First outcomes of the fairmode & aquila intercomparison exercise on spatial representativeness

Olivier Kracht, José Luis Santiago, Fernando Martin, Antonio Piersanti, Giuseppe Cremona, Gaia Righini, Lina Vitali, Kevin Delaney, Bidroha Basu, Bidisha Ghosh, et al.

FIRST OUTCOMES OF THE FAIRMODE & AQUILA INTERCOMPARISON EXERCISE ON SPATIAL REPRESENTATIVENESS

Oliver Kracht[1], José Luis Santiago[2], Fernando Martin[2], Antonio Piersanti[3], Giuseppe Cremona[3], Gaia Righini[3], Lina Vitali[3], Kevin Delaney[4], Bidroha Basu[5], Bidisha Ghosh[5], Wolfgang Spangl[6], Christine Brendle[6], Jenni Latikka[7], Anu Kousa[8], Erkki Pärjälä[9], Miika Meretoja[10], Laure Malherbe[11], Laurent Letinois[11], Maxime Beauchamp[11], Fabian Lenartz[12], Virginie Hutsemekers[13], Lan Nguyen[14], Ronald Hoogerbrugge[14], Kristina Eneroth[15], Sanna Silvergren[15], Hans Hooyberghs[16], Peter Viaene[16], Bino Maiheu[16], Stijn Janssen[16], David Roet[17] and Michel Gerboles[1]

[1]European Commission - Joint Research Centre (JRC), Ispra, Italy
[2]Research Centre for Energy, Environment and Technology (CIEMAT), Madrid, Spain
[3]Italian National Agency for New Technol., Energy and Sust. Econ. Devel. (ENEA), Bologna, Italy
[4]Environmental Protection Agency (EPA), Dublin, Ireland
[5]Trinity College Dublin (TCD), Dublin, Ireland
[6]Federal Environment Agency - Austria (FEA-AT), Vienna, Austria
[7]Finnish Meteorological Institute (FMI), Helsinki, Finland
[8]Helsinki Region Environmental Services Authority (HSY), Helsinki, Finland
[9]City of Kuopio, Kuopio, Finland
[10]City of Turku, Turku, Finland
[11]National Institute for Industrial Environment and Risks (INERIS), Verneuil-en-Halatte, France
[12]Public Service Scientific Institute (ISSeP), Liege, Belgium
[13]Walloon Air and Climate Agency (AwAC), Jambes, Belgium
[14]Netherlands National Institute for Public Health and the Environment (RIVM), Bilthoven, Netherlands
[15]Environment and Health Administration City of Stockholm, Stockholm, Sweden
[16]Flemish Institute for Technological Research (VITO), Mol, Belgium
[17]Flanders Environment Agency (VMM), Aalst, Belgium

Abstract: We are presenting an initial evaluation of the outcomes of the FAIRMODE & AQUILA intercomparison exercise (IE) on spatial representativeness (SR). To the best of our knowledge, this study provides the first attempt to quantitatively compare the range of methods used for estimating the spatial representativeness of air quality monitoring stations in Europe. As a common working basis, a shared dataset has been selected comprising modelling data and auxiliary information from the city of Antwerp. Based on this, 11 teams from 10 different countries provided their SR estimates for PM_{10} and NO_2 at one traffic site, and for PM_{10}, NO_2 and O_3 at two urban background sites. The main objective of this exercise was to evaluate the possible variety of spatial representativeness results obtained by applying the range of different contemporary approaches to a jointly used example case study. The results of the IE revealed a considerable range of variation between the different SR estimates - not only in terms of the extent and position of the SR perimeters, but also in the technical procedures and the extent of input data effectively used. These outcomes do also underline the need for (i) a more harmonized definition of the concept of "the area of representativeness" and (ii) consistent and transparent criteria used for its quantification.

Key words: *spatial representativeness, representativeness area, intercomparison exercise, air pollution, air quality monitoring, measurements, modelling, harmonization*

INTRODUCTION

The elementary concept of spatial representativeness (SR) is based on determining the area to where the information observed at a monitoring site can be extended. For the case of an air quality monitoring station, the key question about SR is thus as to what extent a point measurement at this station is representative of the ambient air pollutant concentrations around it. In practical applications SR has sometimes been described by rather (over-) simplified geometrical concepts. However, subject to the site specific conditions and to the different SR conceptualisation deployed, SR areas can in reality have quite complex, irregular and even discontinuous shapes.

