Science of the Total Environment 988 (2025) 179824

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Estimating the air quality standard exceedance areas and the spatial representativeness of urban air quality stations applying microscale modelling

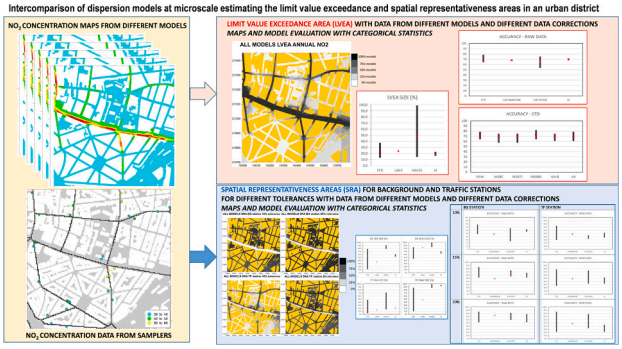
F. Martín [a,*], V. Rodrigues [b], J.L. Santiago [a], J. Sousa [c], J. Stocker [d], S. Janssen [c], R. Jackson [d], F. Russo [e], M.G. Villani [e], G. Tinarelli [f], D. Barbero [f], R. San José [g], J.L. Pérez-Camanyo [g], G. Sousa-Santos [h], L. Tarrason [h], J. Bartzis [i], I. Sakellaris [i], Z. Horváth [j], L. Környei [j], X. Jurado [k], N. Reiminger [k,l], N. Masey [m], S. Hamilton [m], E. Rivas [a], C. Cuvelier [n], P. Thunis [n]

[a] CIEMAT, Research Center for Energy, Environment and Technology, Avenida Complutense 40, 28040 Madrid, Spain
[b] CESAM & Department of Environment and Planning, University of Aveiro, 3810-193 Aveiro, Portugal
[c] VITO NV, Flemish Institute for Research and Technology, Boeretang 200, 2400 Mol, Belgium
[d] Cambridge Environmental Research Consultants (CERC), UK
[e] ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, 40129 Bologna, Italy
[f] ARIANET S.r.l., via Crespi 57, 20159 Milano, Italy
[g] Computer Science School, Technical University of Madrid (UPM), Campus de Montegancedo, s/n, 28660 Madrid, Spain
[h] NILU - The Climate and Environmental Research Institute, Norway
[i] University of Western Macedonia (UOWM), Dept. of Mechanical Engineering, Sialvera & Bakola Str., 50132 Kozani, Greece
[j] SZE, Széchenyi István University, Győr, Hungary
[k] AIR&D, Strasbourg, France
[l] ICUBE Laboratory, UMR 7357, CNRS/University of Strasbourg, F-67000 Strasbourg, France
[m] Ricardo, UK
[n] European Commission, Joint Research Centre (JRC), Ispra, Italy

HIGHLIGHTS

- Microscale models predict limit value exceedance and spatial representativeness areas rather accurately.
- The CFD based modelling systems seem to provide better results predicting limit value exceedance and spatial representativeness areas.
- Scenario-based CFD approaches perform as well as full-year unsteady simulations
- Lagrangian, Artificial Intelligence and Gaussian models with street parametrizations yield quite good results.
- Model accuracy predicting spatial representativeness varies with station type being worse for traffic station.

GRAPHICAL ABSTRACT



* Corresponding author.
E-mail address: fernando.martin@ciemat.es (F. Martín).

https://doi.org/10.1016/j.scitotenv.2025.179824
Received 26 February 2025; Received in revised form 23 May 2025; Accepted 1 June 2025
Available online 6 June 2025

ARTICLE INFO

Editor: Pavlos Kassomenos

Keywords:
Urban air pollution
Microscale modelling intercomparison
Limit value exceedance area
Spatial representativeness area

ABSTRACT

This study builds upon the findings of a FAIRMODE intercomparison exercise conducted in a district of Antwerp, Belgium, where a comprehensive dataset of air pollutant measurements (air quality stations and passive samplers) was available. Long-term average NO_2 concentrations at very high spatial resolution were estimated by several dispersion modelling systems (Martín et al., 2024) to investigate the ability of these to capture the detailed spatial distribution of NO_2 concentrations at the microscale in urban environments. In this follow-up research, we extend the analysis by evaluating the capability of these modelling systems to predict the NO_2 annual limit value exceedance areas (LVEAs) and spatial representativeness areas (SRAs) for NO_2 at two reference air quality stations. The different modelling approaches used are based on CFD, Lagrangian, Gaussian, and AI-driven models.

The different modelling approaches are generally good at predicting the LVEA and SRAs of urban air quality stations, although a small SRA (corresponding to low concentration tolerances or the traffic station) is more difficult to predict correctly. However, there are notable differences in performance among the modelling systems. Those based on CFD models seem to provide more consistent results predicting LVEAs and SRAs. Then, lower accuracy is obtained with AI-based systems, Lagrangian models, and Gaussian models with street canyon parameterizations. The Gaussian models with street-canyon parametrizations show significantly better results than models using simply a Gaussian dispersion parametrization.

Furthermore, little differences are observed in most of the statistical indicators corresponding to the LVEA and SRA estimates obtained from the unsteady full month CFD simulations compared to those from the scenario-based CFD simulation methodologies, but there are some noticeable differences in the LVEA or SRA (traffic station, 10 % tolerance) sizes. The number of scenarios does not seem to be relevant to the results. Different bias correction methodologies are explored.

1. Introduction

The European Union (EU) has established a comprehensive framework to monitor and regulate air quality across its member states, aimed at protecting human health and the environment. This framework is primarily embodied in the Ambient Air Quality Directive (AAQD, EU, 2024/2881). The AAQD specifies limit values for the concentration of these pollutants in ambient air. Exceedance of these limit values occurs when pollutant concentrations surpass the thresholds established by the directive, either in the short term (e.g., daily or hourly exceedances) or over a longer period (e.g., annual averages). When exceedances are detected, member states are required to implement air quality plans to reduce pollution levels.

The AAQD 2024/2881 define "spatial representativeness" as an assessment approach whereby the air quality metrics observed at a sampling point are representative for an explicitly delineated geographical area to the extent that air quality metrics within that area do not differ from the metrics observed at the sampling point by more than a pre-defined tolerance level. This is a critical factor in the interpretation of air quality measured data and in the assessment of compliance with air quality standards. Each air quality sampling point is strategically placed based on specific criteria, including population density, traffic density, and industrial activity, to ensure that it captures the air pollution levels of the area it represents. However, the area over which a station's data is considered representative can vary depending on the type of pollutant, the station's location, the spatial distribution of pollutant emissions and local environmental conditions. For example, a station located in a densely populated urban area might be representative of only a few square kilometres, while a rural background station might represent a much larger area. Accurately determining the spatial representativeness of a station is crucial for ensuring that air quality measurement used to assess compliance with standards is valid and reliable. This involves not only selecting appropriate locations for monitoring stations but also using modelling techniques to understand how pollution disperses across different landscapes and under varying meteorological conditions.

The assessment of limit value exceedance areas and the spatial representativeness of air quality stations can be conducted using measured data, data from numerical modelling techniques or a combination (assimilation or fusion) of both. At urban microscale scale, modelling techniques need to be able to account for the influence of complex urban morphologies on the dispersion of pollutants identifying possible urban hot-spots. There are several types of modelling approaches. Some of them are based on models that calculate numerical solutions to fundamental fluid flow equations near complex urban geometries (e.g. Computational Fluid Dynamics, CFD); other numerically simpler models use parameterisations to take account of the influence of buildings on wind flow (e.g. Gaussian models). In between, there are Lagrangian models and artificial intelligence/machine learning models trained with CFD models. Martín et al. (2024) presented a discussion of the relative success of different approaches in simulating correctly the monthly NO_2 concentration distribution in an urban district of Antwerp (Belgium). The main study finding was that the more complex modelling techniques predicted the most realistic spatial distribution of NO_2 in terms of concentration gradients. However, it is important to additionally evaluate how well local-urban microscale modelling estimates other important indicators such as the limit value (or air quality standard) exceedance area (LVEA) and the spatial representativeness area (SRA) of urban air quality stations.

Air quality modelling techniques have been widely used for estimating the SRA of air quality monitoring stations primarily for non-urban or urban background stations. These include statistical and geostatistical methods (Janssen et al., 2012; Mishra et al., 2012; Righini et al., 2014; Li et al., 2019; Yatkin et al., 2020), probabilistic methods using kriging techniques (Beauchamp et al., 2018), and Chemical Transport Models (CTM) (Martin et al., 2014; Duyzer et al., 2015; Piersanti et al., 2015; Vitali et al., 2016), among others. Additionally, innovative approaches such as the integration of satellite data (Yu et al., 2018; Zhu et al., 2020; Bai et al., 2022, 2023) and machine learning techniques (Luo et al., 2023) have further expanded the possibilities for SRA assessment. More recently, Su et al. (2022) developed a framework using multidimensional Euclidean distance and K-means clustering to assess and improve the spatial representativeness of China's regional air quality monitoring network. Perillo et al. (2022) analyzed the impact of traffic-related air pollution in Dublin, Ireland, comparing air quality across different periods to better understand the representativeness of monitoring stations. These diverse methods highlight the continuous advancement in the field and the importance of selecting appropriate tools. For urban or traffic air quality stations, where high pollution levels are more frequent, there is a significant need to quantify areas exceeding air quality standards or limit values. However, this evaluation is not straightforward due to the inhomogeneous nature of urban domains.

Kracht et al. (2017) published the first attempt to investigate differences in SRAs associated with the application of distinct modelling approaches to a common dataset. This study details the outcomes of a collaboration between members of FAIRMODE and the European Network of Air Quality Reference Laboratories (AQUILA). It presents the results of an intercomparison exercise on SRAs for PM10 and NO_2 at one traffic site, and for PM10, NO_2, and O_3 at two urban background sites. The findings revealed a considerable range of dissimilarities between the different modelling systems, particularly in terms of the extent and location of the SRA. These outcomes highlighted the need for a more harmonized definition of the SRA concept and consistent criteria for its quantification. Hooyberghs et al. (2020) proposed a tiered approach to estimate the SRA of air quality stations, while Tarrasón et al. (2020) went beyond focusing on the siting of the sampling points considering the macroscale and microscale criteria as defined by the European directives. Janssen et al. (2023, 2024) provide recommendations for the estimation of SRA of monitoring stations and the estimation of the Exceedance Situation Indicators, both relevant parameters when reporting under AAQD.

The exceedance of air quality standards is usually estimated with measurements at air quality stations (see for example, EEA, 2025).

Since air quality models provide generally better spatial coverage than stations networks, they are usually applied for estimating the spatial extension of the area exceeding the air quality standards since >25 years ago (Van Aalst et al., 1998).

At the mesoscale, CTM models have been widely used for annual air quality assessment with the aim of determining what areas exceed the air quality standards. For example, Vivanco et al. (2018, among others) applied the CHIMERE model to define the areas where the limit values (European air quality standards) were exceeded for several pollutants as a contribution to the Spanish annual air quality assessment following the European legislation. Jiménez- Guerrero et al. (2008) used the CMAQ model coupled to the WRF meteorological model to estimate the concentrations of O_3, NO_2, CO and PM10 and compare with the limit values to determine the exceedances areas in Catalonia (Spain). Monteiro et al. (2007) also used the CHIMERE model for air quality assessment in Portugal determining LVEA areas for several pollutants including NO_2, ozone, etc. A relatively complex modelling approach based on the ADMS model and on measurements was used by Stedman et al. (2007) in the UK to estimate the PM10 and PM2.5 concentrations and determine the LVEA areas.

However, fewer studies have been conducted to determine LVEA in urban areas. Furthermore, the use of microscale modelling in air quality assessment is quite scarce. Among others, we can point out the works of (Johansson et al., 2022) that applied the urban air quality model ENFUSER, which is a combination of Gaussian puff and Gaussian plume modelling coupled to a regional model and a measurement assimilation method. Rivas et al. (2019) applied a scenario-CFD-simulation-based methodology for air quality assessment, spatial representativeness and health impacts evaluation in Pamplona (Spain). Using almost the same methodology, Santiago et al. (2022b) analyzed the NO_2 LVEA in 3 urban hot spots in Madrid and how they change when emission reduction strategies are applied. Criado et al. (2023) used the CALIOPE-Urban model with a data fusion methodology for correcting bias to determine the NO_2 limit value exceedance areas in Barcelona. Rafael et al. (2021) applied a Gaussian air quality model (WRF-URBAIR) with different methodologies (full-year simulation, scenarios for representative days and scenarios for a set typical meteorological conditions) to estimate PM10 and NO_2 annual concentrations to identify exceedances of EU and WHO limit values in Aveiro (Portugal) and Bristol (UK) regions. Reiminger et al. (2020a) applied CFD modelling in an urban hot spot at very high resolution (1 m) testing a scenario methodology for computing long-term NO_2 concentrations considering different number of scenarios related to the wind rose directions and wind speed distributions.

Lefebvre et al. (2013) developed a RIO-IFDM-street canyon model (ATMO-Street) chain to air quality assessment estimating exceedances areas in Belgium cities with a very high resolution.

Another example of using microscale air quality models at very high resolution in urban areas is Zhong et al. (2024) where the ADMS urban was applied for the West Midlands (UK). The objective was to determine the effects of traffic emission reductions on the LVEA.

Beauchamp et al. (2018) did a very interesting study on how to compute the SRA and LVEA areas in urban areas of France using geostatistical methodologies (kriging) to air quality station and passive sampler NO_2 concentration data to obtain high resolution maps of NO_2 concentrations and then applying probabilistic approach to estimate SRA and LVEA.

Following these last referred works, our paper focuses on analyzing how microscale modelling systems perform estimating SRA and LVEA in urban hot spots. An intercomparison of the microscale modelling systems based on high spatial resolution dispersion models has been carried out. The intercomparison is based on an exercise conducted over an urban district in Antwerp (Belgium). The first part of the study involved model performance inter-comparisons concerning the calculation of period-average spatial distributions of NO_2 (Martín et al., 2024). The primary objective of this current, follow-on study is to assess the suitability of distinct modelling systems for estimating the spatial representativeness of two urban air quality stations, and the determination of areas where air quality standards are likely to be exceeded. The paper is organized as follows: the studied domain, model configurations, and validation methodologies are described in Section 2; the results are presented and discussed in Sections 3 and 4, respectively, and finally, the conclusions are presented in Section 5.

2. Methodology

2.1. Description of the input data from the case study domain and exercise settings

This model intercomparison exercise is the second part of the study undertaken by Martín et al. (2024). It relates to the same urban district of the city of Antwerp (800 × 800 m^2), Belgium; which is an urban built-up area consisting of a mix of street canyons, open areas, 2–3 floors residential houses with small private gardens and some commercial and residential buildings of varying heights. As shown in Fig. 1, the study area is crossed by a main road (Plantin en Moretuslei) from East-West direction.

NO_2, PM10, PM2.5, O_3 and BC concentration data are collected on an hourly basis at two automated fixed monitoring air quality (AQ) stations (a traffic and an urban background stations), operated by the Flemish Environment Agency (VMM), and located at the main road (Fig. 1). NO_2 background concentrations were provided by the RIO model (Janssen et al., 2008).

A campaign involving the distribution of about 2000 passive samplers among Antwerp's citizens was organized by the University of Antwerp in the context of the citizen science project CurieuzeNeuzen (https://ringland.be/academie/curieuzeneuzen/overzicht/). 73 of the passive samplers from this NO_2 concentration dataset, corresponding to the period 2016-04-30 to 2016-05-28 and within the study area, were used for model evaluation.

Hourly meteorological data (wind speed and direction, temperature, relative humidity and total radiation) recorded at the nearby VMM measurement station were used by some of the modelling systems.

Traffic NOx and NO_2 emissions were available for a selection of major and secondary roads (see Fig. 2) from the official Flemish FASTRACE traffic emission model (version 2.1), based on COPERT 5 emission factors.

More details about the input data, experimental data and exercise settings can be seen in Martín et al. (2024).

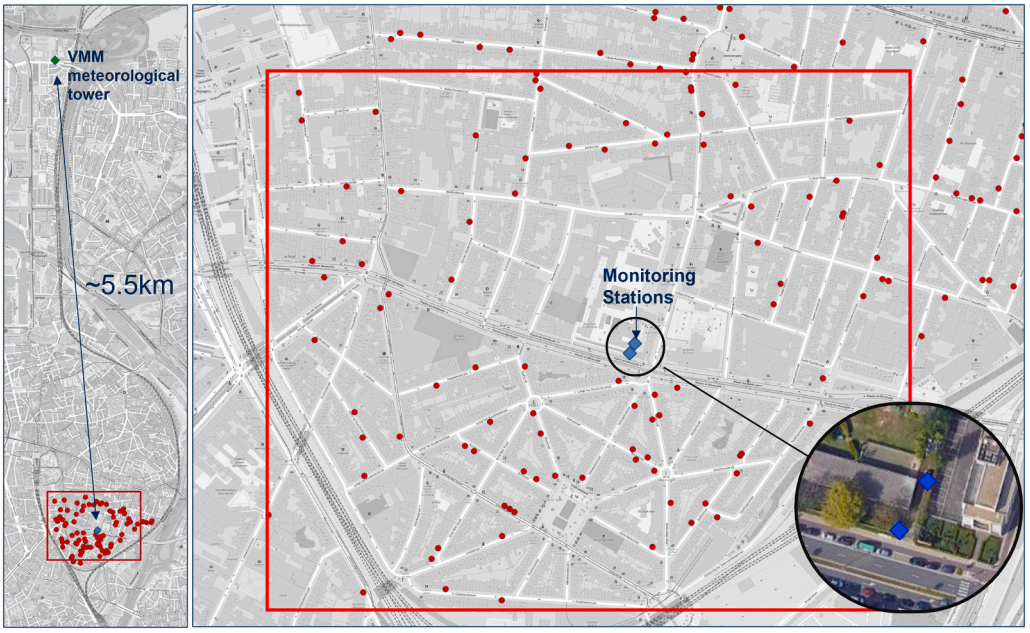


Fig. 1. The modelling domain (aprox 800 × 800 m^2) in the city of Antwerp, Belgium (red rectangle) and the sampler locations (red dots) are shown in two maps. The left map also shows the location of the VMM meteorological tower, whose meteorological data has been used by the modellers in this intercomparison exercise. The right map with an aerial picture includes the locations of the official monitoring stations 42R801 and 42R802 close to the major road Plantin en Moretuslei.

Fig. 2. Line segments with modelled traffic flow used in the FASTRACE model to derive traffic emissions. The red square represents the modelling domain.

2.2. Description of the modelling systems

The same set of modelling systems used in the previous study (Martín et al., 2024) was used to compute the long-term averaged pollutant concentrations; detailed information is provided in Tables 1–3. Additional results from a new Gaussian modelling system were included in this study corresponding to the contribution of the company RICARDO with the RapidAir modelling system. A modelling system is defined as a

Table 1
a. Description of the modelling systems (and model setup) used for directly estimating the long-term average concentrations with full period simulations on an hourly basis.

