Milestone M16 (M3.5)

Definition of metrics for sub-grid variability

RI-URBANS

Research Infrastructures Services Reinforcing Air Quality Monitoring Capacities in European Urban & Industrial AreaS (GA n. 101036245)

By

CNRS/ENPC, NOA, METNO & FORTH

30th May 2023

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

Milestone M16 (M3.5): Definition of metrics for sub-grid variability

Authors: Karine Sartelet (CNRS/ENPC), Eleni Athanasopoulou (NOA), Bruce Rostald Denby (METNO), Evangelos Gerasopoulos (NOA), Hilde Fagerli (METNO) & Maria Kanakidou (FORTH)

Work package (WP)	WP3 Improving modelling and emission inventories for policy assessment
Milestone	M16 (M3.5)
Lead beneficiary	CNRS/ENPC
Means of verification	sensitivity metrics defined
Estimated delivery deadline	M16 (31/05/2023)
Actual delivery deadline	30/05/2023
Version	Final
Reviewed by	Work Package leaders
Accepted by	Project Coordination Team
Comments	This document summarizes several proxies that can be used to characterize the sub-grid variability of pollutants and will allow us to connect the city scale 1 km x 1 km maps that will be delivered by the mesoscale models in T3.3 with the local city and street scale models.

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

Table of Contents

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

1. About this document

WP3 focuses on improving modelling and emission inventories and enhancing ability air quality (AQ) and citizens' health policy assessment. This WP will develop a service tool (ST) coupling urban and regional-scale models for simulating novel AQ health indicators, their source apportionment and emissions to support policy implementation at the city scale and will be demonstrated in WP4 pilot studies. WP3 specific objectives:

i) Better quantifying emissions of gaseous and particulate pollutants at spatial-temporal scales relevant to urban areas.
ii) Improved simulation of urban pollutants disentangling the urban, regional and long-range transport contributions.
iii) Coupling urban and regional scales and optimizing use of new observations, source contributions and remote sensing data;
iv) Innovative modelling of the source contribution of reactive oxygen species (ROS), non-exhaust vehicle PM emissions, nanoparticles, secondary PM, PM and O_3 precursors, such as VOCs, NOx and NH_3.

In this context T3.4 focuses on implementing novel AQ indicators in tools supporting policy decision making to improve citizen health. This task tests a pilot system that can be used to assess efficiency of policy measures at different scales (city, national, EU), building on the new emission inventories (T3.2) and the improved modelling (T3.3) using the novel (ROS, source apportionment (SA) information, nanoparticles) and AQ indicators. The models cover a range of scales starting from regional and reaching out to urban areas. Several strategies will be used to quantify the local vs. non-local contribution to city AQ, incl. particulate SA technology and the novel "local fraction" methodology available in the EMEP MSC-W model. The models will be run for the European domain for a period of several years covering different meteorological and emissions conditions. These will simulate the changes in novel AQ indicators and SA resulting from specific local (WP4), national and EU policies are assessed by comparing the model results to the source apportionment data from WP1, historic data and other observational data available (e.g. from European Environment Information and Observation Network). The information on source apportionment at 1x1 km2 resolution is further nested over the pilots (T4.3) at higher resolution in local scale models to refine the guidance for upscaling advanced PM exposure indicators in T5.5. The sub-grid variability (< 1 km x 1 km) are reflected through the provision of sensitivity metrics that are defined in this milestone (M16 (M3.5)) and that will provide variability range in the source specific 1 km x 1 km maps for population exposure studies (WP2).

An assumption is made in the models that the grid boxes are homogeneous, i.e., that inside each grid model the pollutant concentrations (emissions and meteorological parameters) are everywhere the same. However, we know that this is not true and in the real world there is significant variability, in particular, for short lived pollutants and close to the sources. This so-called sub-grid variability characterizing the variability of the concentrations of pollutants, is introducing uncertainties in the computed pollutant concentrations, in particular, for models having relatively large grid boxes and needs to be evaluated. In cities, this sub-grid variability may be particularly high in streets, affecting the exposure indicators. Metrics of sub-grid variability are therefore defined for concentrations, and they can be applied to the concentrations of any pollutant that is simulated at both the local and regional scales. Sub-grid variability varies with pollutant, since it depends also on pollutant lifetime, but also with time. So, it is important to also provide the time horizon to which the sub-grid variability refers.