The European Commission has worked intensively on the implementation of a harmonized programme for the monitoring of air pollutants. The harmonization program relies on the European Directives 2008/50/EC and 2004/107/EC, which endeavour to improve the quality of measurements and data collection, and to ensure that the information collected on air pollution is sufficiently representative and comparable across the community. However, though these directives include several considerations about the order of magnitude of the SR of a monitoring site, no detailed provisions on the methods for assessing the SR are provided. Also in the scientific literature, there is no unified agreement to address this complex problem, and no well-established procedure for assessing SR has been identified so far.

DATASETS AND METHODS

As an initial preparatory step for this intercomparison exercise (IE), a feasibility study has been conducted by CIEMAT (Spain) during the first half of 2015 (Martin et al. 2015). This first step aimed at consulting the expert communities within FAIRMODE and AQUILA in order to identify (i) prospective candidate methodologies to be considered in an IE on SR, (ii) the requirements on shared datasets, (iii) the assessment of the comparability of the different types of SR results, and (iv) the limitations to be expected. Based on the outcomes of this feasibility study, the FAIRMODE Cross-Cutting Activity Group on Spatial Representativeness agreed that the scope of the IE should be a case study to estimate the SR of selected air quality monitoring stations located in the urban area of the city of Antwerp. As a central task, it was decided that the IE should aim at assessing the participants SR estimates for PM_{10} and NO_2 at one traffic station, and for PM_{10}, NO_2 and O_3 at two urban background stations. In the following, a suitable dataset had been prepared to be used in the IE (Kracht et al. 2016). The employed model results, comprising gridded annual means and time series for a number of 341 virtual receptor points, have been prepared by VITO (Belgium) by applying the RIO-IFDM-OSPM model chain (e.g., Lefebvre et al. 2013). Furthermore, the dataset compiled by VITO includes data from measurements of the regular Antwerp monitoring stations, individual sampling campaigns, emissions, traffic, population density, building information, and gridded CORINE land use data. An overview of the Antwerp modelling domain and the example of the annual average concentration field for NO_2 is provided in **figure 1**.

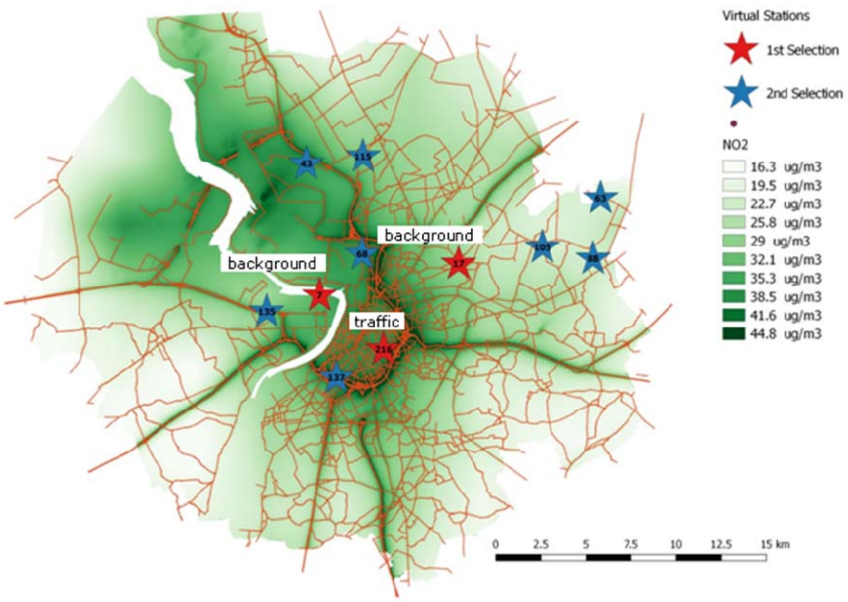

Figure 1. Overview of the Antwerp modelling domain, showing the annual average concentration field for NO_2 (green colours), the basic road network, and the positions of monitoring stations for which spatial representativeness estimates had been requested. Red stars highlight the three monitoring stations Linkeroever (7), Schoten (17) and Borgerhout (216), which had been selected for the main task of the exercise.