Modelling System	VITO-ATMOStreet	RICARDO-RAPIDAIR	CERC-ADMS
Microscale Model	ATMOStreet	RAPIDAIR	ADMS
Description	Gaussian plume (IFDM) with street canyon parameterization (OSPM)	Steady-state Gaussian plume model (AERMOD) with urban parameterization (AEOLIUS street canyon model)	Quasi-Gaussian
Urban morphology	Building footprints and heights	Data provided by VITO.	Buildings data processed to generate road by road street canyon, and gridded building, parameters. Parameters used to model dispersion in street canyons and within the urban canopy.
Computational Domain	Full Antwerp city	Full Antwerp city	Model input horizontal domain: 7822 m × 7750 m. Vertically, model extends to above the boundary layer.
Grid type	Receptor points	Regular grid	Irregular, source oriented. Higher receptor network where concentrations gradients are highest.
Grid resolution	10 m	1 m. Results provided at 1 and 2 m resolution	From 0.3 m to 25 m
Total number of cells			
Emissions data	Computed using the FASTRACE model, and applying time profiles	Traffic emission data + time profile provided by VITO	Traffic emission data + time profile provided by VITO
Chemistry	Simple chemistry	NOx as passive and NO_2 derived using the conversion scheme outlined in the SHERPA model (Degraeuwe et al., 2021), assuming constants as outlined in Romberg et al. (1996)	NO_x photolytic chemistry module, which accounts for fast, near-road oxidation of NO by O_3 to form NO_2
Atmospheric Stability	Stable, neutral and unstable	Stability dependent simulations. Hourly meteorology data is fed into AERMET/AERMOD and all stability calculations are controlled by these models	Full range, characterized by ratio of boundary layer height to Monin-Obukhov length
Meteorological boundary conditions	Data from the meteorological station: Antwerpen-Luchtbal.	Measured meteorological data in Antwerp (VITO data) converted into AERMET-ready formats. Upper air data estimated from this surface data to feed into AERMET	Measured meteorological data in Antwerp (VITO data) adjusted to account for the change in roughness at the meteorological measurement station compared to the dispersion site.
Meteorology within the domain	OSPM model includes an urban parametrization to modelling the street-canyons	AERMET model	Urban parameterization to modelling the street canyons (accounting for separate carriageways, pavements canyon asymmetry and porosity)
Pollution boundary conditions/ background concentrations		Background concentrations provided by RIO model.	Background concentrations provided by RIO model.
Assumptions and/or limitations	Gaussian plume parameterization. No account for explicitly 3D effects. Street canyon parameterization (OSPM model)	Gaussian plume model with urban parameterization (AEOLIUS street canyon model)	Street canyon flow and dispersion assumes uniform properties along the length of each modelled road with well-established flow patterns (no specific modelling of individual buildings or junctions); thermal effects on in-canyon flow (e.g., differential solar heating) neglected; concentration at a selected hour only depends on the emissions within the modelling domain and meteorological conditions at that hour.
References	Hooyberghs et al. (2022)	Masey et al. (2018)	Zhong et al. (2021), Hood et al. (2021), Biggart et al. (2020), Owen et al. (2000), Carruthers et al. (1994)

Table 1.b

Modelling System	ENEA-PMSS	NILU-EPISODE	SZE-OpenFOAM-U
Microscale Model	PMSS	EPISODE	OpenFOAM
Description	Mass consistent diagnostic reconstruction + conservation of momentum for meteorology & obstacle aware Lagrangian dispersion	3D grid Eulerian dispersion model with imbedded Gaussian dispersion model	CFD Unsteady RANS with k-eps turbulence modelling
Urban morphology	Data provided by VITO. Buildings simplified and grouped in blocks.	None. This implies that the Gaussian dispersion model does not have a street canyon parametrization.	Data provided by VITO.
Computational Domain	Buildings within an 800 m × 800 m region centered in the AQ station. No distance is needed from buildings to lateral boundaries; height of domain = 500 m; Hmax = height of tallest building	Larger domain defining all the road net available. The 3D grid had the resolution of 1 kmx1kmx20m. Then in the smaller 800mx800m domain, a set receptor points at the passive samplers' locations and regularly in 20mx20m and at a 2 m height.	2.6 km × 2.6 km with buildings in AOI. 8.6kmx8.6 km total simulation domain. Height of domain = 400 m
Grid type	Regular horizontal grid, terrain following vertical grid with 17 levels	regular	Octree mesh.
Grid resolution	3 m horizontal resolution	1kmx1km and 20mx20m	2 m on ground level, up to 32 m in the outside of the area of interest
Total number of cells		81 and 1600	3.3 million
Emissions data	Traffic emission data, together with the temporal profiles provided by VITO	Traffic emission data, together with the temporal profiles provided by VITO, except for hourly profiles	Traffic emission data, together with the temporal profiles provided by VITO

(*continued on next page*)

Table 1 (*continued*)

Table 1.b

Modelling System	ENEA-PMSS	NILU-EPISODE	SZE-OpenFOAM-U
Chemistry	No chemistry, passive pollutant. NOx/NO_2 transformation as post-processing based on NOx/NO_2 ratio of RIO model	Calculations of NO_2 are based on using photochemical equilibrium between the three fast-cycle reactive compounds NO, NO_2 and O3.	No chemistry yet. Passive scalar pollutant.
Atmospheric Stability	Depending on meteorological conditions and parameterized through Surface Layer scaling variables	Pasquill-Guifford stability class.	Neutral conditions
Meteorological boundary conditions	Data from the meteorological station: Antwerpen-Luchtbal.	WRF configured in a system with 3 one-way nested domains until the 1 kmx1km resolution for the larger domain.	Logarithmic boundary wind profile based on data from the meteorological station: Antwerpen-Luchtbal.
Meteorology within the domain	Diagnostic reconstruction using a simplified CFD	Data from the meteorological model WRF in the 1 kmx1km horizontal resolution. No additional wind fields for the street canyon locations were either included.	Diffusion coefficients depend on PBL height and turbulence diffusion coefficient.
Pollution boundary conditions/ background concentrations	Background concentrations provided by RIO model.	For NO_2 and NO we used RIO background and for O3 we used observations available.	Background concentrations provided by RIO model.
Assumptions and/or limitations		Gaussian plume model for the dispersion from the roads. No buildings accounted. Meteorology data from a model with coarse resolution, which means that for the 800mx800m we worked with two data points. Simplified meteorology in the 3D eulerian grid.	Diffusion coefficients equal in each direction.
References	Oldrini et al. (2017), Trini et al. (2018), Veratti et al. (2020), Russo et al. (2021), Villani et al. (2021), Barbero et al. (2021)	Hamer et al. (2020)	Horváth et al. (2016), Környei et al. (2021)

chain of models and submodels, including all necessary input data, and any post-processing (see Janssen et al., 2024, or EU-AAQD, 2024/2881). A clear example of modelling system could be the case of the methodology based on CFD models simulating representative scenarios and postprocessing methods to retrieve monthly or annual pollutant concentrations distribution in a selected spatial domain. Other modelling systems were based on performing a full month simulation CFD unsteady state simulations or running non-CFD models such as Gaussian, Lagrangian or AI/machine learning (convolutional neural network) models. All the CFD simulations were performed neglecting chemical reactions by assuming a passive scalar pollutant and only neutral atmospheric stability conditions were imposed. All the models provided their results for the same common domain (800 × 800 m^2) as described in the previous section but with their own grid resolution. More details of the modelling approaches used in this study can be found in Martín et al. (2024).

2.3. Methodology of modelling system evaluation

The aim of this study is to quantify how well the described modelling systems predict the long-term limit value exceedance and station spatial representativeness areas. Therefore, it is necessary to compare with observed long-term average pollutant concentrations. However, measurements at urban district level are difficult to obtain due to the limited air quality stations deployed in urban areas. Ideally, good spatial data coverage is needed, which captures the strong near-road pollutant concentration gradients. Hence, a dense sampling or measuring point network is required (Martín et al., 2024).

Such type of information can be obtained from campaigns deploying large numbers of passive samplers in the studied area. In this study, 73 passive samplers measuring NO_2 concentrations were deployed from April 30th to May 28th, 2016 (almost one month) throughout an 800 m × 800 m domain of the city of Antwerp (Belgium) (see Martín et al., 2024). 28 out of these 73 samplers were located on streets where traffic emission data was available (see Fig. 3). The NO_2 concentration data from these 28 passive samplers were used in this study. The rationale for utilizing only the 28 samplers from streets with emission data is to eliminate errors in the model predictions at the remaining sampler locations caused by the absence of emission data from nearby sources.

To evaluate the modelling system performance estimating LVEA and SRA, the concentrations measured by samplers in the campaign period (April 30–May 28, 2016) are used. Since the number of samplers is limited (28 samplers), we are not able to delimitate accurately the LVEA and SRA areas. However, we can identify the samplers with observed concentrations exceeding the limit value for NO_2 of 40 μg/m^3, which will be inside the "observed" LVEA. On the contrary, the samplers not exceeding the limit value would be out of the "observed" LVEA. In the same way, for the case of the "observed" SRAs of the two air quality stations located within the domain, the samplers with NO_2 concentration within the interval of the NO_2 concentration at each station plus/minus a tolerance margin (three tolerances were considered 10, 15 and 20 %) were considered inside the "observed" SRA of one of the stations with the selected tolerance.

Concerning the "modelled" LVEA and SRA, the samplers, where a modelling systems predict concentrations (averaged for the same period of the campaign, April 30–May 28, 2016) exceeding the limit value or being within the concentration similarity interval, will be part of the LVEA or SRA estimated by such modelling system, respectively. The "modelled" LVEAs and SRAs obtained with the different modelling systems were compared with the "observed" LVEA and SRA, as previously defined.

Since the data to be compared are categorical (inside or outside the LVEA and SRA from observations or from modelling results), the statistical metrics needs to be categorical as well. So, several categorical indexes were computed:

- Accuracy index. Trying to answer the question: How good are the models predicting the samplers in and out the LVEA or SRAs?

$$A = \frac{a + c}{a + b + c + d}$$

being:

a = number of hits predicting a sampler, which is inside the LVEA or SRA.

b = number of misses predicting a sampler, which is inside the LVEA or SRA.

c = number of hits predicting a sampler, which is outside the LVEA or SRA.

Table 2
a. Description of the models included in the modelling systems based on scenario simulations for retrieving monthly average concentrations through scenarios-based approaches.

Modelling System	VITO-OpenFOAM	CIEMAT-simple/windfactor/ detailed and CERC-CIEMAT	UOWM-ADREA	SZE-CxC
Microscale Model	OpenFOAM	Star CCM+	ADREA-HF	OpenFOAM
Description	CFD-RANS steady-state	CFD-RANS steady state	CFD-RANS steady state	CFD-RANS steady state
Schmidt number	0.5	0.3	0.7	0.5
Urban morphology	LOd1.2: 3D reconstruction based on building footprints and heights. Used CITY4CFD.	Data provided by VITO. Buildings simplified and grouped in blocks.	Data provided by VITO. No simplifications.	Data provided by VITO. No simplifications.
Computational Domain	Buildings: 1.5 × 1.5Km2. Avg distance to boundary ~500 m	Buildings within a rectangle 1 km × 1 km. Distance from buildings to lateral boundaries = 8H; Height of domain = 8H; Hmax = height of tallest building	Buildings within 1 km × 1 km. The horizontal domain has been extended by 20 m in each for the four (4) sides flat with roughness 0.3 m. Vertical dimension 300 m	2.6 km × 2.6 km with buildings in AOI. 8.6kmx8.6 km total simulation domain. Height of domain = 400 m
Grid type	Unstructured Polyhedral mesh. Build with SnappyHexMesh	Irregular. Polyhedral with hexahedral prism layers close to buildings and ground	structured, irregular	Octree mesh with cell sizes from 2 to 32 m.
Grid resolution	Resolution at building level ≈1 m	1 m	Horizontal resolution uniform 5 m; Vertical resolution nonuniform (2.0 m, 10.0 m)	2 m at ground level
Total number of cells	19 million	1,180,000	204x204x60	3.3 million
Emissions data	Computed using the FASTRACE model, and applying time profiles	Traffic emission data, together with the temporal profiles provided by VITO	Traffic emission data, together with the temporal profiles provided by VITO	Traffic emission data, together with the temporal profiles provided by VITO
Chemistry	Passive scalar. NO_2/NO_x correction based on measurements	No chemistry. Passive scalar pollutant.	No chemistry. Passive scalar pollutant.	No chemistry. Passive scalar pollutant.
Atmospheric Stability	Neutral atmospheric stability conditions.	Neutral atmospheric stability conditions.	Neutral atmospheric stability conditions.	Neutral atmospheric stability conditions.
Meteorological boundary conditions	Data from the meteorological station: Antwerpen-Luchtbal	Data from the meteorological station: Antwerpen-Luchtbal.	Data from the meteorological station: Antwerpen-Luchtbal	Data from the meteorological station: Antwerpen-Luchtbal
Meteorology within the domain	Computed by the microscale models	Computed by the microscale models	Computed by the microscale models	Computed by the microscale models
Pollution boundary conditions/ background concentrations	Background concentrations provided by RIO model.	Background concentrations provided by RIO model.	Background concentrations provided by RIO model.	Background concentrations provided by RIO model.
Assumptions and/or limitations	Concentrations ~1/v Passive pollutants; No thermal effects; Concentration at an hour only depends on the emissions within the modelling domain, NO_2/NOx ratio and meteorological conditions at that hour.	Concentrations ~1/v Passive pollutants; No Thermal effects; Concentration at an hour only depends on the emissions within the modelling domain, NO_2/NOx ratio and meteorological conditions at that hour.	Concentrations ~1/v Passive pollutants; No Thermal effects; Concentration at an hour only depends on the emissions within the modelling domain, NO_2/NOx ratio and meteorological conditions at that hour.	Concentrations ~1/v Passive pollutants; No Thermal effects; Concentration at an hour only depends on the emissions within the modelling domain, NO2/NOx ratio and meteorological conditions at that hour.
References	Janssen et al. (2008), Vranckx et al. (2015), Sousa et al. (2018), Sousa and Gorlé (2019), Hooyberghs et al. (2022), Paden et al. (2022)	Rivas et al. (2019); Santiago et al. (2013, 2017, 2021, 2022a); Santiago and Martin (2015), Sanchez et al. (2017); Parra et al. (2010)	Bartzis et al. (2015, 2020a,b, 2021, 2022); Sakellaris et al. (2022)	Horváth et al. (2016); Környei et al. (2021)

Table 2.b

Modelling System	AIR&D-CFD	AIR&D- AI	UPM-PALM4U
Microscale Model	OpenFOAM	FOLLOWAIR	PALM4U
Description	CFD – RANS unsteady state (time-average after convergence)	Encoder decoder Convolutional Neural Network (CNN) trained on CFD-RANS unsteady averaged results	CFD-LES
Schmidt number (*Sc*)	0.7	0.7	
Urban morphology	Data provided by VITO. No simplifications.	Data provided by VITO. No simplifications.	Data provided by VITO. 3D buildings as cubes (5x5x5 m)
Computational Domain	Buildings within a rectangle 1.3 km × 1.3 km. Distance from buildings to lateral boundaries = 200 m; Height of domain = 358 m; Hmax = 70 m	Buildings within a rectangle 1.3 km × 1.3 km.	3D grid: 200 (x) x 200 (y) x 70 (z) grid cells. 5 m spatial resolution.
Grid type	Regular. Unstructured. Cubical	Regular.	Regular.
Grid resolution	0.5 × 0.5 m (proximity to walls) up to 16 × 16 m (for the highest altitudes)	Resolution of 1 × 1 m^2	5 m, 200 × 200 grid cells
Total number of cells	12 million		2.8 million
Emissions data	Traffic emission data, together with the temporal profiles provided by VITO	Traffic emission data, together with the temporal profiles provided by VITO	Traffic emission data, together with the temporal profiles provided by VITO
Chemistry	Passive scalar. NO_2/NO_x correction based on DERWENT and BACHLIN parametrizations	Passive scalar. NO_2/NO_x correction based on DERWENT and BACHLIN parametrizations	Photo-stationary Chemistry mechanism
Atmospheric Stability	Neutral atmospheric stability conditions.	Neutral atmospheric stability conditions.	Stability dependent simulation

(*continued on next page*)

Table 2 (*continued*)

Table 2.b

Modelling System	AIR&D-CFD	AIR&D- AI	UPM-PALM4U
Meteorological boundary conditions	Data from the meteorological station: Antwerpen-Luchtbal	Data from the meteorological station: Antwerpen-Luchtbal	From WRF/Chem simulation 1 km spatial resolution Off-line nesting, BCs frequency 10 min. Wind components (u, v, w) Potential temperature and humidity Soil temperature and moisture.
Meteorology within the domain	Computed by the microscale models	Computed by the microscale models	Computed by the microscale models
Pollution boundary conditions/ background concentrations	Background concentrations provided by RIO model.	Background concentrations provided by RIO model.	From WRF/Chem simulation 1 km spatial resolution Off-line nesting, BCs frequency 10 min.
Assumptions and/or limitations	Concentrations ~1/v Passive pollutants; No thermal effects; Concentration at an hour only depends on the emissions within the modelling domain, NO_2/NOx ratio and meteorological conditions at that hour.	Concentrations ~1/v Passive pollutants; No thermal effects; Concentration at an hour only depends on the emissions within the modelling domain, NO_2/NOx ratio and meteorological conditions at that hour.	
References	Reiminger et al. (2020a, 2020b); Jurado et al. (2021)	Jurado et al. (2022, 2023)	Belda et al. (2021), Maronga et al. (2019), San Jose and Perez-Camanyo (2022, 2023),San Jose et al. (2021)

d = number of misses predicting a sampler, which is outside the LVEA or SRA

- False Alarm Rate (FAR). Answering to: what is the rate of prediction of samplers inside the LVEA or SRAs, when they actually are outside of them?

$$\mathrm{FAR} = \frac{\mathrm{d}}{\mathrm{a}+\mathrm{d}}$$

- Categorical Bias (BIAS). Answering to: are the models under (BIAS <1) or overpredicting (BIAS >1) the samplers inside the LVEA or SRA?

$$\mathrm{BIAS} = \frac{\mathrm{a}+\mathrm{d}}{\mathrm{a}+\mathrm{b}}$$

These metrics were computed separately for the case of LVEA and for the SRAs with the three studied tolerances (10, 15 and 20 %).

