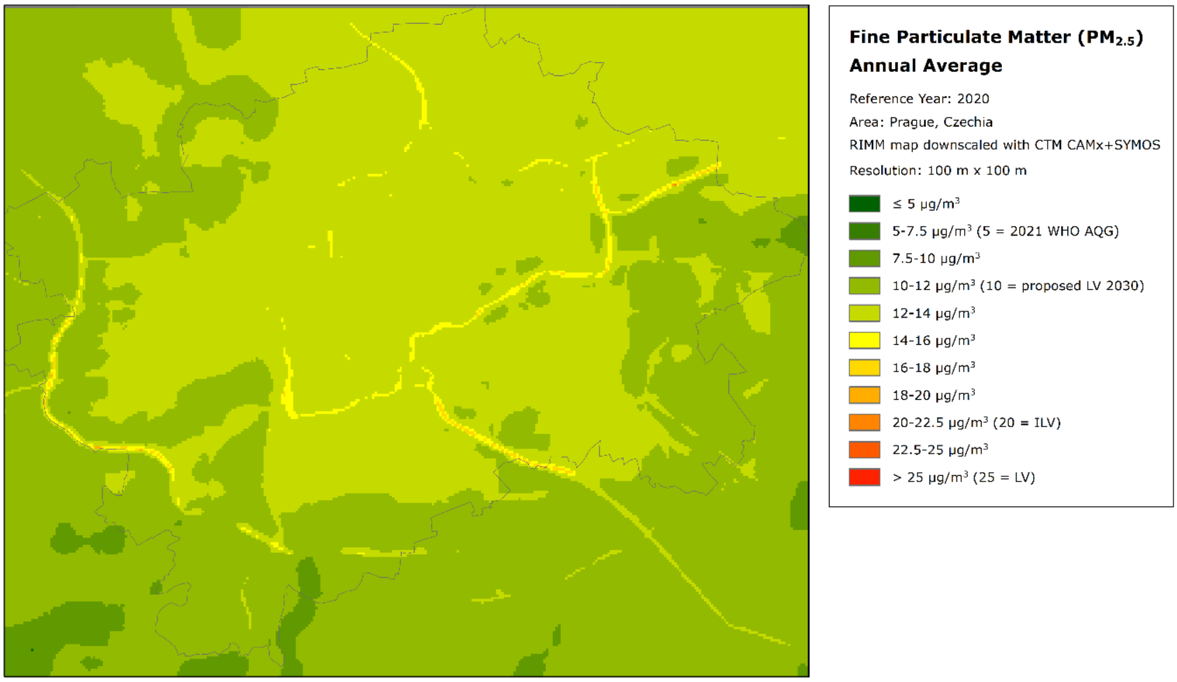


Table 3.6 presents the population frequency distribution for a limited number of exposure classes to $PM_{2.5}$ concentrations and the population-weighted concentration as calculated based on the different mapping results.

Table 3.6: Population exposure and population-weighted concentration for $PM_{2.5}$ annual average 2020 in Prague calculated based on different mapping results

Mapping Variant		Population [inhbs]	$PM_{2.5}$ Annual Average						
			Exposed Population						Pop.-w. Conc. [µg/m³]
			< 5 µg/m³	5-10 µg/m³	10-15 µg/m³	15-20 µg/m³	20-25 µg/m³	> 25 µg/m³	
O_bf		1 421 394	0	992	1 420 402	0	0	0	12.6
O_1k	Coarse RIMM map, 1 km x 1 km	1 421 394	0	992	1 420 402	0	0	0	12.6
a2p	Area-to point kriging, 100 m x 100 m	1 421 394	0	1 234	1 420 160	0	0	0	12.5
bl	Bilinear interpolation, 100 m x 100 m	1 421 394	0	815	1 420 578	0	0	0	12.5
LR	Linear regression, 100 m x 100 m	1 421 394	0	0	1 415 124	6 262	8	0	12.3
LR.a2p	LR + a2p on its residuals, 100 m x 100 m	1 421 394	0	1 318	1 415 551	4 525	0	0	12.5
LR.bl	LR + bl on its residuals, 100 m x 100 m	1 421 394	42	993	1 415 597	4 763	0	0	12.5

4 Conclusion

The report examines the city-level air quality mapping at the fine resolution 100 m x 100 m for the area of Prague, as a case study. Several methods of statistical downscaling have been implemented and evaluated. The reference downscaling methodology used is the linear regression and the interpolation of its residuals by the area-to-point kriging. As an alternative, bilinear interpolation is used for the interpolation of the linear regression´s residuals. In addition and for comparison purposes, methods using the interpolation only and the linear regression only have been also examined. The results of different downscaling methods have been compared based on the data from the monitoring stations of Prague, separately for urban background and traffic areas.

Based on the analysis, it can be concluded that the best results for NO_2 are given by the downscaling methods of the linear regression and the interpolation of its residuals, either by the area-to-point kriging or the bilinear interpolation. Both these methods improve the results given by the original map in the coarse 1 km x 1 km resolution. The downscaled results show almost no bias in the urban background areas. The traffic areas are still somewhat underestimated in the fine resolution, however the level of this bias is reduced, compared to the coarse resolution map.

For PM_{10}, the downscaling methods of the linear regression and the interpolation of its residuals give the best ones for the traffic areas, however only the second best ones for the urban background areas, for which the best results are given by simple linear regression. Compared to the original coarse resolution map, the downscaled maps give only slightly better results.

For $PM_{2.5}$, the best downscaling results are given by simple linear regression, followed by the methods of the linear regression and the interpolation of its residuals. However, due to very limited number of the monitoring stations available fo $PM_{2.5}$, one cannot make strong conclusions.