COURSE OF THE INTERCOMPARISON EXERCISE

Within the IE, 11 different teams from 10 different countries provided their SR estimates for the agreed sites and pollutants. As it was the main objective of this IE to evaluate the possible variety of spatial representativeness results obtained by applying the possible range of different contemporary approaches, all participating teams worked by applying their own selected methods and by using those parts of the dataset that they would normally require. In order to focus and to reasonably narrow down the range of conceivable SR approaches and definitions, it was however suggested to use the area of SR of the monitoring sites as a general concept to work with. In the course of the exercise, this concept of the SR area in fact turned out to be a useful indicator, and 10 of 11 teams were able to define shapes surrounding the stations under investigation, whereas 1 team rather worked towards a classification of the stations, as this was more common practice for SR evaluations in their member state. Participants were furthermore asked to provide estimates for the number of inhabitants within their calculated areas of representativeness. In this second step, most of the teams decided to compute zonal statistics for the population within the SR areas by overlaying the SR-area polygons with a population density raster file which was provide on a 100 m x 100 m grid. However, two teams (CIEMAT and RIVM) followed slightly different approaches adapted to the specific structure of their SR outcomes[1, 2].

RESULTS AND DISCUSSION

The SR methodologies applied within this exercise can roughly be distinguished as methods relying on air quality measurements, methods relying on proxy data, and methods relying on air quality model outputs. However, certain overlapping between these categories exists. From a rough categorization based on the selection of input data, 4 out of the 11 teams deployed the high resolution annual average concentration fields (made available from the RIO-IFDM-OSPM model outputs) as an immediate starting point, 3 teams performed a geostatistical interpolation of the 341 modelled virtual stations (based on time series and / or annual averages), 2 teams primarily focused on the use of concentration proxies, 1 team deployed their own computational fluid dynamics (CFD) model, and 1 team worked on principal component analyses (PCA) of concentration measurements.

The results obtained by the different teams revealed a considerable range of variation of the SR estimates. For a brief illustration[3], **figure 2** exemplifies six different estimates of the spatial representativeness areas obtained for NO_2, demonstrating a largely varying extent, shape and position of the SR area perimeter. Top panels of **figure 3** summarize the sizes of the SR areas obtained for NO_2 from all participating teams, whereas panels on the bottom of **figure 3** recapitulate the numbers of inhabitants within these areas of representativeness. Names of the reporting teams can be distinguished from the x-axis. These bar charts are sorted in descending order by the size of the SR area, or by the magnitude of the population within these areas, respectively. In cases where no results have been reported for that combination of site and pollutant, the team names are parenthesised and follow in alphabetical order. It has to be pointed out that the estimates concerning the size of the SR area and the results obtained for the number of inhabitants within these SR areas cannot be linked immediately to each other in a simple linear way. The relationship between these two different strata of the final results is in fact rather complex, as it emanates from the intersection of (i) SR areas that do not only differ in the extent of their perimeter, but also in their exact shape and position, with (ii) a spatially distributed and heterogeneous population density field. Finally, from a methodological point of view, it was an important observation that even on a shared dataset, individual teams made a clearly diverse choice on those subsets of the data (average concentrations,

---

[1] As a consequence of the CFD approach used by the CIEMAT team, areas covered by buildings have by principle not been part of their estimated SR area. For calculating the amount of population within the SR area, the gridded population density data therefore needed to be adjusted beforehand to correct for the proportion of built-up areas within in each grid cell. In this way the CIEMAT team re-allocated the part of the population density that was intersecting with the built-up area to the open area of each grid cell. Afterwards these adjusted population density data could be intersected with the SR-area polygons as described before.

[2] The RIVM team worked towards a station classification based on PCA. It was then assumed that the geometry of the traffic station is such that is is representative of (at least) 100 meters of street, as required by the AQD. For background stations a representative area of 1000x1000 meters was assumed. The numbers of inhabitants within the SR areas stated by RIVM are therefore providing a lower limit (>=), as opposed to the finite numbers derived by the other teams.

[3] Unfortunately, within the scope of this extended abstract only a small excerpt of the comprehensive results can be shown.

concentration time series, emissions, population densities, traffic, land cover, building geometries, etc.) that had been substantial or supportive to their particular SR method. It should deserve a closer look, as how far this indicates that harmonized recommendations should be established for the input requirements, too.