The predictions of concentrations obtained with the modelling systems can show some bias respect to the observed concentrations at the air quality stations (see Martín et al., 2024). In such case, some post-processing could be required for removing it and then, probably improving the estimates of SRAs and LVEA. Five different modelling data were used:

- Original or raw modelling data, i.e., the output data of using the modelling systems without any correction -.
- Modelling data modified using averaged sampler bias (*MSBC*) Modelling concentration estimated at sampler locations are modified using the average concentration bias of each modelling system's monthly NO_2 concentration prediction relative to the samplers' observations.

$B = \overline{C_{MS} - C_{OS}}$, $MSBC_S = C_{MS} + B$.

where C_{MS} is the concentration predicted by a modelling system at the sampler site S, C_{OS} is the concentration measured by the sampler S, B is the mean concentration bias of the predictions of a modelling system, $MSBC_S$ is the bias modified concentration at the sampler site S.

- Modelling data modified using reference station bias (*MSBBG* for background station, *MSBTF* for traffic station). The difference of monthly NO2 concentration prediction of modelling system at each station location respect to the observed data is used to modify all the sampler estimates from each modelling system. It was done for the background and traffic stations, respectively.

$\Delta C_{BGS} = C_{MBGS} - C_{OBGS}$; $MSBBG_S = C_{MS} + \Delta C_{BGS}$
$\Delta C_{TFS} = C_{MTFS} - C_{OTFS}$; $MSBTF_S = C_{MS} + \Delta C_{TFS}$

where C_{MBGS} is the concentration predicted by a modelling system at the sampler site where the background station is located *BGS*, C_{MTFS} is the same but for the sampler located at traffic station (*TFS)*. C_{OBGS} is the concentration measured by the sampler where the background station (*BGS)* is located, while C_{OTFS} is the same but for the sampler of the traffic station (*TFS)*. $MSBBG_S$ and $MSBTF_S$ are the modified concentration at the sampler location S by the concentration at the background and traffic stations, respectively.

- Modelling data modified using linear regression ($Ax + B$). The linear regression functions were computed for every concentration data set of the modelling systems with the observed concentrations from the samplers and the slopes A and the intercepts B were computed for every modelling system.
- Modelling data modified using linear regression with intercept zero (Ax). The coefficients A were computed for every modelling system as in the former case.

Results can be found in the Supplementary Material section.

3. Results

3.1. Annual average NO_2 concentration maps

In Figs. 4–7, the maps of 2016 annual averaged of NO_2 concentrations for the studied domain obtained from using the different modelling systems used in this study are shown. The maps are qualitatively very similar to those of the monthly corresponding to the samplers' campaign period (see Martín et al., 2024). In the Discussion section, these results will be extensively analyzed.

Table 3
a. Description of averaging methodologies involved in the modelling systems based on scenario simulations for retrieving monthly/annual average concentrations through scenarios-based approaches.

Modelling System	SZE-CxC	CERC-CIEMAT	CIEMAT detailed	CIEMAT simple	CIEMAT wind-factor
Microscale Model	OpenFOAM	Star CCM+	Star CCM+	Star CCM+	Star CCM+
Set of simulations required/ no. of scenarios	4–32 wind sectors	16 wind sector scenarios + wind speed bins from VITO met measurements + roadside & background pollutant concentration measurement	16 wind sector scenarios	16 wind sector scenarios	16 wind sector scenarios + wind speed bins
Criteria for selecting scenarios	Measured wind direction (VITO meteorological data)	Measured wind speed, wind direction, and peak or off-peak traffic flows	Using the measured wind direction (VITO meteorological data)	Using the measured wind direction (VITO meteorological data)	Using the measured wind direction (VITO meteorological data)
Procedure for retrieving annual/ monthly concentrations	Series of hourly concentration maps (gridded data) computed by assigning one scenario each hour and applying corrections by wind speed or by hour/date (emissions) and adding background concentrations. Annual/monthly concentration maps estimated averaging for the time series of maps for a year/month	1) Derive wind-speed and traffic emissions correction factors from measured concentration data. 2) Calculate the frequency of occurrence of each wind speed / traffic emissions combination on a sector by sector basis (monthly / annual). 3) Calculate weighted sum of 16 CFD gridded concentration maps	Series of hourly concentration maps (gridded data) computed by assigning one scenario each hour and applying corrections by wind speed or by hour/date (emissions) and adding background concentrations. Annual/monthly concentration maps estimated averaging for the time series of maps for a year/month	Weighted average based on probability density function (wind direction) + corrections by hour/date (emissions) and wind speed and adding background concentrations.	Weighted average based on probability density function (wind direction and wind speed scenarios) + corrections by hour/date (emissions) and wind speed and adding background concentrations.
Assumptions and/ or limitations		Dependence of concentrations on wind speed and traffic emissions is spatially homogeneous; Influence of atmospheric stability on solution dependent through wind speed magnitude only; Non-reactive pollutants; Thermal effects negligible; Concentration at a selected hour only depends on the emissions within the modelling domain and meteorological conditions at that hour.			

Table 3.b

Modelling System	VITO-Open FOAM	UOWM- ADREA	AIR&D-CFD	AIR&D-AI	UPM-PALM4U
Microscale Model	OpenFOAM	ADREA-HF	OpenFOAM	FOLLOWAIR	PALM4U
Set of simulations required/ no. of scenarios	36 Wind sectors scenarios	32 wind sector scenarios	18 wind sector scenarios	36 wind sector scenarios	Representative days 1 per period (month or year)
Criteria for selecting scenarios	Measured wind direction (VITO meteorological data)	Using the measured wind direction (VITO meteorological data)	Using the measured wind direction (VITO meteorological data)	Using the measured wind direction (VITO meteorological data)	Three rules: 1) Bias of the station data daily average vs monthly/annual average value (min. < 30 %) 2) Bias of the WRF/ Chem daily average vs station data monthly/ annual average value (min. < 30 %) 3) Correlation coefficient of hourly data of the WRF/Chem and station data for the representative day (min. > 0.6)
Procedure for retrieving annual/ monthly concentrations	Series of hourly concentration maps (gridded data) computed by assigning one scenario each hour and applying corrections by wind speed or by hour/date (emissions). NO_2/NOx ratio per hour and adding	32 Steady state reference simulations (i.e. one (1) per each sector) with wind velocity at the meteorological station 5 m/s. These 32 simulations are utilized to estimate the 8764 h applying corrections by wind speed	Series of hourly concentration maps (gridded data) computed by assigning one scenario each hour and applying corrections by wind speed or by hour/date (emissions) and adding background concentrations.	Series of hourly concentration maps (gridded data) computed by assigning one scenario each hour and applying corrections by wind speed or by hour/date (emissions) and adding background concentrations.	Averaging the 24 hourly maps of the representative day

(*continued on next page*)

Table 3 (*continued*)

Table 3.b					
Modelling System	VITO-Open FOAM	UOWM- ADREA	AIR&D-CFD	AIR&D-AI	UPM-PALM4U
Assumptions and/ or limitations	background concentrations. Annual/monthly concentration maps estimated averaging for the time series of maps for a year/month	and by hour/date (emissions) and adding background concentrations.	Annual/monthly concentration maps estimated averaging for the time series of maps for a year/ month	Annual/monthly concentration maps estimated averaging for the time series of maps for a year/ month	

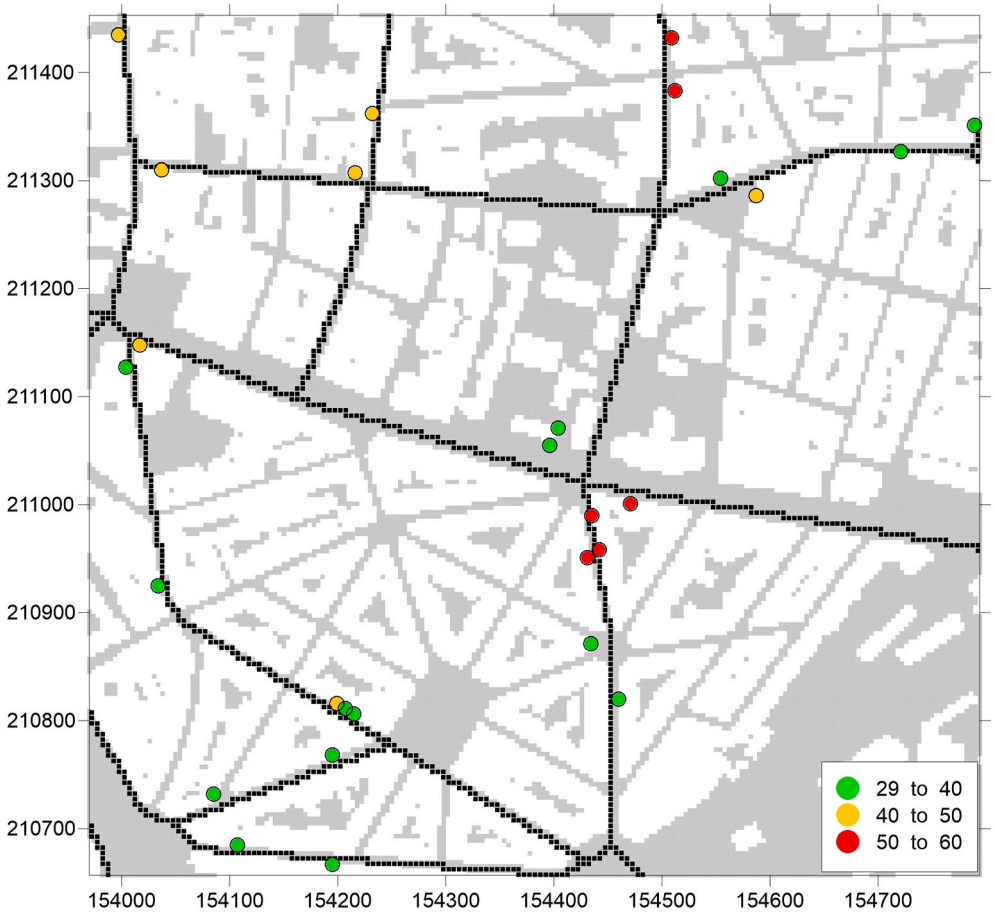


Fig. 3. *Map showing the location of the 28 passive samplers (colored dots) measuring NO_2 concentrations (μg/m^3) during the campaign period (April 30th to May 28th, 2016), which are sited in streets with available traffic emission data (black lines). Colors of the dots refers to the measured NO_2 concentrations. White areas are buildings, whereas gray zones are streets, squares or parks.*

3.2. Evaluation of the monthly LVEA estimates

In Fig. 8, the results of the categorical metrics for the predictions of samplers inside or outside the LVEA for the direct estimates of the modelling systems are shown. On average the values of the accuracy index for the estimates is roughly 70 % but ranges from 55 % to almost 80 %. The FAR is mostly low (about 20 % on average), but with BIAS generally lower than 1. It means that the modelling systems seem to mostly underpredicting the number of samplers inside the LVEA. Indeed, the modelling systems predict better the no-exceedance samplers (>80 % of hits) than the exceedances ones (<60 %) for the case of the raw predictions (see supplementary material, Fig. SM.1).

The CFD, AI and Lagrangian (in this order) modelling systems provide significantly high averaged values of the accuracy index (around 70 %). However, the Gaussian modelling systems provide a lower false alarm rate (around 10 % on average with maximum of 33 %), while the other model types have average values close to 20 % but with maxima of >35 % in the case of the CFD and AI modelling systems. The CFD and AI models have higher values of BIAS (0.78 and 0.75, respectively but underpredicting) but with large variability (0.35–1.30 for the CFD cases and 0.39–1.10 for the AI ones). The Lagrangian model has a rather low BIAS value (strong underprediction) close to the mean value of all the Gaussian models. However, Gaussian models have a very significant variability (0.0–0.90).

In addition, the same statistics were computed for the different modified modelling concentration results as indicated in Section 2.3. The detailed results can be seen in Supplementary Material Section (Fig. SM.4). The results will be discussed briefly in Section 4.2.2.

Additionally, it is important to investigate how different are the predictions of LVEA when using the same model but with different type of simulations or number of scenarios. This was done using the simulations of the SZE group, which consisted of a full-month unsteady

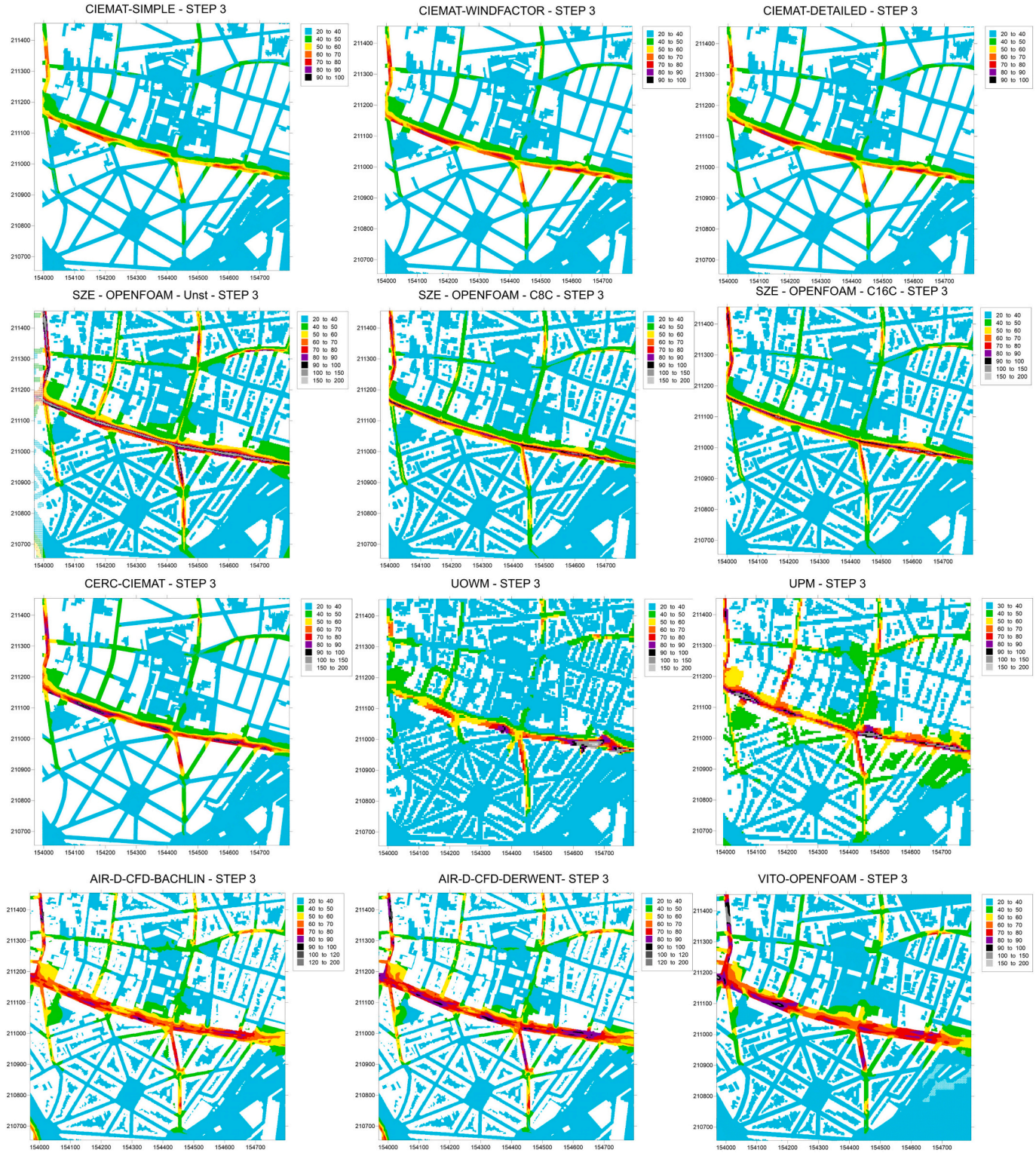


Fig. 4. Maps of 2016 annual average of NO_2 concentrations ($\mu g/m^3$) computed with the modelling systems based on CFD models.

OpenFOAM simulations and the application of a methodology for retrieving monthly NO_2 concentration data from different number of CFD simulated wind sector scenarios (4, 8, 16, and 32). In Fig. 9, the results of the accuracy index, false alarm rate and bias for the analyzed SZE simulations are depicted. There are no significant differences in the accuracy index among the unsteady full-month CFD simulations and the approaches using different number of scenario-based CFD simulations. FAR and BIAS for the unsteady full-month CFD simulation is slightly higher than for the cases of using scenario-based CFD simulations. As with most of the modelling systems, the unsteady full-month CFD simulation and the scenario-based CFD simulations are better predicting the non-exceedance areas (73–80 % of hits) than the exceedance ones (54–62 %). The full-month simulation has higher percentage of hits predicting the LVEA and lower for the non-exceedance area. The four cases using the scenario-based methodology have the same percentage of hits (see Fig. SM.2 in the Supplementary Material section).

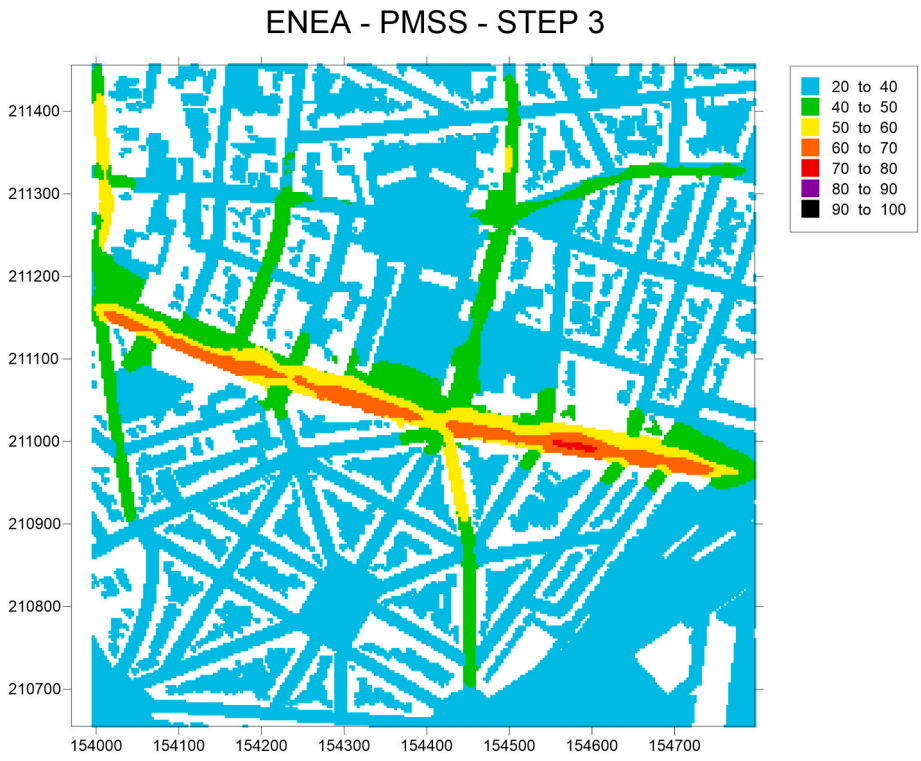


Fig. 5. Maps of 2016 annual average of NO_2 concentrations (μg/m^3) computed with the Lagrangian ENEA-PMSS model.