These findings agree with those obtained from other downscaling methods (e.g. the uEMEP downscaling approach, Denby et al., 2020) that also tend to perform best for NO_2 and show significantly poorer results for particulate matter. The reasons can vary somewhat from method to method, however in general downscaling methods will perform better for pollutants such as NO_2, for which the pollution typically remains spatially relatively close to their emissions sources.

The downscaled maps in 100 m x 100 m resolution have been constructed for three pollutants, namely for NO_2, PM_{10} and $PM_{2.5}$. In the maps, one can see overall realistic spatial patterns, for example the main roads in Prague are visible through higher air pollution levels. This is distinct especially for NO_2, while for PM_{10} and $PM_{2.5}$ the differences between road increments and urban background are smaller as would be expected. Based on the downscaled mapping results, the population exposure estimates have been also calculated, in addition.

The results of the case study for Prague have proven the usefulness of the statistical downscaling for the air quality mapping, especially for NO_2. It is recommended further to examine the downscaling for more cities of Europe (ideally for all cities of the Eurostat´s Urban Audit), including the population exposure estimates. Such examination is dependent on available proxy data in fine resolution. Next to the statistical downscaling, applying the existing mapping methodology but exploiting a high-resolution CTM output (e.g. from the uEMEP model, Denby et al., 2020) should be also examined if possible, as recommended in Horálek et al. (2022a).

List of abbreviations

Abbreviation	Name	Reference
AQ	Air Quality	
AQG	Air Quality Guideline	
CAMx	Comprehensive Air Quality Model with Extensions	https://www.camx.com/
CHMI	Czech Hydrometeorological Institute	https://www.chmi.cz/
CTM	Chemical Transport model	
EEA	European Environment Agency	www.eea.europa.eu
ETC HE	European Topic Centre on Human health and the Environment	https://www.eionet.europa.eu/etcs
EU	European Union	https://european-union.europa.eu
GRIP	Global Roads Inventory Dataset	
ILV	Indicative Limit Value	http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:152:0001:0044:EN:PDF
LV	Limit Value	http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:152:0001:0044:EN:PDF
NILU	Norwegian Institute for Air Research	https://www.nilu.no/
NO_2	Nitrogen dioxide	
PM_{10}	Particulate Matter with a diameter of 10 micrometres or less	
$PM_{2.5}$	Particulate Matter with a diameter of 2.5 micrometres or less	
R^2	Coefficient of determination	
RIMM	Regression – Interpolation – Merging Mapping	
RMSE	Root Mean Square Error	
SAMIRA	Satellite based Monitoring Initiative for Regional Air Quality	https://samira.nilu.no
WHO	World Health Organization	https://www.who.int/

References

CHMI, 2016, *Model SYMOS* (https://www.chmi.cz/aktualni-situace/stav-ovzdusi/modelovani-kvality-ovzdusi/model-symos) accessed 30/11/2022.

Denby, B., et al., 2020, Description of the uEMEP_v5 downscaling approach for the EMEP MSC-W chemistry transport model, *Geoscientific Model Development, 13(12),* 6303-6323 (https://doi.org/10.5194/gmd-13-6303-2020) accessed 22/12/2022.

EEA, 2022, *Air Quality e-Reporting.* Air quality database *(*https://www.eea.europa.eu/data-and-maps/data/aqereporting-9). Data extracted in March 2022.

ENVIRON, 2011, *User´s Guide to the Comprehensive Air Quality Model with Extensions (CAMx), Version 5.40* (http://www.camx.com) accessed 12/03/2020.

ESRI, 2021*, Cell size and resampling in analysis* (https://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/cell-size-and-resampling-in-analysis.htm) accessed 20/12/2022.

Eurostat, 2014, *GEOSTAT 2011 grid dataset*, Population distribution dataset (http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography) accessed 26/08/2020.

Eurostat, 2022, *Total population for European states for 2020* (https://ec.europa.eu/eurostat/web/main/data/database) accessed 29 August 2022.

Horálek, J., et al., 2022a, *European-wide city level mapping*, Eionet Report ETC/ATNI 2021/20 (https://doi.org/10.5281/zenodo.6802606) accessed 28/11/2022.

Horálek, J., et. al., 2022b, *European air quality maps for 2020*, Eionet Report ETC HE 2022/12 (https://www.eionet.europa.eu/etcs/etc-he/products/etc-he-report-2022-12-european-air-quality-maps-for-2020-pm10-pm2-5-ozone-no2-nox-and-benzo-a-pyrene-spatial-estimates-and-their-uncertainties) accessed 09/02/2023.

Kyriakidis, P. C., 2004, A geostatistical framework for area-to-point spatial interpolation, *Geographical Analysis 36,* pp. 259–289 (https://doi.org/10.1111/j.1538-4632.2004.tb01135.x) accessed 23/11/2022.

Park, N.-W., 2013, Spatial downscaling of TRMM precipitation using geostatistics and fine scale environmental variables*, Advances in Meteorology, Volume 2013*, Article ID 237126.

Stebel, K., et al., 2021, SAMIRA-SAtellite Based Monitoring Initiative for Regional Air Quality, *Remote Sensing, 13(11),* 2219 (https://doi.org/10.3390/rs13112219) accessed 11/12/2021.

European Topic Centre on
Human Health and the Environment
https://www.eionet.europa.eu/etcs/etc-he

The European Topic Centre on Human Health and the Environment (ETC HE) is a consortium of European institutes under contract of the European Environment Agency.

European Environment Agency
European Topic Centre
Human health and the environment