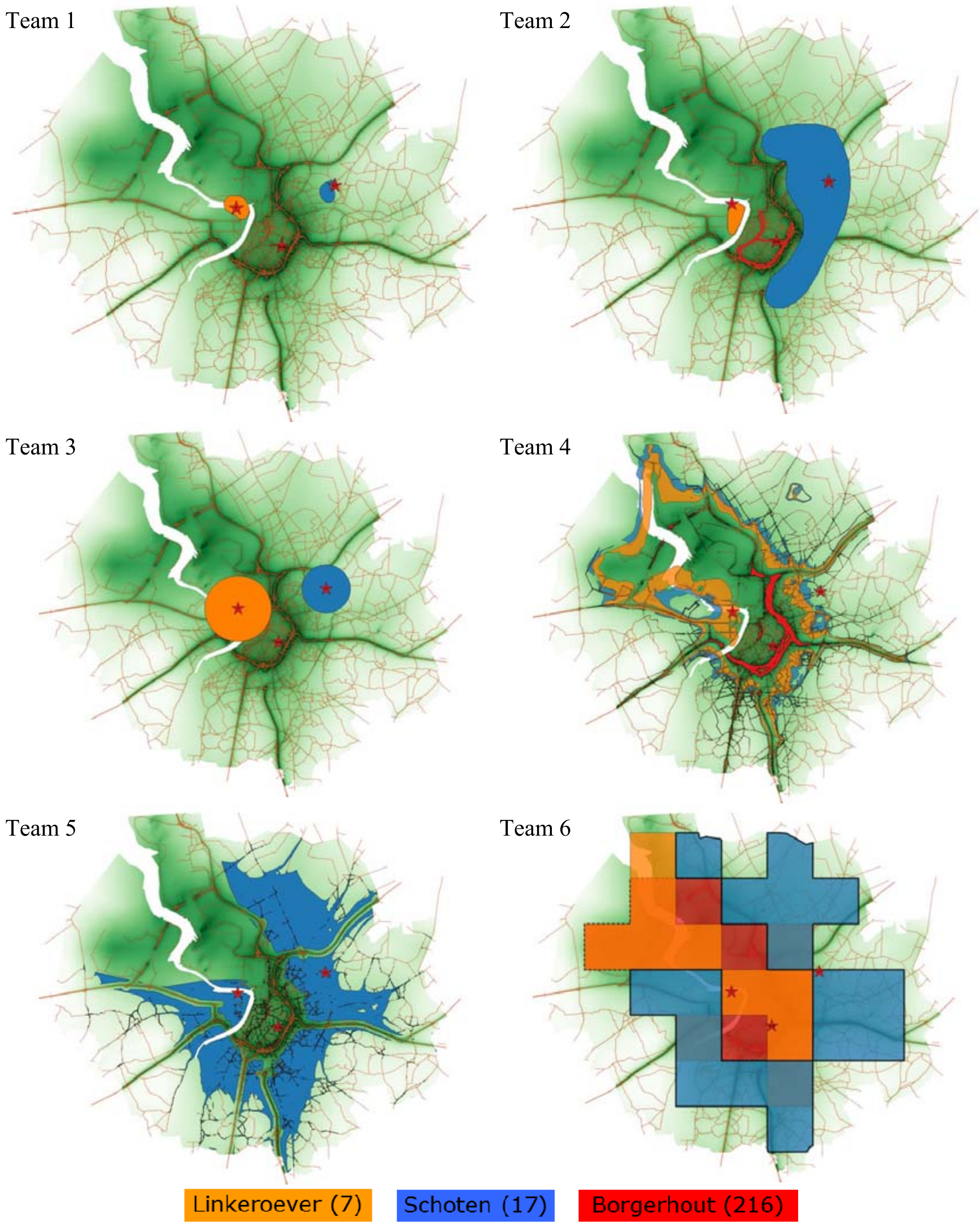

Figure 2. Examples of six different estimates obtained for the spatial representativeness areas for the pollutant NO_2 at the urban-background stations Linkeroever (orange) and Schoten (blue), and at the traffic station Borgerhout (red). Note that for the simplification of this figure the spatial representativeness areas of the three different stations have been overplotted (Linkeroever first, then Schoten, and finally Borgerhout on the top layer).