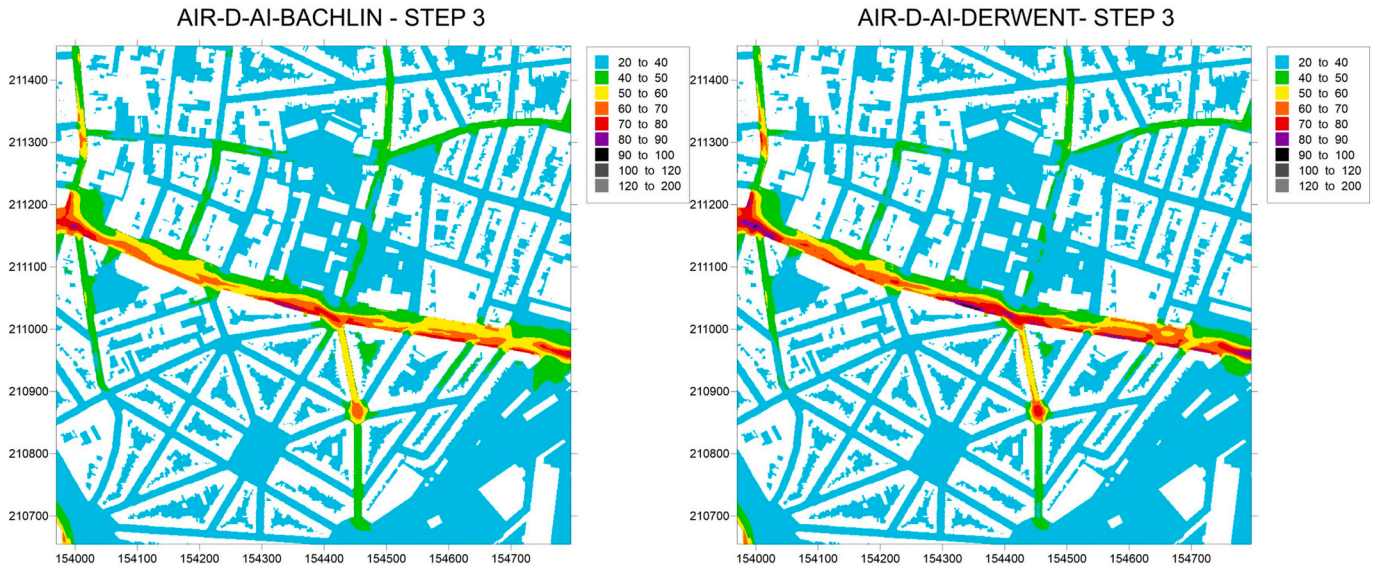


Fig. 6. Maps of 2016 annual average of NO_2 concentrations (μg/m^3) computed with the modelling systems based on AI models.

3.3. Comparison of estimated 2016 annual LVEAs

Fig. 10 shows the overlapping of the LVEAs estimated from the raw outputs of all the modelling systems. The LVEAs, which are common for most of the modelling systems, are darker, whereas the light gray areas correspond to the LVEA estimated by only a few modelling systems. It is evident that the LVEAs predicted by most of the modelling systems are the main streets and avenues, but there are differences in the shape and size of the LVEAs. These noticeable differences can be partly attributed to the different types of models (Figs. 11 and 12) but with significant variability.

3.4. Evaluation of the estimates of monthly SRAs

Fig. 13 shows the results of the accuracy index for the predictions of samplers inside or outside the SRAs of the two air quality stations for different tolerances and for the different types of modelling systems.

It is evident that accuracy index is higher for the SRA of the background station and for larger tolerances. For the background station, the accuracy index increases from 55 % on average for 10 % tolerance to roughly 80 % for 20 % tolerance. For the traffic station, the accuracy index also grows notably from 45 % on average for 10 % tolerance to 55 % for 20 % tolerance but with a very important variability. Accordingly, the modelling systems for low (10 %) tolerance predict better the SRA-

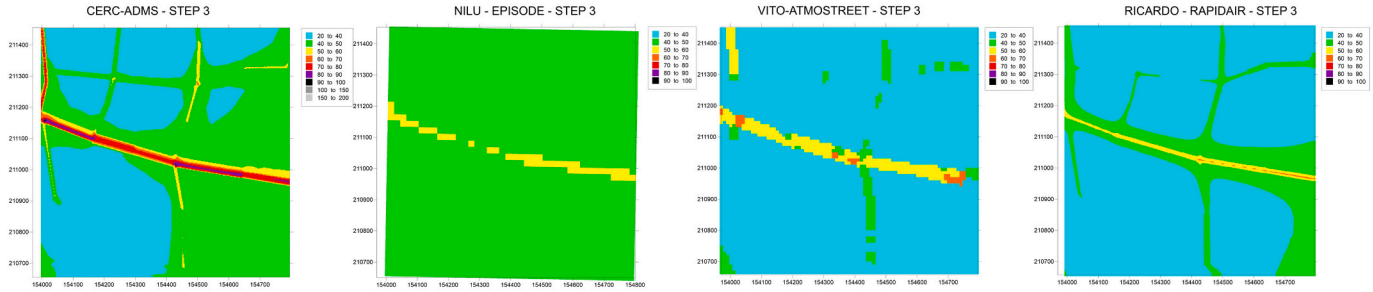


Fig. 7. Maps of 2016 annual average of NO_2 concentrations ($\mu g/m^3$) computed by the Gaussian modelling systems used in this study.

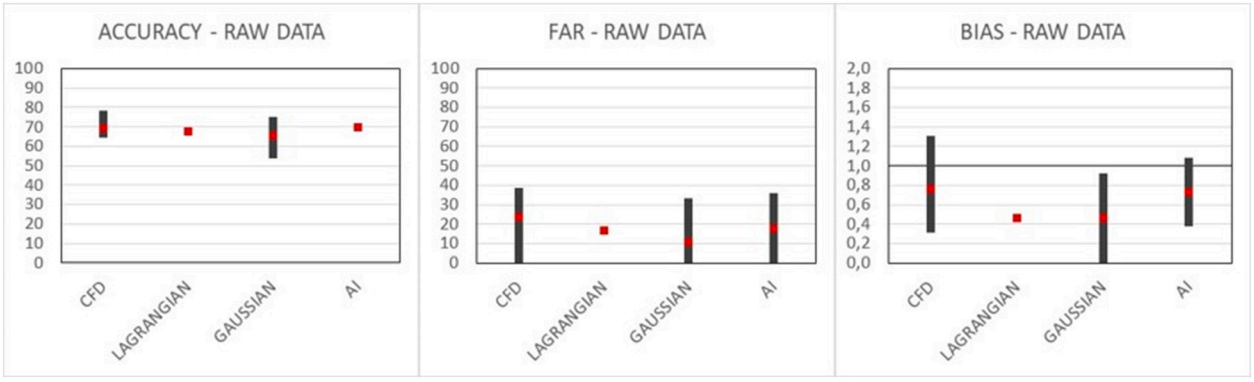


Fig. 8. Accuracy Index, False Alarm Rate (FAR) and Categorical Bias (BIAS) (maximum-minimum range in black and mean value in red) for predictions of the limit value exceedance area corresponding to the raw estimates of the modelling systems grouped by model types (CFD, Lagrangian, Gaussian and Artificial Intelligence (AI)).

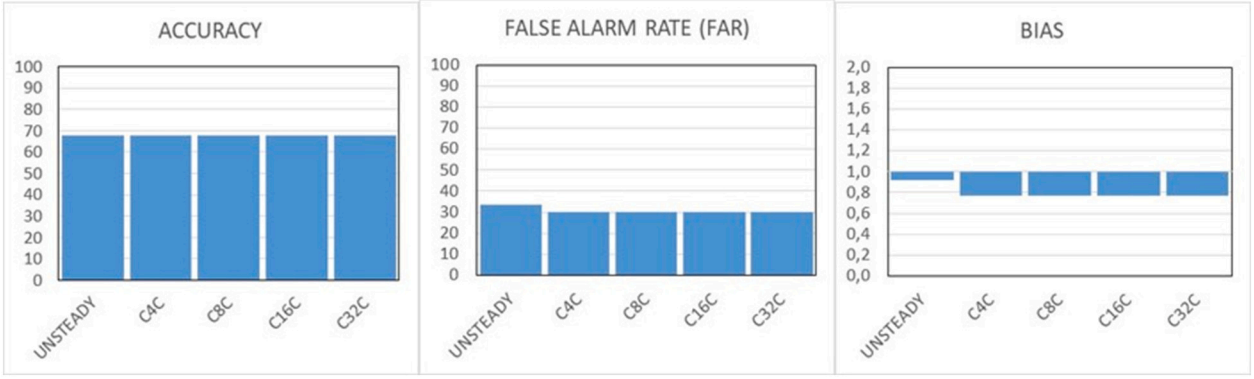


Fig. 9. Accuracy index (above), False Alarm Rate (FAR) (middle) and Categorical Bias (BIAS) (below) for the prediction of LV exceedance samplers predicted by the full-month unsteady OpenFOAM simulations and by applying methodologies based on wind sector scenarios (4, 8, 16 and 32) OpenFOAM simulations.

outside samplers (60 % of hits for both stations) than the SRA-inside ones (55 % of hits for the background station and <30 % for the traffic one) for the case of the non-modified concentration predictions. However, for high (20 %) tolerances the hits of samplers inside the SRAs increase a lot for both stations (see supplementary material, Figs. SM.5. and 6).

The CFD modelling systems provide the highest average of accuracy in all cases (different tolerances and both stations) ranging their average values from 60 to 75 % and 45–60 % for the background and traffic stations, respectively. Their accuracy values improve as tolerance increases and are better for the background station. Nevertheless, there is a strong variability in the results among the individual models, especially in the case of the traffic station, spanning from 22 to 85 % for 20 % tolerance (the most variability case). We have to keep in mind that 12 CFD modelling systems have been used in this study, while only one Lagrangian model, three Gaussian and two AI were involved.

The accuracy results for the AI models is good for the background station with a slight improvement when increasing tolerance, while such an improvement is not observed for the traffic station case. The differences between the two AI models seems to be somewhat larger for the traffic station and increase with tolerance for both stations.

The Lagrangian model scores its best average accuracy (73 %) for the background station with 20 % tolerance, while the average accuracy is roughly 50 % in the remaining cases (10 and 15 % tolerances for the background station, and all tolerances for the traffic one).

The accuracy values for the Gaussian models show a notable improvement when increasing the tolerances reaching averaged values of 72 and 58 % for the background and traffic stations, respectively for 20 % tolerances. Nevertheless, there are remarkable differences among the three Gaussian models, which are more important for 10 % tolerance

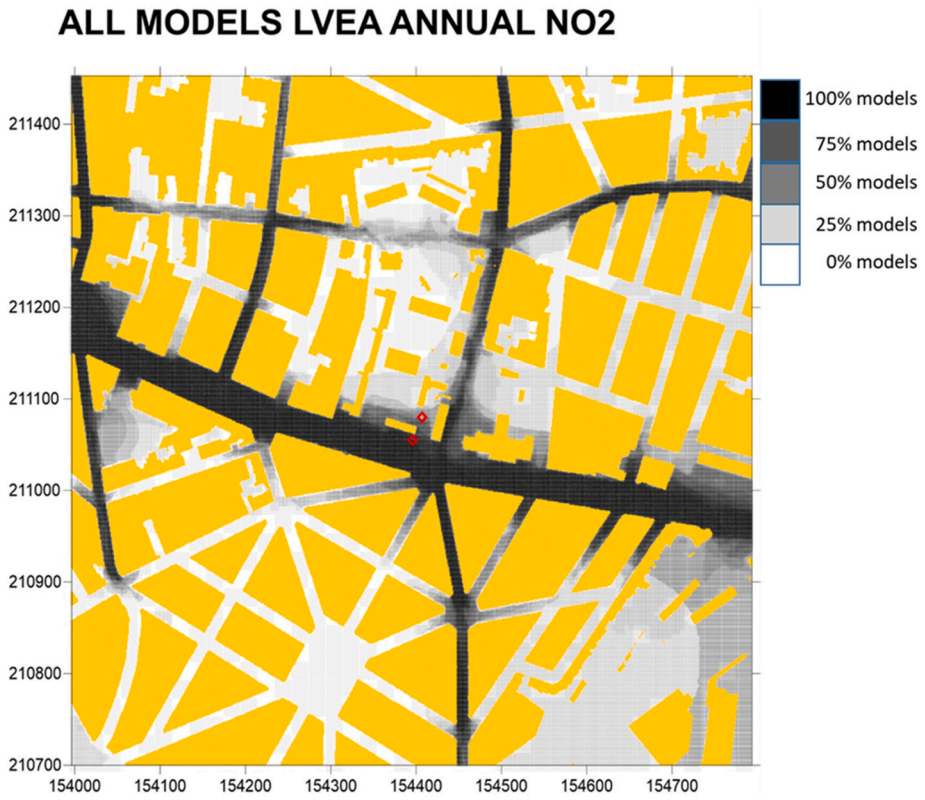


Fig. 10. Overlapping of the limit value exceedances areas computed by all the modelling systems. The darker the pixel, the more models estimate it is within the LVEA. The white color pixel indicates none model predicts it within the limit value exceedance area. Yellow means buildings.

for the background station (25–68 %) and for 15 and 20 % tolerances for the traffic one (32–78 %).

The false alarm rate (FAR) values are shown in Fig. 14. FAR is clearly higher for the traffic station (ranging the model-type-averaged FAR from 48 % to 58 % for 10 % tolerance and from 25 % to 54 % for 20 % tolerance) than for the background (ranging from 29 % to 55 % for 10 % tolerance and from 15 % to 29 % for 20 % tolerance).

Gaussian models have better FAR for both stations than other model types but with a noteworthy variability among the three Gaussian models, especially for the traffic station for all tolerances (0–82 % for 10 % tolerance, for example). For the traffic station, the AI models have higher FAR values for all tolerances (>50 %), whereas the Lagrangian model has the highest FAR values (>50 %) for the background station for 10 and 15 % tolerances. CFD modelling systems reach average FAR values ranging from 20 % (for 20 % tolerance) to 40 % (for 10 % tolerance) for the background station, and from 35 % (for 20 % tolerance) to 55 % (for 10 % tolerance) for the traffic station. Additionally, the CFD modelling systems (11 models) show a remarkable variability similar to the case of Gaussian ones (3 models).

The average values of BIAS (Fig. 15) are close to unity except for the traffic station for low tolerances in which the values are around 0.6 indicating a substantial underprediction. In general, BIAS grows as the tolerance increases. For the background station, SRA is slightly underpredicted for 10 % tolerance, and there is some overprediction for 20 % tolerance. For 15 % tolerance neither over nor underprediction is detected. For the traffic station, the values of BIAS are very close to one for 20 % tolerance.

For medium and high tolerances for both stations, the CFD modelling systems have an average BIAS close to 1 for the background station for 10 and 15 % tolerances and for the traffic one for 20 % tolerance. In the other cases, the CFD models seem to underpredict the samplers inside the SRAs but more strongly for the traffic station with 10 % tolerance. We can highlight that the AI models have the closest BIAS values to 1 for almost all tolerances and both stations, while the Lagrangian model has BIAS close to 1 for the background station with low tolerance and for traffic one with medium and high tolerances. The Lagrangian model underpredicts the SRA of traffic station for 10 % tolerance and overpredicts the background station SRA for 20 % tolerances. The Gaussian modelling systems tend to overpredict the SRAs for the background stations for all tolerances and for the traffic station with 15 and 20 % tolerances. In the case of SRA of the traffic station for 10 % tolerance, the Gaussian modelling systems underpredict to a similar extent as the CFD or Lagrangian modelling systems.

In general, the variability in the BIAS values is very similar for the CFD, Gaussian and AI modelling systems in the case of the background station SRA, while the largest variability for the traffic station SRA is observed in the CFD and Gaussian modelling systems.

In addition, the same statistics were computed for the different modified modelling concentration results as indicated in Section 2.3. The detailed results can be seen in Supplementary Material Section (Figs. SM.10–12). The results will be discussed briefly in Section 4.3.2.

It is also relevant to investigate the sensitivity of the predicted SRAs when using the same model but with different type of simulations or number of scenarios. This was done for two opposite cases. One of very large SRA, that is, the background station SRA for 20 % tolerance, and another one for a case of small SRA, such as the traffic station SRA for 10 % tolerance. Simulations from the SZE group were used. They consisted of a full-month unsteady OpenFOAM simulations and the application of a methodology for retrieving monthly NO_2 concentration data from different number of CFD simulated wind sector scenarios (4, 8, 16, and 32).