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

In cities, sub-grid variability along traffic axes and streets is particularly important for NO_2, black carbon (BC), organic matter (OM), number of particles and $PM_{2.5}$ and PM_{10}.

Sub-grid variability is also important for exposure. In the present milestone we present several proxies that can be used to characterize the sub-grid variability of pollutants and will allow us to connect the city scale 1 km x 1 km maps that will be delivered by the mesoscale models in T3.3 with the local city and street scale models.

This is a public document, available in the RI-URBANS website (https://riurbans.eu/work-package-3/#milestones-wp3). The document will be distributed to all RI-URBANS Partners for their use and submitted also to European Commission as the RI-URBANS milestone M16 (M3.5).

2. Variability of the concentrations on the grid cells of the regional domain

The variability of the concentrations may be estimated for each 1 x 1 km^2 grid cell of the regional domain.

To estimate this variability, the average, minimum and maximum concentrations of the local-scale model should be computed, as detailed in section 2.1. The statistics used to estimate the variability are then defined in section 2.2.

2.1. Average grid-cell concentrations of the local-scale modelling

The average concentrations at the local scale $\overline{C_{loc}}$ may be compared to the average concentration of the regional-scale model, as in Lugon et al. (2021). To do so, the concentrations need to be averaged over each grid cell of the regional domain. The minimum and maximum concentrations at the local scale may also be compared to the minimum and maximum concentrations at the regional scale.

2.1.1 Case of evenly gridded local-scale models

The concentrations should be averaged over the grid cells determined at finer resolution, e.g., 100 x 100 m^2 grid cell:

$$\overline{C_{loc}} = \frac{1}{n} \sum_{i \in I, j \in J} C_{loc}\,(i,j)$$

where *i,j* are the indices over the finer resolution, and *I,J* are the indices of the regional-scale grid, and *n* is the number of local cells in a regional-scale grid.

2.1.2 Case of non-gridded local-scale models, e.g., street-network models

In the case of street-network models, the concentrations should be averaged over the streets that are in the grid cells of the regional model. Because the streets have different lengths, the concentrations in each street are weighted by the length of the street:

$$\overline{C_{loc}} = \frac{1}{\sum_{s \in I,J} L_s} \sum_{s \in I,J} L_s C_{loc}\,(s)$$

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

2.2. Comparison of average concentrations at local and regional scales

The averaged concentrations over a specific time period (a month or a year) may be compared. However, as the sub-grid scale variability is not only a spatial variability but also a temporal variability, it is recommended to compare not only the differences (bias) in averaged concentrations, but also the error. The normalised mean bias (NMB), normalised mean absolute error (NMAE) and standard deviation (SD) are defined as

$$NMB = \frac{\overline{C_{loc}} - C_{reg}}{C_{reg}} \qquad \text{with } \overline{C_{loc}} = \sum_t \overline{C_{loc}^t}$$

$$NMAE = \frac{\sum_t \left| \overline{C_{loc}^t} - C_{reg}^t \right|}{C_{reg}}$$

$$SD = \sqrt{\sum_t \left(\overline{C_{loc}^t} - C_{reg}^t \right)^2}$$

and $C_{reg} = \sum_t C_{reg}^t$, where C is the concentration, *loc* stands for local-scale modelling , and *reg* for regional-scale modelling (1 x 1 km^2 grid cell), and *t* represents the time. The average concentrations $\overline{C_{loc}}$ are defined on the grid cells of the regional-scale model. They correspond to the average computed in 2.1. To best characterize the variability, it is recommended to consider hourly variation of the concentration (or at shorter intervals).

Maps of the NMAE or the SD can be used to characterise the spatial variability over a specific time period (monthly or yearly), considering the hourly variability of the concentrations.