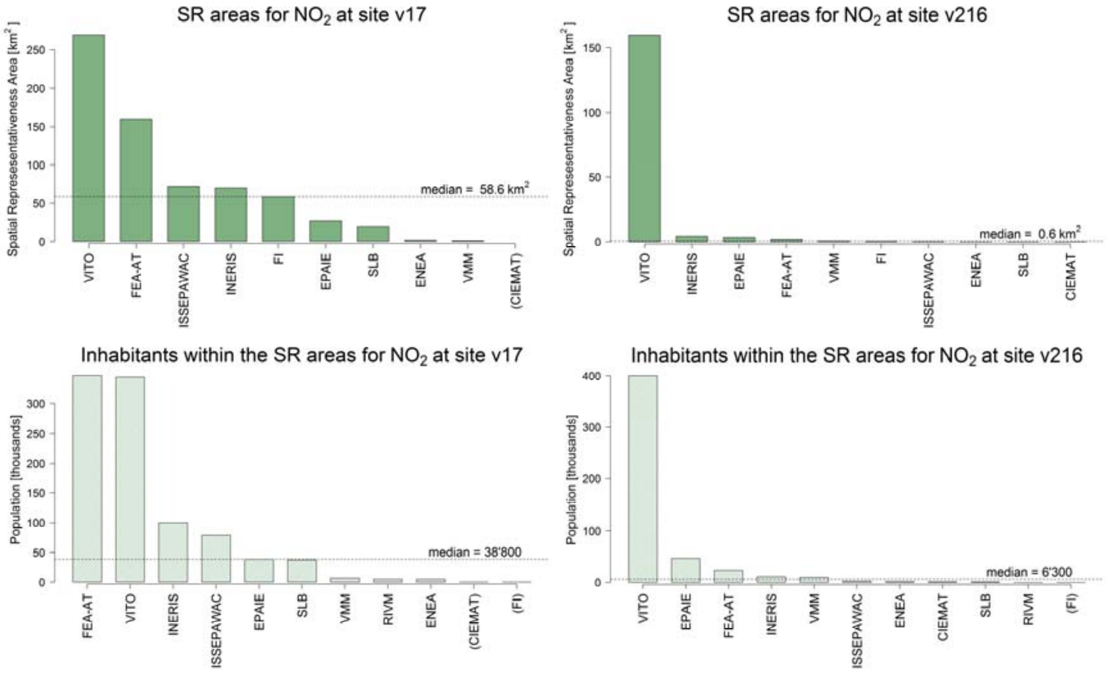

Figure 3. Spatial representativeness estimates obtained for the pollutant NO_2 at the urban-background site Schoten (v17) and at the traffic site Borgerhout (v216). Panels on the top display the size of the SR areas. Panels on the bottom summarize the number of inhabitants within these areas of representativeness. Parenthesised team names indicate that no results have been reported for that combination of site and pollutant.

SUMMARY & CONCLUSIONS

To the best of our knowledge, this study is the first attempt to investigate systematically the differences in spatial representativeness (SR) estimates that are achieved by applying a large set of SR approaches to the same common dataset. Our tests are providing quantitative evidence for a large variation in between the results obtained by the range of different contemporary methods. The considerable spread of the results obtained by the different teams nevertheless concerns the extent and position of the estimated SR area perimeters, but also the technical procedures and the extent of input data effectively used. In fact, though the general concept of the area of SR proved to be a useful indicator to work with, some important differences revealed regarding the details of the underlying concepts and definitions employed. These differences require detailed evaluations in order to identify the major factors triggering and controlling this variability, which could be found amongst (1) the basic principles of the methods, (2) the parameterization of the similarity criteria and thresholds, (3) the effective use of input data[4], and (4) the detailed conceptualization and definitions of SR. We consider that the diversity observed in this exercise requires the experts community to take further efforts towards a harmonized definition of the concept of "the area of representativeness" and in eliminating unnecessary differences in the SR methodologies.

REFERENCES

Kracht, O., Hooyberghs, H., Lefebvre, W., Janssen, S., Maiheu, B., Martin, F., Santiago, J.L., Garcia, L. and Gerboles, M. (2016): FAIRMODE Intercomparison Exercise - Dataset to Assess the Area of Representativeness of Air Quality Monitoring Stations, 267 pp, *JRC Technical Reports, **102775***, EUR 28135 EN, ISBN 978-92-79-62295-3 (PDF), DOI 10.2790/479282.

Lefebvre, W., Van Poppel, M., Maiheu, B., Janssen, S., and Dons, E (2013): Evaluation of the RIO-IFDM-street canyon model chain, *Atmospheric Environment*, ***77***, 325-337.

Martin, F., Santiago, J. L., Kracht, O., Garcia, L. and Gerboles, M. (2015): FAIRMODE Spatial Representativeness Feasibility Study, 82 pp, *JRC Technical Reports, **96827***, EUR 27385 EN, ISBN 978-92-79- 50322-1 (PDF), DOI 10.2788/49487.

[4] Though the same set of input data had been provided to all participants, sub-selection from these input data was effectively different in between teams.