There are very few differences in the values of the accuracy, FAR and BIAS between the unsteady full-month simulation and the scenario-based estimates for the background station SRA. For the traffic station SRA, the unsteady simulation has slightly worse accuracy (45 %), higher FAR (57 %) and underpredicts more (BIAS equal to 0.50) than the

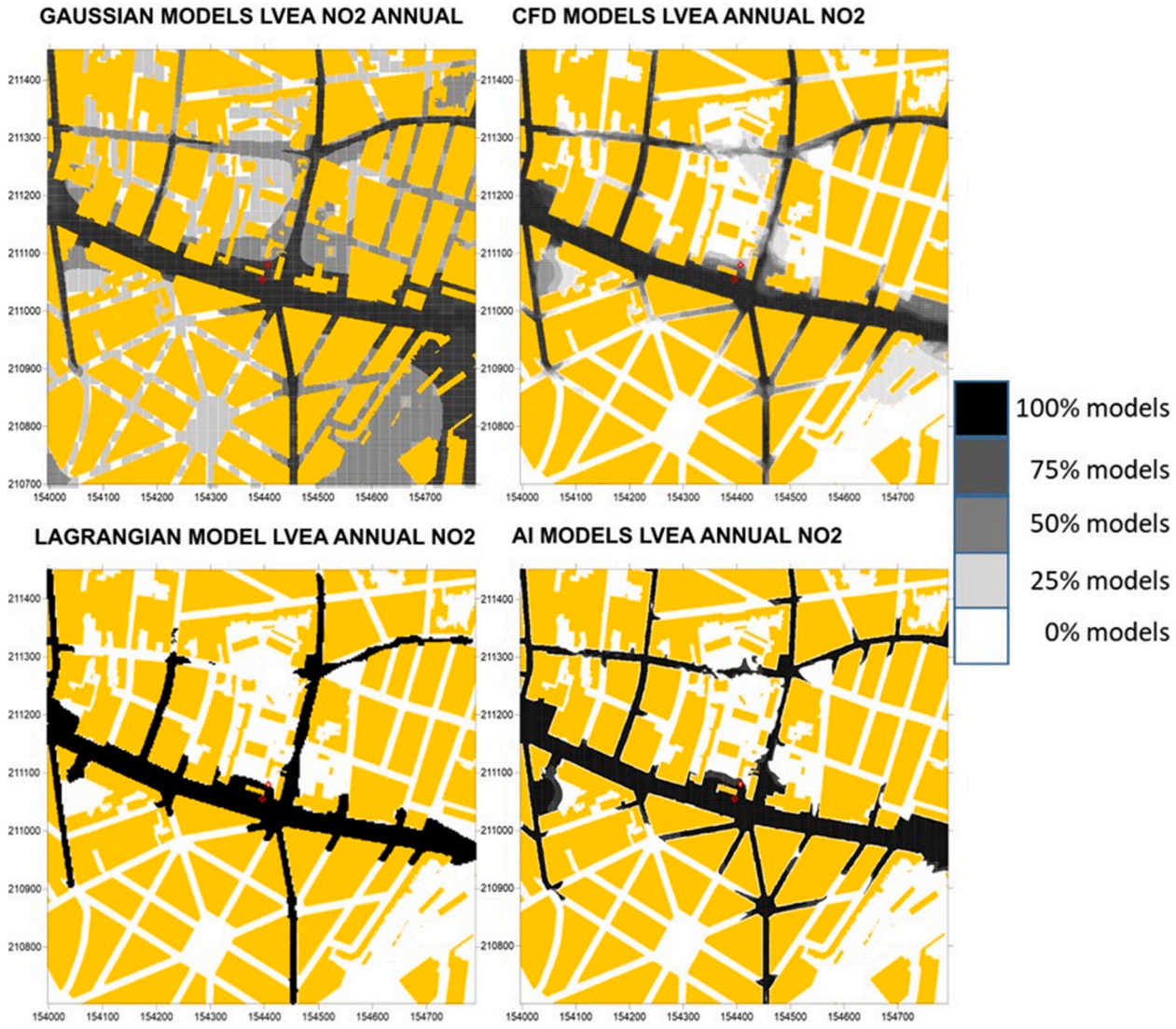


Fig. 11. Overlapping of the limit value exceedances areas computed by the different types of modelling systems (Gaussian models at upper left, CFD models at upper right, Lagrangian model at lower left and Artificial Intelligence at lower right). The darker the pixel, the more models estimate it is within the LVEA. The white color pixel indicates none model predicts it within the LVEA. In the case of the Lagrangian model the pixels shown in black are the LVEA computed by the model. Yellow means buildings.

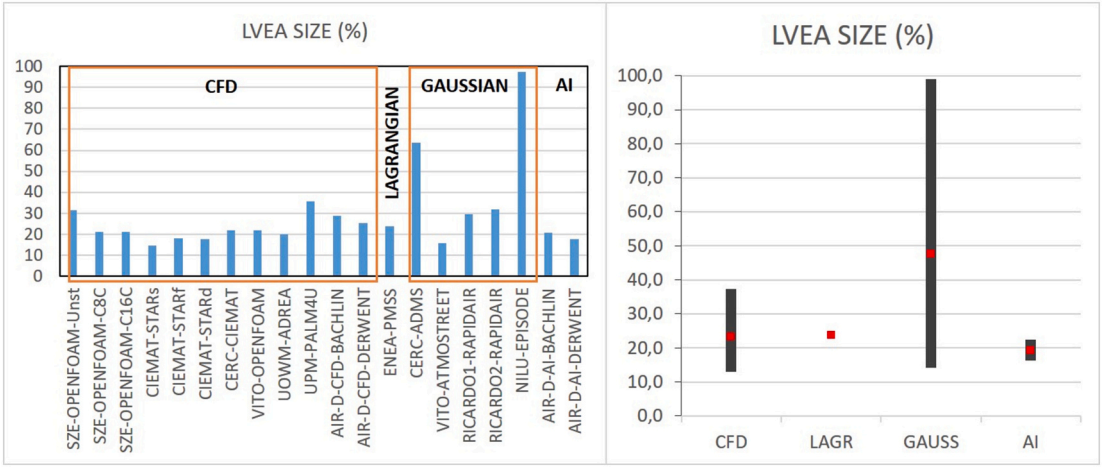


Fig. 12. Size (% of non-built area of the domain) of the limit value exceedances areas computed by every modelling system (left) and the mean and range of size for the different types of modelling systems (right). CFD models, GAUSS = Gaussian models, LAGR = Lagrangian model, and AI = Artificial Intelligence models.

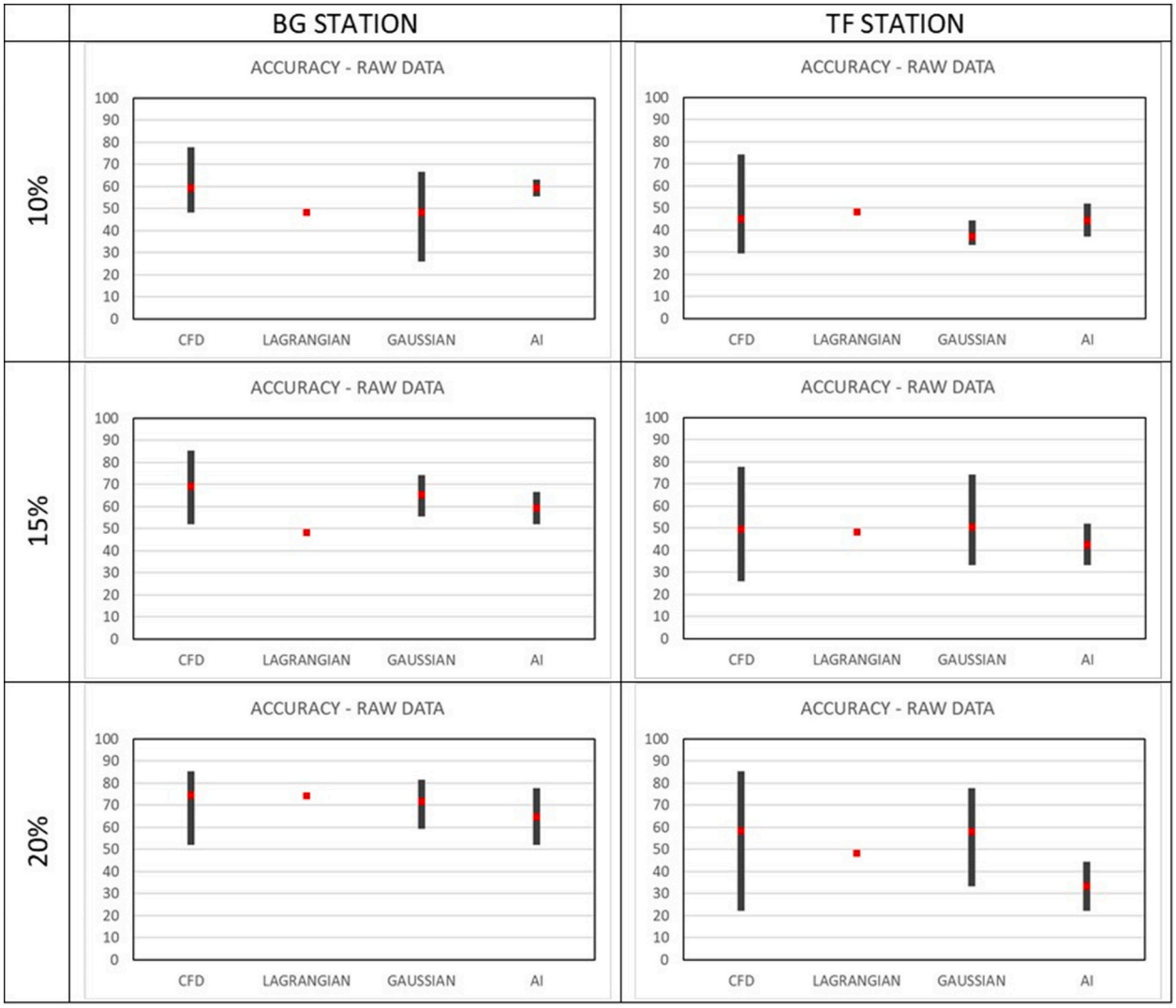


Fig. 13. Accuracy Index (maximum-minimum range in black and mean value in red) for predictions of the spatial representativeness area of the background (BG, left) and traffic (TF, right) stations corresponding to the raw estimates of modelling systems grouped by model types (CFD, Lagrangian, Gaussian and Artificial Intelligence (AI)). Three different tolerances were used (10 % (above), 15 % (middle) and 20 % (below)).

scenario-based estimates. In the group of the scenario-based estimates, there are no significant differences in the accuracy, FAR and BIAS when different number of scenarios is used. Very little differences are observed for the 32 scenarios case for the SRA of background station (Fig. 16).

3.5. Comparison of 2016 annual SRAs

The maps of the annual SRAs from the raw model data for both stations and for 10 % and 20 % tolerance are shown in Fig. 17. There are many common areas among the SRAs computed with results of the different modelling systems, but the SRAs of the background station are larger than those of the traffic one. In addition, as expected higher tolerance yields larger SRAs.

The maps of resulted SRAs for both stations and for 10 % and 20 % tolerances grouped by modelling system types are shown in Fig. 18, while the information about the SRAs' sizes can be seen in Figs. 19 and 20. These plots reveal some significant differences in shape and size of the SRAs computed by the different modelling systems. These results will be analyzed in detail in the Discussion section.

4. Discussion

4.1. Annual average NO_2 concentration maps (Figs. 4–7)

The maps of annual averaged NO_2 concentrations for the studied domain obtained from the different modelling systems are qualitatively very similar to those of the monthly averaged concentrations analyzed by Martín et al. (2024). Despite the resulted annual maps having many qualitative similarities estimating the maximum concentration in almost the same zones, there are notable differences in the magnitude of the predicted concentrations. The results from the CFD (Fig. 4), Lagrangian (Fig. 5), and AI (Fig. 6) modelling systems seem to simulate better the Street-Canyon effects as maximum concentration areas are shifted to a sidewalk. Gaussian modelling systems (Fig. 7) provide smoother concentration maps (much more notable for the NILU-EPISODE model, which does not include any street-canyon parametrization and uses a simplified meteorology) with weak gradients and/or maximum concentrations areas centered in the street axis. VITO-ATMOSTREET and CERC-ADMS predict maxima at the street crossings, while CFD modelling systems do not show these hotspots due to the explicit modelling of higher ventilation in crossing areas.

There are also significant differences in the magnitude of the maxima

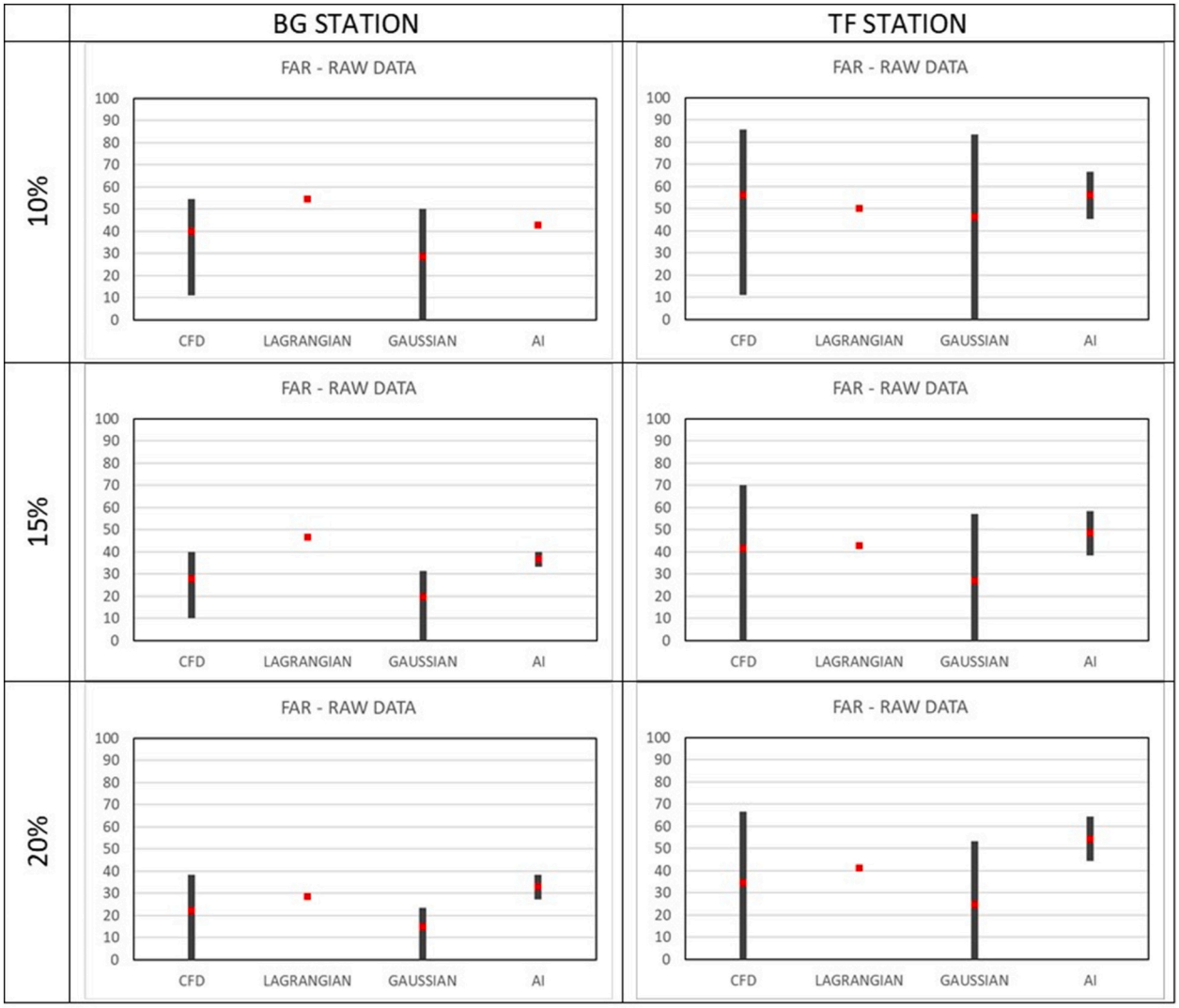


Fig. 14. False Alarm Rate (FAR) (maximum-minimum range in black and mean value in red) for predictions of the spatial representativeness area of the background (BG, left) and traffic (TF, right) stations corresponding to the raw estimates of the modelling systems grouped by model types (CFD, Lagrangian, Gaussian and Artificial Intelligence (AI)). Three different tolerances were used (10 % (above), 15 % (middle) and 20 % (below)).

in the results of the non-Gaussian modelling systems, and the location of some of them (Figs. 4–6). The higher ones are obtained by UOWM-ADREA, UPM-PALM4U, CERC-CIEMAT, AIR-D-CFD-DERWENT, AIR-D-CFD-BACHLIN and SZE-OpenFOAM unsteady simulation (all are CFD modelling systems). In contrast, ENEA-PMSS (Lagrangian) and CIEMAT-SIMPLE (CFD) predict a lower magnitude for the maxima than the average. These differences could be related to particular features or parameter configuration of the modelling systems, or the way in which the emissions data were processed prior to input into the modelling systems, or the numerical methodology for post-processing the scenario simulations to compute annual average concentrations (in the case of non-full-year simulations). In general, the Gaussian modelling systems (except CERC-ADMS) predict lower maxima than CFD modelling systems (Fig. 7). There are remarkable differences in the maxima concentration among the Gaussian models' estimates due to the different street-canyon parameterizations used by three of them (NILU-EPISODE does not include any street canyon parameterization). All of these may point to the importance of accounting for street-canyon effects in detail, which influences the formation of pollutant hotspots in the urban environment.

Martín et al. (2024) shows a detailed evaluation of the different modelling systems for computing monthly NO_2 concentrations, which are the same ones used in this study. The outcomes of that study contribute to a better understanding of the concentration maps generated for the current study. That study concluded that the models that account for complex urban geometries (i.e. CFD, Lagrangian, and AI models) appear to provide better estimates of the spatial distribution of one-month NO_2 average concentrations in the urban canopy. Approaches based on steady CFD-RANS (Reynolds Averaged Navier Stokes) model simulations of meteorological scenarios seem to provide good results with similar quality to those obtained with an unsteady one-month period CFD-RANS simulations. In contrast, Gaussian models were not able to provide detailed information, unless they include building data and street-canyon parameterizations as the case of CERC-ADMS, VITO-ATMOSTREET or RICARDO-RAPIDAIR.