2.3. Sub-grid spatial variability

Spatial sub-grid variability within a regional model grid, averaged over the time, can be characterised by the standard deviation of the sub-grids within the grid. These distributions may tend to be more log-normally distributed than following normal distribution. The normalised standard deviation (NSD) of the sub-grids within any regional model grid can be written as

$$NSD = \frac{\sqrt{\frac{1}{n} \sum_{i,j} \left(C_{loc}(i,j) - \overline{C_{loc}} \right)^2}}{\overline{C_{loc}}}$$

where $\overline{C_{loc}}$ is the mean of the sub-grids and *n* the total number of sub-grids within the regional model grid. Note that $\overline{C_{loc}}$ is used rather than C_{reg}. This indicator then does not include the NMB which is the difference between $\overline{C_{loc}}$ and C_{reg}.

In addition, it is useful to know the ranges involved so the normalised max and min within a grid can also be determined as

$$NMax = \frac{\max(C_{loc})}{\overline{C_{loc}}} \quad \text{and} \quad NMin = \frac{\min(C_{loc})}{\overline{C_{loc}}}$$

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

3. Further characterisation of temporal variability

In section 2, the statistics to define the variability are averaged over a certain period of time (e.g., month) using a finer temporal representation of the concentrations (e.g., hourly). The sub-grid-scale variability may be much larger at specific hours of the day, for example during the traffic peak hours. To quantify this variability, the statistics defined in section 2 may be estimated over selected time intervals (rush hours on morning and evening of the day) by summing over the days of interest, and allowing to define daily cycles for variability.

4. Variability over a specific domain or sub-domain

The local-scale models may not be run over the whole domain of the regional-scale model, because all the regional cells may not correspond to urban areas for example. Therefore, a sub-domain focusing on the city or part of the city may be defined. Also, defining a sub-domain for the study of the variability may be useful to characterise the variability over specific areas, e.g., city center, immediate surburbs of the city etc. The formulae in section 2 may be modified to characterise the variability over the sub-domain. For example, the NMB, NMAE and SD can be re-written as follows:

$$NMB = \frac{\sum_{I,J\in D}\left(\overline{C_{loc}}(I,J) - C_{reg}(I,J)\right)}{\sum_{I,J\in D} C_{reg}(I,J)}$$

$$NMAE = \frac{\sum_{I,J\in D}\left|\overline{C_{loc}}(I,J) - C_{reg}(I,J)\right|}{\sum_{I,J\in D} C_{reg}(I,J)}$$

$$SD = \sqrt{\sum_{I,J\in D}\left(\overline{C_{loc}}(I,J) - C_{reg}(I,J)\right)^2}$$

where *I,J* are the regional-grid indices.

Characterizing the variability of subdomains allows us to evaluate it for specific subdomains, such as city centre, suburbs and rural areas, without maps.

5. Variability of exposure

The sub-grid variability of (the health-related indicator of) population exposure to air pollution may provide a more realistic estimate of exposure within a grid. Population exposure corresponds here to outdoor exposure. The outdoor population exposure is assessed by multiplying the population data at the residential address and the pollutant concentrations (C), simulated either in the 1 x 1 km^2 grid cell or using the finer resolution (100 x 100 m^2 or street resolution), as in Lugon et al. (2022).

An 'exposure scaling factor' may be defined to estimate the sub-grid covariance of the sub-grid population and the sub-grid concentrations. It corresponds to the ratio of the outdoor population exposure in the 1 x 1 km^2 grid cell to the average of the outdoor population exposure in the sub-grid cells/streets.

The population weighted concentration (PWC) in any model grid *(I,J)* with sub-grids *(i,j)* within that grid, is calculated using the following equation

$$PWC(I,J) = \frac{\sum_{i\in I,j\in J} pop(i,j)\, C_{loc}(i,j)}{\sum_{i\in I,j\in J} pop(i,j)}$$