4.2. LVEA

There are several key questions to answer considering the results of statistical analysis results:

4.2.1. How good are the models' results predicting the LVEA? (Fig. 8, SM.1)

Considering the high accuracy index values (70 % on average) and low false alarm rates (<25 % on average), the modelling systems are

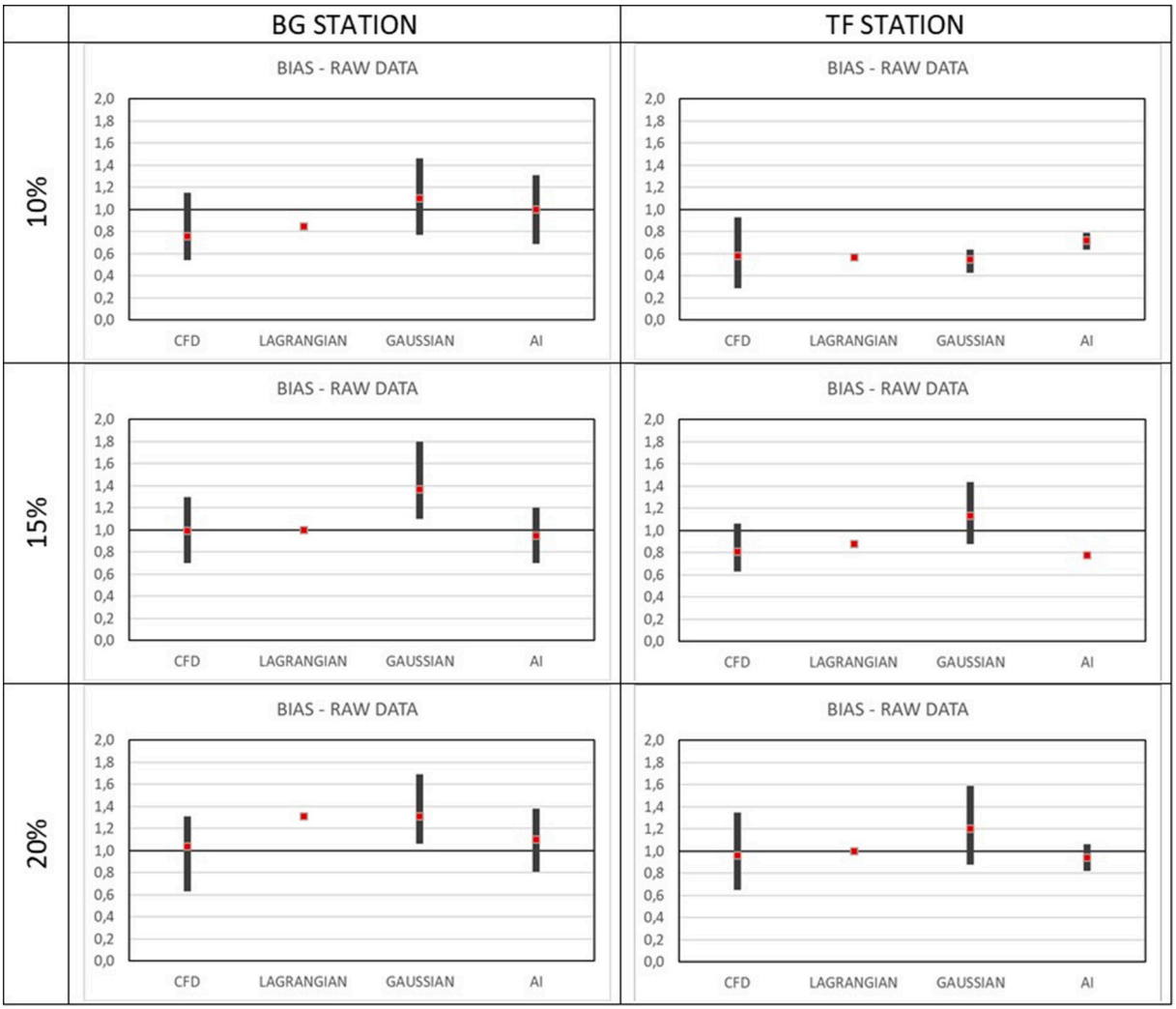


Fig. 15. Categorical Bias (BIAS) (maximum-minimum range in black and mean value in red) for predictions of the spatial representativeness area of the background (BG, left) and traffic (TF, right) stations corresponding to the raw estimates of the modelling systems grouped by model types (CFD, Lagrangian, Gaussian and Artificial Intelligence (AI)). Three different tolerances were used (10 % (above), 15 % (middle) and 20 % (below)).

generally good at predicting the limit value exceedance areas. However, they tend to underpredict the number of samplers exceeding the limit value, that is, they underpredict the spatial extent of LVEA. This is consistent with the fact that the modelling systems predict the no-exceedance areas (80 % of hits) better than the exceedances ones (<60 %) due to the underprediction of NO_2 concentration (Martín et al., 2024). We have to bear in mind that the emission data correspond to the road traffic of only the relevant streets (emissions from smaller streets were not available), and other sources such as domestic heating were not consider because of the very low contribution in spring. Nevertheless, these results could vary depending on the urban morphology and meteorological conditions. Very simplified modelling of stability conditions and background concentrations could also yield to this concentration underpredictions.

4.2.2. What type of modelling systems provides better predictions of the LVEA? (Figs. 8, 9 and SM.2–4)

CFD models seem to provide more consistent results because they have better scores for many statistics and are less sensitive to the different types of model data modification. However, there is a significant variability among the 12 CFD-based modelling systems, especially in FAR and BIAS (see Fig. 8), but such variability is of the same order as the Gaussian or AI modelling systems with much less cases.

With respect to the statistical results of the estimates of LVEA with full-month simulations and the scenario simulations, there are almost no differences for accuracy index and false alarm rates (FAR), but some slight improvement (less underprediction) is observed for the BIAS of the full-month simulation (Fig. 9). It is in agreement with the fact that the scenario-based methodologies underpredict more intensely the NO_2 concentrations at samplers' location (see Martín et al., 2024).

Other model types such as AI and Lagrangian (in this order) also achieve pretty good results (Fig. 8). The group of Gaussian models seem to obtain the lowest values of accuracy and categorical bias (BIAS) (strong underprediction), but in contrast they have the lowest false alarm rate which could be related with the strong underprediction of samplers exceeding the limit value and the NO_2 concentrations (see Martín et al., 2024). However, at this point, it is important to analyze the individual model results (see supplementary material figs. SM.3), because there is a strong variability. The Gaussian modelling systems with street-canyon parametrizations (VITO-ATMOSTREET and CERC-ADMS) parametrizations provide much better results than the simplest Gaussian setups (NILU-EPISODE), which do not have neither high spatial resolution meteorological input representing the flows in the streets nor a street-canyon parametrization to take into account the effect of buildings along the street on the pollution dispersion patterns. The results of the Gaussian modelling systems with street-canyon

Fig. 16. Accuracy Index, False Alarm Rate (FAR) and Categorical Bias (BIAS) (maximum-minimum range in black and mean value in red) for predictions of the spatial representativeness areas for the background (20 % tolerance, left) and traffic (10 % tolerance, right) station corresponding to the full-month unsteady OpenFOAM simulations and methodologies based on wind sector scenarios (4, 8, 16 and 32) OpenFOAM steady-state simulations.

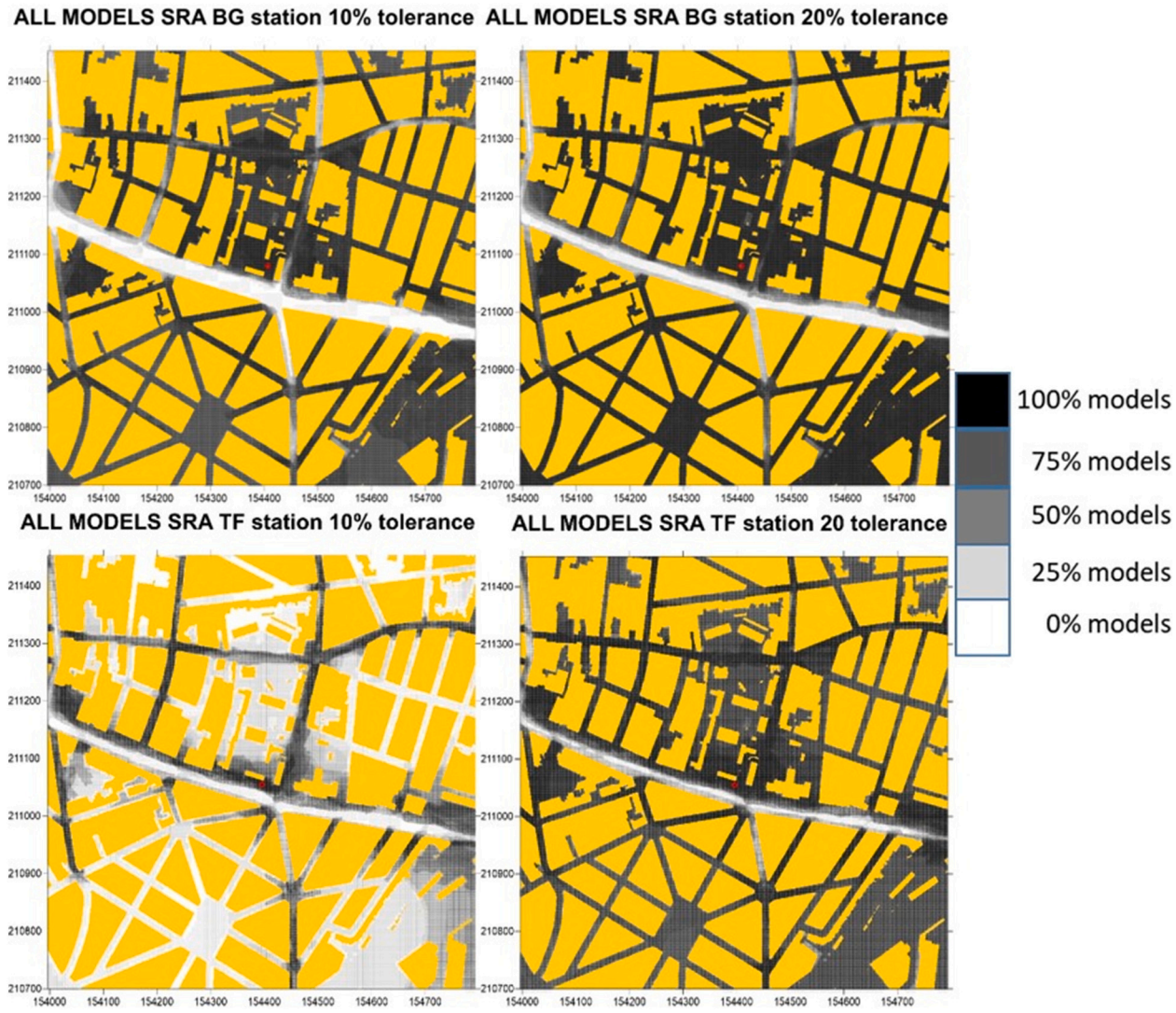


Fig. 17. Overlapping of the spatial representativeness areas computed by all the modelling systems for the background (above) and traffic station (below) and for 10 % (left) and 20 % (right) tolerances. The darker the pixel, the more modelling systems estimate it is within the SRA. The white color pixel indicates that none of the modelling systems predict it within the SRA. Yellow means buildings.

parametrizations are relatively close to the results from CFD, AI or Lagrangian approaches.

The predictions of concentrations obtained with the modelling systems can show some bias respect to the observed concentrations at the air quality stations (see Martín et al., 2024). Removing such concentration bias could improve the estimates of LVEA or SRA. Several model data modification methodologies have been tested in this study (see Section 2.3). All of them consist in modification to compensate the bias of the model concentration estimates using the concentration bias at the air quality stations or the concentration bias at the samplers' locations. It is described extensively in Supplementary Section and in Fig. SM.4. Concerning to the LVEA predictions, the main conclusion was that, it is not clear what modelling data modification could be better, but it seems to be clear that in this case, the model output modification based on correcting the bias with the station concentration data (*MSBBG* and *MSBTF*) does not provide significant improvements in the prediction of the LVEA respect to the raw model data. However, more studies in other urban areas could be needed to obtain clearer conclusions.

4.2.3. Are there significant differences in the annual LVEA when using different types of averaging methodologies in the CFD modelling systems? (Figs. 10–12)

Despite the annual LVEA estimated by the different modelling systems have many common zones such as the main streets and avenues; there are notably differences in the shape and size of the annual LVEA (Figs. 10–12). They can be due to the different types of models. The larger LVEAs are from the Gaussian modelling systems, and there is the strongest variability as well (from 15 % of LVEA size respect to the maximum area for the VITOATMOSTREET model to >95 % for the NILU-EPISODE model setup). No variability is observed for the Lagrangian and AI models because there are only 1 and 2 cases for such model types, respectively.

The CFD, Lagrangian and AI modelling systems provide quite similar sizes and shapes of the LVEA, but in the case of CFD modelling systems there is also significant variability. The larger LVEAs are for PALM4U and OpenFOAM unsteady full-year simulation from SZE (>30 % of the maximum area), while the smaller are the CIEMAT-STAR-CCM cases (14 %–18 %). The little differences among the CIEMAT-STAR-CCM results are noticeable. We keep in mind that these 3 different LVEA size estimates correspond to 3 different methods for retrieving the annual mean NO_2 concentration considering 16 wind sector scenarios (see

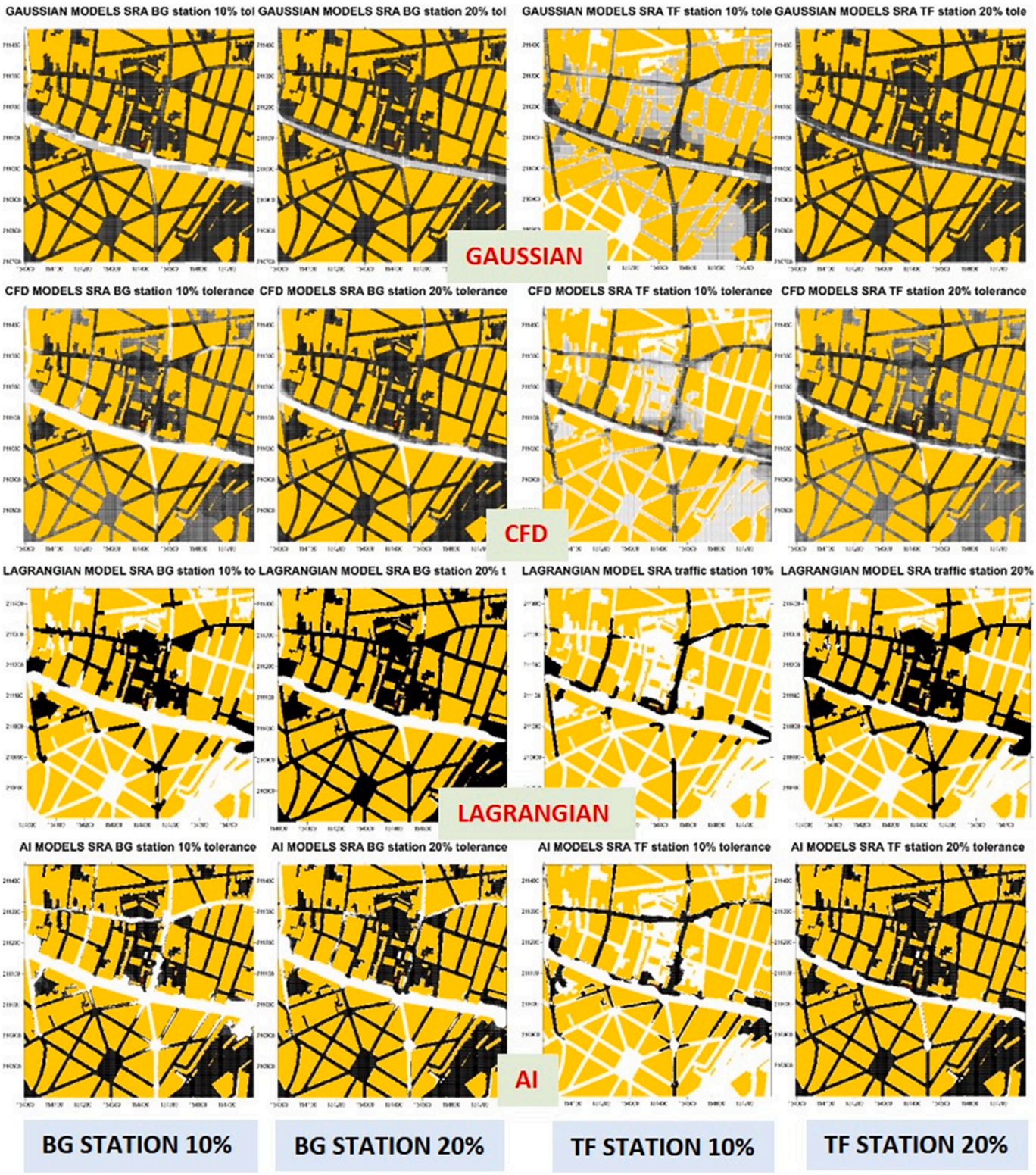


Fig. 18. Overlapping of the spatial representativeness areas computed by the different types of modelling systems for the background (left) and traffic station (right) and for 10 % and 20 % tolerances. The darker the pixel, the more modelling systems estimate it is within the SRA. The white color pixel indicates none model predicts it within the SRA. In the case of the Lagrangian model the pixels shown in black are the SRA computed by the model. Yellow means buildings.

Table 3.a and b of Section 2.2). The simplest CIEMAT-STAR-CCM (STARS) provides the smaller LVEA (14 %) due to this method also gives the lower NO_2 concentrations. The other two more complex (STARF and STARD) methods give a somewhat larger LVEA (18 % approximately). CERC also used the CFD simulations of CIEMAT with a more complex methodology for retrieving annual NO_2 concentrations. In this case, the estimated concentrations are higher and consequently, the annual LVEA is larger (with a size of 21 %).

It is also relevant the fact that the full-year unsteady simulation with OpenFOAM done by SZE (SZE—U) estimate a LVEA size (31 %)

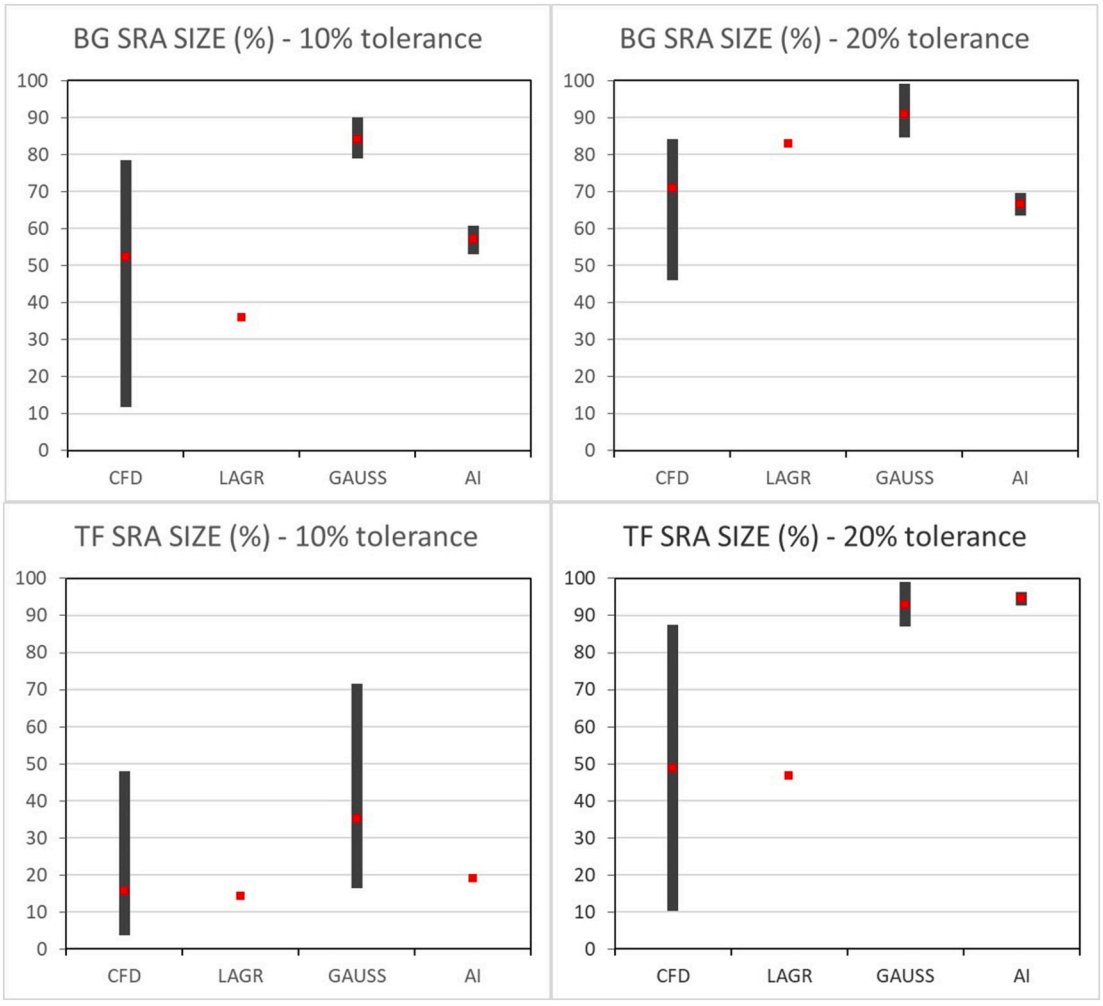


Fig. 19. Mean and range of size of the spatial representativeness areas for the different types of modelling systems for the background (above) and traffic station (below) and for 10 % (left) and 20 % (right) tolerances. GAUSS = Gaussian modelling systems, CFD modelling systems, LAGR = Lagrangian modelling systems, and AI = Artificial Intelligence modelling systems.

remarkably larger than the cases of using a wind direction scenario approach (SZE-C8C and SZE-C16C, with 8 and 16 wind direction scenarios, respectively) with LVEAs of 20 % of the maximum possible area. In addition, there is not significant differences when using 8 or 16 scenarios.