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

where *I,J* denote the regional grid values, the *i,j* denote the sub-grid model values, and $pop(i,j)$ is the population with residential address within the grid cell i,j. The $PWC(I,J)$ may differ from the regional grid concentration $C_{reg}(I,J)$ due to any correlation between the population and concentration sub-grids within the regional model grid. An exposure scaling factor $ESF(I,J)$ can then be written for each regional grid that relates the calculated regional grid concentration $C_{reg}(I,J)$ to the derived $PWC(I,J)$ from the sub-grid model as follows:

$$ESF(I,J) = \frac{PWC(I,J)}{C_{reg}(I,J)}$$

In many cases this factor is > 1, which tells us that the sub-grid population is positively correlated with the sub-grid concentrations or that the sub-grid bias is positive. This value can also be < 1 if there is negative correlation between the sub-grid population and concentrations or if the sub-grid bias (NMB) is negative. This last point can happen when downscaling emissions that are emitted at height above the surface, such as residential combustion, shipping or industry.

The same statistics (NMB, NMAE) as in section 2 may also be defined using the outdoor population exposure instead of the concentrations. Note that the NMB for exposure is related to the ESF, i.e., ESF = NMB (exposure) +1.

6. Previous work and work planed using these metrics in RI-URBANS

The street-network MUNICH model coupled to the CTM model Polair3D has been applied over the Paris area (France) for the year 2014 (Lugon et al. 2021, 2022). Lugon et al. (2022) estimated that about 25% of the population of the Paris city was exposed in 2014 to concentrations higher than 20 µg m^{-3}. However, this percentage increased up to 90% when the local-scale (street) concentrations were taken into account in computing the exposition. In RI-URBANS, the sub-grid scale variability will be evaluated over Paris using the street-network MUNICH model coupled to the CTM model CHIMERE (1 km^2 over the Paris area). Two time periods will be studied: the winter 2020-2021 and the summer 2022. The spatial temporal variability will be studied for $PM_{2.5}$, the different compounds of $PM_{2.5}$, the number of particles and NO_2 (urban health-related pollutants), through the statistical indicators provided above.

Episode-CityChem (local-scale CTM model) is applied over the Greater Athens Area (Greece). The model has an ability to simulate the photochemical and transport processes for each 1 km^2 of the domain, with a concurrent output at the sub-grid scale (100 m). Therefore, the sub-grid variability of hourly mass concentration outputs over the urban —— of Athens will be studied. To this purpose, representative warm and cold-periods, pre-covid months will be selected to study the spatio-temporal variability for $PM_{2.5}$ and NO_2 (urban health-related pollutants), through the statistical indicators provided above. Averages over selected time intervals (rush hours of the day, typical daily cycles) will reveal the spatial variability of air pollution model outputs in the selected time frame.

The uEMEP/EMEP modelling system has been applied to Europe previously (Mu et. al., 2022) and will be further applied in RI-URBANS for Europe using EMEP grids at 0.1° and uEMEP grids at 250-100 m resolutions. The uEMEP calculations are carried out on average mean concentrations for this region. Further to this, the hourly calculations can be carried out for individual pilot cities, notably Paris and Athens for comparison with the other sub-grid modelling. For compatibility with the other models 1 km grids will be made from aggregated uEMEP sub-grids scaled by the NMB of the regional model. The spatial variability will be studied for $PM_{2.5}$, NO_2 and the Oxidation potential (OP).

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

7. References

Lugon L., Sartelet K., Kim Y., Vigneron J., Chrétien O. (2021), Simulation of primary and secondary particles in the streets of Paris using MUNICH. Faraday Discuss., 226, 432-456, doi: 10.1039/D0FD00092B.

Lugon L., Kim Y., Vigneron J., Chrétien O., André M., André J.-M., Moukhtar S., Redaelli M., Sartelet K. (2022), Effect of vehicle fleet composition and mobility on outdoor population exposure: A street resolution analysis in Paris. Atmospheric Pollution Research, 13, 5, 101365, doi: 10.1016/j.apr.2022.101365.

Mu, Q., Denby, B. R., Wærsted, E. G., and Fagerli, H., 2022: Downscaling of air pollutants in Europe using uEMEP_v6, Geoscientific Model Development, 15, 449–465, doi:10.5194/gmd 15-449-2022, URL: https://gmd.copernicus.org/articles/15/449/2022 .

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245