In spite of the comparison with the monthly data of samplers shows that the number of scenarios used does not seem not to be relevant for estimating the LVEA using methodologies based on CFD simulations of wind direction sector scenarios (see former section), the LVEA size resulted from the data of the full-year CFD simulation could be significantly larger for the full year simulation. However, surely further studies are needed using other modelling systems and urban domains.

4.3. SRAs

4.3.1. How good are the modelling systems predicting the SRAs? (Figs. 13–15, SM.5–6))

SRAs for low tolerances are more difficult to predict than SRAs for larger tolerances. It is supported by the clear improvement of the accuracy index, false alarm rates (FAR) and categorical bias (BIAS) from 10 % to 20 % tolerances especially in the case of the traffic station SRA. In general, the SRA of background station is better predicted than for the traffic station one (lower accuracy, higher false alarm rate and more underprediction except for 20 % tolerances). On average of all the modelling systems, the SRA for high tolerances is rather good for the background station (accuracy index >65 %, false alarm rates <35 % and bias between 1 and 1.3). In the case of the traffic station for high tolerances, the modelling systems did not predict the SRA as well as the SRA of the background station. The improvement predicting SRA is higher when passing from 10 % tolerance to 15 % tolerance than from 15 % to 20 % for the background station. For the traffic station, the improvement rate with increasing tolerance is more regular. Additionally, there is an important variability in the results of the different metrics, which shows an important influence of the modelling types.

As in the case of the LVEA, the modelling systems for low tolerances predict better the samplers outside of the SRAs than those being inside. It is more notable for the traffic station with a smaller SRA. In contrast, for higher tolerances, the modelling systems predicts better the samplers inside of SRA than those being outside (see Supplementary material, SM.5 and 6). It is much more related to the size of the SRA, which is larger when tolerances are higher.

4.3.2. What type of modelling systems provide better predictions of SRAs? (Figs. 13–16, SM.7–12)

It seems the CFD modelling systems provide more consistent results with better scores for many statistics for the SRAs of both stations for most of the tolerances in spite of some variability among the different CFD-based modelling systems (especially in accuracy and FAR for the SRA of the traffic station) (Figs. 13–15). Furthermore, their predictions appear to be little sensitive to the different analyzed modelling data

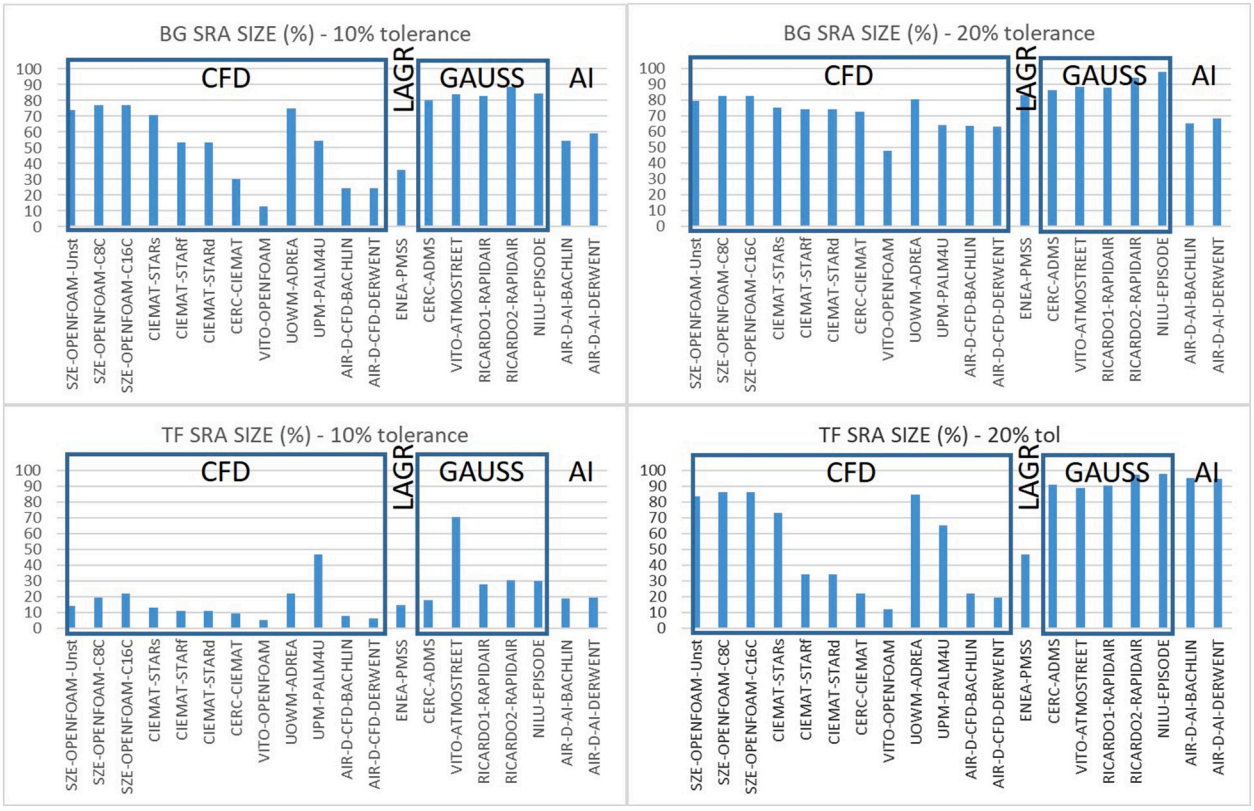


Fig. 20. Size (% of non-built area of the domain) of the spatial representativeness areas computed by every modelling system for the background (above) and traffic (below) stations and for 10 % (left) and 20 % (right) tolerances. GAUSS = Gaussian modelling systems, CFD modelling systems, LAGR = Lagrangian modelling systems, and AI = Artificial Intelligence modelling systems.

modifications except for AX+B case (see Figs. SM.10–12).

It is important to highlight that scenario-based CFD simulation methodologies (SZE-CnnC) provides results as good as the unsteady full month CFD simulations with OpenFOAM (SZE—U) in many cases especially for higher tolerances (larger SRAs). For smaller tolerances (smaller SRAs), the SZE-CnnC modelling system (based on OpenFOAM simulations for *nn* wind sector scenarios) performs even better than the unsteady full month CFD simulations with OpenFOAM (SZE—U). (Fig. 16).

Other modelling system types have also generally quite good results but with some shortcomings. For example, the Lagrangian model overpredicts the SRA of the background station for 20 % tolerance; the accuracy index is relatively low for the background station for 10 and 15 % tolerances and for the traffic station for 20 % tolerances. The false alarm rates are the highest for the background station for 10 and 15 % tolerances. In the case of the AI modelling systems, the accuracy index is significantly lower and the false alarm rate higher for the traffic station and 20 % tolerance, (Figs. 13–15).

The Gaussian modelling systems provide its best results for the SRA of the background station. The FAR values are generally the lowest of all the model types. However, the accuracy for low tolerances is relatively low. For 10 % and 15 % tolerances, the Gaussian modelling systems overpredict clearly the SRA for both stations. It could be much related to the fact that the Gaussian model does not simulate well enough the concentration gradients inside the streets (Martín et al., 2024). However, there is notable variability in the results depending on the modelling systems. The Gaussian modelling systems with street-canyon parametrizations provides better results than the simplest Gaussian modelling systems (NILU-EPISODE) for the SRA of background station being closer to the results of other model as CFD. However, the improvement is no evident for the SRA of traffic station (see Supplementary material, Figs. SM.7–9)

With respect to what modelling output data modification (for correcting the bias of the concentration estimates) could provide better estimates of SRA, a detailed discussion is shown in the Supplementary Material Section based on the results depicted in Figs. SM.10–12. The main conclusion, which is very similar to that for LVEA, is that it is not possible to determine what model data modification could be the most suitable because there is not a clear tendency to improve the raw data. It seems to slightly depend on the modelling approach but it could be very dependent on the studied case. Perhaps, it could be related to the ability of the modelling systems to simulate the concentration gradients. Nevertheless, surely further studies with other urban configurations should be needed to clarify this question.

4.3.3. Are there significant differences in the annual SRAs when using different types of averaging methodologies in the CFD modelling systems? (Figs. 17–20)

Analyzing the maps of the annual SRAs for both stations and for 10 % and 20 % tolerance (Fig. 17), it is noteworthy that the SRAs estimated by the different modelling systems share many areas in spite there are remarkable differences. Additionally, the SRAs are larger for the background station than for the traffic one, and for the 20 % tolerance, the SRAs are larger than for the 10 %. This increase in SRA size from 10 % to 20 % tolerance is more evident for the traffic station.

It is remarkable the fact that the most part of the main avenue (where high NO_2 concentration is estimated) is out of any of the SRAs from both stations and both tolerance (except for the SRA estimated by the simplest Gaussian model, NILU-EPISODE). It is an example of is very difficult to locate an air quality station to be representative of the air quality in all parts of avenues or streets with high traffic intensity and pollutant emissions. It is a big challenge for air quality network design.

Analyzing the results grouped by modelling system types (Figs. 18–20), we found that the SRAs for both tolerances are clearly

larger for the Gaussian modelling systems, with NILU-EPISODE yielding the largest SRAs (>95 % for the background station for both tolerances and for the traffic station for 20 % tolerance). However, there is a high variability of SRA computed with Gaussian modelling systems for the traffic station with 10 % tolerance, but not for the background station with both tolerances. It is also evident the remarkable variability of SRAs for both tolerances computed with CFD modelling systems, especially for the traffic station.

It is noticeable that there are little differences among the SRA sizes computed with the full year simulation with the OpenFOAM model (SZE—U) and with wind sector scenario OpenFOAM simulations (SZE-C8C and SZE-C16C), but it is slightly shorter for the SZE-U being more notable for the traffic station with 10 % tolerance (when the SRA is smaller). The SRA sizes obtained with OpenFOAM simulations with 8 and 16 wind sector scenarios (SZE-C8C and SZE-C16C, respectively) are very similar showing that the number of scenarios do not seem to be relevant.

On the other hand, analyzing the SRAs estimated by 4 different methods for retrieving the annual mean NO_2 concentration considering 16 wind sector scenarios simulated with the STAR-CCM CFD model (STARs, STARF, STARd and CERC-CIEMAT), some significant differences are detected. The simplest method (STARs) provides larger SRAs than the other 3 methods, while the most complex (CERC-CIEMAT) yields smaller SRAs for both stations and tolerances. For the background station and 20 % tolerance, all the methodologies provides similar sizes of SRAs (>70 %). It means the SRA sizes depends on the methodology for retrieving the annual mean NO_2 concentration.

5. Conclusions

The main conclusions obtained of this study are:

- The analysis of annual concentrations maps computed the different modelling systems types (CFD, Lagrangian, Gaussian and AI) used in this intercomparison exercise point to the importance of accounting for street-canyon effects in detail, which influences the formation of pollutant hotspots in the urban environment with very strong concentration gradients. This is in agreement with the results of Martín et al. (2024).
- Most of the analyzed microscale modelling systems generally are good predicting the LVEA and SRA, but there are significant differences among the different type of modelling systems. CFD modelling systems (despite some variability) seem to provide somewhat better predictions of LVEA and SRA, but Lagrangian, AI and Gaussian models with built-in street-canyon parametrizations provide also quite good results. However, the modelling systems seem to predict better the SRA for the background station than for the traffic one. Small SRAs are much more difficult to predict correctly, which is the case of traffic stations with 10 % tolerance.
- The annual LVEA and SRA predicted by the different modelling systems is quite similar with several common streets but there are some relevant variability among the different modelling system types and among the different methodologies for retrieving the annual mean NO_2 concentration using the same model system. The LVEA and SRA sizes are clearly larger for the Gaussian modelling systems but much larger for the Gaussian modelling system with no street-canyon parametrizations.
- A stricking result was that most of the SRAs' estimates do not cover most of the main avenue where high NO_2 concentrations are predicted, which means that the measurement of the closest stations is not representative of the pollution in the avenue.
- Scenario-based CFD simulation methodologies provide statistical results as good as the unsteady full month CFD simulations predicting the LVEA and the SRA for higher tolerances (larger SRA). However, for smaller tolerances (smaller SRA), the scenario-based methodology improves slightly the results of the full-month CFD simulations. The number of scenarios seems not to be very relevant, in general.
- In addition, there are also relatively notable differences in the sizes of the predicted annual LVEAs and SRAs. The SRAs are larger for the estimates using the scenario-based CFD simulation approach (more notable for the traffic station with 10 % tolerance) than for the estimates of the full-year simulation. In contrast, the sizes of the annual LVEA are larger for the estimates using the unsteady full-year simulations.
- It is not possible to determine what model data modification could be suitable because there is not a clear tendency to improve the raw data. It seems to slightly depend on the modelling approach, but it could be very dependent on the studied case. Surely further studies with other urban configurations would be needed.

An important remark is that air pollution hotspots in a city are specially located in streets flanked continuously by buildings that lead to micro-meteorological effects dominated by the street geometry, leading to canyon effects (reduced ventilation) and formation of air vortexes. In cities with long and narrow streets, it is very important to include the effect of buildings on dispersion, be it parameterized (like street canyon parameterizations in Gaussian modelling or explicitly as in CFD models, for example). That is the case of our domain in Antwerp (Martín et al., 2024). In other urban domains with wider and ventilated streets this may be less crucial. This explains why the NILU EPISODE model without the street-canyon parametrization does not give so good results in this case. However, it is used a base reference for the more complex modelling systems. Nevertheless, NILU-EPISODE is widely used with good results in modelling applications in Norwegian cities, which have wider and ventilated streets (Weydahl et al., 2023, 2024; Høiskar et al., 2023). In the Antwerp application, however, the present results show as street-canyon parametrizations should have been used in Gaussian models.

The novelty of the results shown in this paper is to provide insight on modelling systems prediction accuracy for the areas of limit value (air quality standards) exceedance and spatial representativeness of air quality stations in a complex urban area. However, they could be highly dependent on the urban configuration of the studied domain and it is possible that different results could be expected when other urban areas with different urban morphology, meteorology or emissions are provided. Haeger-Eugensson et al. (2021) made a study comparing the performance of different type of models (a CFD model, a parametric model and two Gaussian models) for 4 different urban environments in some Sweden cities. The purpose was to investigate the effect of building layout on air quality to give recommendations to authorities and consultants when setting the models. The results shown the very notable street-canyon contribution to the NO_2 concentrations in narrow streets and the CFD models provided the best results in all urban configuration with significant difference with the other model types for complex urban configuration. In Martín et al. (2024), further discussion about the performance for predicting NO_2 concentrations of the different type of models used in this intercomparison exercise is shown. Hence, we consider that more investigations using other sets of modelling systems and urban domains is needed.

The results presented in this study along with those of Martín et al. (2024) are contributing to the development of a guidance or best practices document for microscale modelling that considers the urban layout complexity and type of emissions. The objective of such document is to indicate the type of microscale modelling systems applied to air quality assessment, compute exceedance areas, and evaluate the spatial representativeness of urban monitoring stations that are fit for purpose.

CRediT authorship contribution statement

F. Martín: Writing – review & editing, Writing – original draft,

Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **V. Rodrigues:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **J.L. Santiago:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **J. Sousa:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **J. Stocker:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **S. Janssen:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **R. Jackson:** Writing – review & editing, Software, Methodology, Formal analysis. **F. Russo:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **M.G. Villani:** Software, Methodology, Investigation, Formal analysis. **G. Tinarelli:** Software, Methodology, Investigation, Formal analysis. **D. Barbero:** Software, Methodology, Investigation, Formal analysis. **R. San José:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **J.L. Pérez-Camanyo:** Software, Methodology, Investigation, Formal analysis. **G. Sousa-Santos:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **L. Tarrason:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **J. Bartzis:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **I. Sakellaris:** Software, Methodology, Investigation, Formal analysis. **Z. Horváth:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis. **L. Környei:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **X. Jurado:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **N. Reiminger:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis. **N. Masey:** Software, Resources, Methodology, Investigation, Formal analysis. **S. Hamilton:** Software, Resources, Methodology, Investigation, Formal analysis. **E. Rivas:** Software, Methodology, Investigation, Formal analysis. **C. Cuvelier:** Visualization, Software, Methodology, Investigation, Formal analysis. **P. Thunis:** Supervision, Project administration, Investigation, Funding acquisition, Formal analysis.

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European High Performance Computing Joint Undertaking (JU) and Poland, Germany, Spain, Hungary, France. Neither the European Union nor the granting authority can be held responsible for them.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Thanks are due to FCT/MCTES for the financial support to CESAM (UIDP/50017/2020 + UIDB/50017/2020 + LA/P/0094/2020), through national funds, and the co-funding by the FEDER, within the PT2020 Partnership Agreement and Compete 2020.

Funded by the European Union. Part of this work has received funding from the European High Performance Computing Joint Undertaking (JU) and Poland, Germany, Spain, Hungary, France under grant agreement number: 101093457.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2025.179824.

Data availability

Data will be made available on request.

References

Bai, H., Yan, R., Gao, W., Wei, J., Seong, M., 2022. Spatial representativeness of PM2.5 monitoring stations and its implication for health assessment. Air Qual. Atmos. Health 15, 1571–1581. https://doi.org/10.1007/s11869-022-01202-2.

Bai, H., Gao, W., Seong, M., Yan, R., Wei, J., Liu, C., 2023. Evaluating and optimizing PM2.5 stations in Yangtze River delta from a spatial representativeness perspective. Appl. Geogr. 154, 102949. https://doi.org/10.1016/j.apgeog.2023.102949.

Barbero, D., Tinarelli, G., Silibello, C., Nanni, A., Gariazzo, C., Stafoggia, M., Viegi, G., 2021. A microscale hybrid modelling system to assess the air quality over a large portion of a large European city. Atmos. Environ. 264, 118656.

Bartzis, J., Andronopoulos, S., Efthimiou, G., 2020a. Simplified approaches in quantifying exposure statistical behaviour due to airborne hazardous releases of short time duration. In: Proceedings of Abstracts 12th International Conference on Air Quality Science and Application. Thessaloniki, Greece.

Bartzis, J.G., Kalimeri, K.K., Sakellaris, I.A., 2020b. Environmental data treatment to support exposure studies: the statistical behavior for NO_2, O_3, PM10 and PM2.5 air concentrations in Europe. Environ. Res. 181, 108864.

Bartzis, J.G., Efthimiou, G.C., Andronopoulos, S., 2015. Modelling short term individual exposure from airborne hazardous releases in urban environments. J. Hazard. Mater. 300, 182–188.

Bartzis, J.G., Efthimiou, G.C., Andronopoulos, S., 2021. Modelling exposure from airborne hazardous short-duration releases in urban environments. Atmosphere 12 (2).

Bartzis, J.G., Sakellaris, I.A., Efthimiou, G., 2022. On exposure uncertainty quantification from accidental airborne point releases. Journal of Hazardous Materials Advances 6, 100080.

Beauchamp, M., Malherbe, L., de Fouquet, C., L'etinois, L., 2018. A necessary distinction between spatial representativeness of an air quality monitoring station and the delimitation of exceedance areas. Environ. Monit. Assess. 190, 441. https://doi.org/10.1007/s10661-018-6788-y.

Belda, M., Resler, J., Geletic, J., Krc, P., Maronga, B., Sühring, M., Kurppa, M., Kanani-Sühring, F., Fuka, V., Eben, K., Benešová, N., Auvinen, M., 2021. Sensitivity analysis of the PALM model system 6.0 in the urban environment. Geosci. Model Dev. 14 (7), 4443–4464.

Biggart, M., Stocker, J., Doherty, R.M., Wild, O., Hollaway, M., Carruthers, D., Li, J., Zhang, Q., Wu, R., Kotthaus, S., Grimmond, S., Squires, F.A., Lee, J., Shi, Z., 2020. Street-scale air quality modelling for Beijing during a winter 2016 measurement campaign. Atmos. Chem. Phys. 20 (5), 2755–2780.

Carruthers, D.J., Holroyd, R.J., Hunt, J.C.R., Weng, W.S., Robins, A.G., Apsley, D.D., Thompson, D.J., Smith, F.B., 1994. UK-ADMS: a new approach to modelling dispersion in the earth's atmospheric boundary layer. J. Wind Eng. Ind. Aerodyn. 52, 139–153.

Criado, A., Armengol, J.M., Petetin, H., Rodriguez-Rey, D., Benavides, J., Guevara, M., Pérez García-Pando, C., Soret, A., Jorba, O., 2023. Data fusion uncertainty-enabled methods to map street-scale hourly NO_2 in Barcelona: a case study with CALIOPE-urban v1.0. Geosci. Model Dev. 16, 2193–2213. https://doi.org/10.5194/gmd-16-2193-2023.

Degraeuwe, B., Pisoni, E., Christidis, P., Christodoulou, A., Thunis, P., 2021. Sherpa-city: a web application to assess the impact of traffic measures on NO_2 pollution in cities. Environ. Model. Softw. 135, 104904.

Duyzer, J., van den Hout, D., Zandveld, P., van Ratingen, S., 2015. Representativeness of air quality monitoring networks. Atmos. Environ. 104, 88–101. https://doi.org/10.1016/j.atmosenv.2014.12.067.

EEA, 2025. Exceedance of air quality standards in Europe. European Environment Agency. https://www.eea.europa.eu/en/analysis/indicators/exceedance-of-air-quality-standards?activeAccordion.

EU, 2024. Proposal for a revision of the ambient air quality directives. https://environment.ec.europa.eu/publications/revision-eu-ambient-air-quality-legislation_en.

Haeger-Eugensson, M., et al., 2021. Air quality modeling in dense urban areas at ground level—CFD, OSM or Gauss? In: Mensink, C., Matthias, V. (Eds.), Air Pollution Modeling and Its Application XXVII. ITM 2019. Springer Proceedings in Complexity. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-63760-9_37.

Hamer, P.D., Walker, S.E., Sousa-Santos, G., Vogt, M., Vo-Thanh, D., Lopez-Aparicio, S., Schneider, P., Ramacher, M.O.P., Karl, M., 2020. The urban dispersion model EPISODE v10.0 – Part 1: An Eulerian and sub-grid-scale air quality model and its

application in Nordic winter conditions. Geosci. Model Dev. 13, 4323–4353. https://doi.org/10.5194/gmd-13-4323-2020, 2020.
Høiskar, B.A.K., Walker, S.E., Weydahl, T., Markelj, M., Andersen, A., Lopez-Aparicio, S., Grythe, H., et al., 2023. ISBN: 978-82-425-3107-0. https://hdl.handle.net/11250/3062978.
Hood, C., Stocker, Jenny, Seaton, Martin, Johnson, Kate, O'Neill, James, Thorne, Lewis, Carruthers, D., 2021. Comprehensive evaluation of an advanced street canyon air pollution model. J. Air Waste Manage. Assoc. 71 (2), 247–267.
Hooyberghs, H., De Craemer, S., Lefebvre, W., Vranckx, S., Maiheu, B., Trimpeneers, E., Vanpoucke, C., Janssen, S., Meysman, F.J.R., Fierens, F., 2022. Validation and optimization of the atmo-street air quality model chain by means of a large-scale citizen-science dataset. Atmos. Environ. 272, 118946.
Hooyberghs, H., Tarrason, L., Janssen, S., Soares, J., 2020. Assessing the spatial representativeness (SR) of air quality sampling points – sensitivity and feasibility tests for a tiered approach – final report. In: Service Request 5 under Framework Contract ENV.C.3/FRA/2017/0012. Specific Contract: 07.0203/2018/793545/SFRA/ENV.C.3. Report for European Commission - DG Environment. Ares, 4920320 (ED 11492 | Task 1 Issue number 5 | Date 21/12/20. Ricardo Confidential, 2020).
Horváth, Z., Liszkai, B., Istenes, G., Zsebők, P., Szintai, B., Rácz, É.V.P., Környei, L., Harmati, I., 2016. Integrated urban air pollution dispersion modelling framework and application in air quality prediction of the city of Győr. In proceedings of the 17th international conference on harmonisation within atmospheric dispersion modelling for regulatory purposes.
Janssen, S., Dumont, G., Fierens, F., Mensink, C., 2008. Spatial interpolation of air pollution measurements using corine land cover data. Atmos. Environ. 42 (20), 4884–4903.
Janssen, S., Dumont, G., Fierens, F., Deutsch, F., Maiheu, B., Celis, D., Trimpeneers, E., Mensink, C., 2012. Land use to characterize spatial representativeness of air quality monitoring stations and its relevance for model validation. Atmos. Environ. 59, 492–500. https://doi.org/10.1016/j.atmosenv.2012.05.028.
Janssen, S., Tarrasón, L., Ross-Jones, M., 2023. FAIRMODE WG8 – guidance document on the estimation of spatial representativeness and of exceedance situation indicators. FAIRMODE (forum for air quality modelling in Europe). Working group 8 on monitoring design, spatial representativeness and associated exceedance situation indicators. https://fairmode.jrc.ec.europa.eu/document/fairmode/WG8/WG8_Guidance_Document_VS3.pdf.
Janssen, S., Ross-Jones, M., Monteiro, A., Pirovano, G., Denby, B.R., Strużewska, J., Jursins, J., Green, J., Brookes, D., 2024. In: EUROPEAN COMMISSION Directorate-General for Environment. Directorate C — Zero Pollution. Unit C3— Clean Air & Urban policy (Ed.), Air quality modelling for air quality policy. Technical support document on the use of modelling for various application domains under the ambient air quality directive.
Jiménez- Guerrero, P., Jorba, O., Baldasano, J.M., Gassó, S., 2008. The use of a modelling system as a tool for air quality management: annual high-resolution simulations and evaluation. Sci. Total Environ. (2–3), 323–340. https://doi.org/10.1016/j.scitotenv.2007.10.025.
Johansson, L., Karppinen, A., Kurppa, M., Kousa, A., Niemi, J.V., Kukkonen, J., 2022. An operational urban air quality model ENFUSER, based on dispersion modelling and data assimilation. Environ. Model. Softw. 156, 105460.
Jurado, X., Reiminger, N., Vazquez, J., Wemmert, C., 2021. On the minimal wind directions required to assess mean annual air pollution concentration based on CFD results. Sustain. Cities Soc. 71, 102920.
Jurado, X., Reiminger, N., Benmoussa, M., Vazquez, J., Wemmert, C., 2022. Deep learning methods evaluation to predict air quality based on computational fluid dynamics. Expert Syst. Appl. 203, 117294.
Jurado, X., Reiminger, N., Maurer, L., Vazquez, J., Wemmert, C., 2023. Assessment of a deep learning model for monitoring atmospheric pollution: case study in Antwerp, Belgium. Sustain. Cities Soc. 99, 104951.
Környei, L., Horváth, Z., Ruopp, A., Kovacs, A., Liszkai, B., 2021. Multi-scale modelling of urban air pollution with coupled weather forecast and traffic simulation on HPC architecture. In: Proceedings of the International Conference on High Performance Computing in Asia-Pacific Region Companion (HPC Asia 2021).
Kracht, O., Santiago, J., Martin, F., Piersanti, A., Cremona, G., Righini, G., Vitali, L., Delaney, K., Basu, B., Ghosh, B., Spangl, W., Brendle, C., Latikka, J., Kousa, A., Pärjälä, E., Meretoja, M., Malherbe, L., Letinois, L., Beauchamp, M., Lenartz, F., Hutsemekers, V., Nguyen, L., Hoogerbrugge, R., Eneroth, K., Silvergren, S., Hooyberghs, H., Viaene, P., Maiheu, B., Janssen, S., Roet, D., Gerboles, M., 2017. Spatial Representativeness of Air Quality Monitoring Sites: Outcomes of the FAIRMODE/AQUILA Intercomparison Exercise, EUR 28987 EN. Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/60611. ISBN 978-92-79-77218-4. JRC108791.
Lefebvre, W., Van Poppel, M., Maiheu, B., Janssen, S., Dons, E., 2013. Evaluation of the RIO-IFDM-street canyon model chain. Atmos. Environ. 77, 325–337. https://doi.org/10.1016/j.atmosenv.2013.05.026.
Li, H.Z., Gu, P., Ye, Q., Zimmerman, N., Robinson, E.S., Subramanian, R., Apte, J.S., Robinson, A.L., Presto, A.A., 2019. Spatially dense air pollutant sampling: implications of spatial variability on the representativeness of stationary air pollutant monitors. Atmospheric Environment: X 2, 100012. https://doi.org/10.1016/j.aeaoa.2019.100012.
Luo, R., Zhang, M., Ma, G., 2023. Regional representativeness analysis of ground-monitoring PM2.5 concentration based on satellite remote sensing imagery and machine learning techniques. Remote Sens. 15. https://doi.org/10.3390/rs15123040.
Maronga, B., Gross, G., Raasch, S., Banzhaf, S., Forkel, R., Heldens, W., Kanani-Sühring, F., Matzarakis, A., Mauder, M., Pavlik, D., Pfafferott, J., Schubert, S., Seckmeyer, G., Sieker, H., Winderlich, K., 2019. Development of a new urban climate model based on the model palm – project overview, planned work, and first achievements. Meteorol. Z. 28 (2), 105–119.
Martin, F., Fileni, L., Palomino, I., Vivanco, M.G., Garrido, J.L., 2014. Analysis of the spatial representativeness of rural background monitoring stations in Spain. Atmos. Pollut. Res. 5, 779–788. https://doi.org/10.5094/APR.2014.087.
Martín, F., Janssen, S., Rodrigues, V., Sousa, J., Santiago, J., Rivas, E., Stocker, J., Jackson, R., Russo, F., Villani, M., Tinarelli, G., Barbero, D., José, R.S., Pérez-Camanyo, J., Santos, G.S., Bartzis, J., Sakellaris, I., Horváth, Z., Környei, L., Liszkai, B., Kovács, Jurado, X., Reiminger, N., Thunis, P., Cuvelier, C., 2024. Using dispersion models at microscale to assess long-term air pollution in urban hot spots: a FAIRMODE joint intercomparison exercise for a case study in Antwerp. Sci. Total Environ. 925, 171761. https://doi.org/10.1016/j.scitotenv.2024.171761.
Masey, N., Hamilton, S., Beverland, I.J., 2018. Development and evaluation of the RapidAir® dispersion model, including the use of geospatial surrogates to represent street canyon effects. Environ. Model Softw. 108, 253–263.
Mishra, V.K., Kumar, P., Van Poppel, M., Bleux, N., Frijns, E., Reggente, M., Berghmans, P., Int Panis, L., Samson, R., 2012. Wintertime spatio-temporal variation of ultrafine particles in a Belgian city. Sci. Total Environ. 431, 307–313. https://doi.org/10.1016/j.scitotenv.2012.05.054.
Monteiro, A., Miranda, A.I., Borrego, C., Vautard, R., 2007. Air quality assessment for Portugal. Sci. Total Environ. 373 (1), 22–31. https://doi.org/10.1016/j.scitotenv.2006.10.014.
Oldrini, O., Armand, P., Duchenne, C., Olry, C., Moussafir, J., Tinarelli, G., 2017. Description and preliminary validation of the PMSS fast response parallel atmospheric flow and dispersion solver in complex built-up areas. Environ. Fluid Mech. 17, 997–1014.
Owen, B., Edmunds, H.A., Carruthers, D.J., Singles, R.J., 2000. Prediction of total oxides of nitrogen and nitrogen dioxide concentrations in a large urban area using a new generation urban scale dispersion model with integral chemistry model. Atmos. Environ. 34 (3), 397–406.
Paden, I., García-Sánchez, C., Ledoux, H., 2022. Towards automatic reconstruction of 3D city models tailored for urban flow simulations. Front. Built Environ. 8.
Parra, M.A., Santiago, J.L., Martín, F., Martilli, A., Santamaría, J.M., 2010. A methodology to urban air quality assessment during large time periods of winter using computational fluid dynamic models. Atmos. Environ. 44 (17), 2089–2097.
Perillo, H.A., Broderick, B.M., Gill, L.W., McNabola, A., Kumar, P., Gallagher, J., 2022. Spatiotemporal representativeness of air pollution monitoring in Dublin, Ireland. Sci. Total Environ. 827, 154299. https://doi.org/10.1016/j.scitotenv.2022.154299.
Piersanti, A., Vitali, L., Righini, G., Cremona, G., Ciancarella, L., 2015. Spatial representativeness of air quality monitoring stations: a grid model-based approach. Atmos. Pollut. Res. 6, 953–960. https://doi.org/10.1016/j.apr.2015.04.005.
Rafael, S., Rodrigues, V., Oliveira, K., Coelho, S., Lopes, M., 2021. How to compute long-term averages for air quality assessment at urban areas? Sci. Total Environ. 795, 148603.
Reiminger, N., Jurado, X., Vazquez, J., Wemmert, C., Blond, N., Wertel, Jonathan, Dufresne, Matthieu, 2020a. Methodologies to assess mean annual air pollution concentration combining numerical results and wind roses. Sustain. Cities Soc. 59, 102221.
Reiminger, N., Vazquez, J., Blond, N., Dufresne, M., Wertel, J., 2020b. CFD evaluation of mean pollutant concentration variations in step-down street canyons. J. Wind Eng. Ind. Aerodyn. 196, 104032.
Righini, G., Cappelletti, A., Ciucci, A., Cremona, G., Piersanti, A., Vitali, L., Ciancarella, L., 2014. GIS based assessment of the spatial representativeness of air quality monitoring stations using pollutant emissions data. Atmos. Environ. 97, 121–129. https://doi.org/10.1016/j.atmosenv.2014.08.015.
Rivas, E., Santiago, J.L., Lechón, Y., Martín, F., et al., 2019. CFD modelling of air quality in Pamplona City (Spain): assessment, stations spatial representativeness and health impacts valuation. Sci. Total Environ. 649, 1362–1380.
Romberg, E., Bosinger, R., Lohmeyer, A., Ruhnke, R., Roth, E., 1996. N0-N02-Umwandlungsmodell fur die Anwendung bei Immissionsprognosen fur Kfz-Abgase. Gefahrstoffe-Reinhaltung Luft 56, 215–218.
Russo, F., Villani, M.G., D'Elia, I., D'Isidoro, M., Liberto, C., Piersanti, A., Tinarelli, G., Valenti, G., Ciancarella, L., 2021. A study of traffic emissions based on floating car data for urban scale air quality applications. Atmosphere 12 (8).
Sakellaris, I.A., Bartzis, J.G., Neuhäuser, J., Friedrich, R., Gotti, A., Sarigiannis, D.A., 2022. A novel approach for air quality trend studies and its application to European urban environments: the ICARUS project. Atmos. Environ. 273, 118973.
San Jose, R., Perez-Camanyo, J.L., 2022. Modelling effects of type of trees on urban air pollution with a computational fluid dynamics model. Euro-Mediterr. J. Environ. Integr. 7, 381–389.
San Jose, R., Perez-Camanyo, J.L., 2023. High-resolution impacts of green areas on air quality in Madrid. Air Qual. Atmos. Health 16, 37–48.
San Jose, R., Perez-Camanyo, J.L., Gonzalez-Barras, R.M., 2021. The use of LES CFD urban models and mesoscale air quality models for urban air quality simulations. In: Studies in Systems, Decision and Control, 7, pp. 185–199.
Sanchez, B., Santiago, J.L., Martilli, A., Martin, F., Borge, R., Quaassdorff, C., de la Paz, D., 2017. Modelling NOx concentrations through CFD-RANS in an urban hot-spot using high resolution traffic emissions and meteorology from a mesoscale model. Atmos. Environ. 163, 155–165.
Santiago, J.L., Martin, F., 2015. Use of CFD modeling for estimating spatial representativeness of urban air pollution monitoring sites and suitability of their locations. Fis. Tierra 191–221.
Santiago, J.L., Martín, F., Martilli, A., 2013. A computational fluid dynamic modelling approach to assess the representativeness of urban monitoring stations. Sci. Total Environ. 454-455, 61–72.

Santiago, J.L., Borge, R., Martin, F., de la Paz, D., Martilli, A., Lumbreras, J., Sanchez, B., 2017. Evaluation of a CFD-based approach to estimate pollutant distribution within a real urban canopy by means of passive samplers. Sci. Total Environ. 576, 46–58.

Santiago, J.L., Borge, R., Sanchez, B., Quaassdorff, C., de la Paz, D., Martilli, A., Rivas, E., Martín, F., 2021. Estimates of pedestrian exposure to atmospheric pollution using high-resolution modelling in a real traffic hot-spot. Sci. Total Environ. 755, 142475.

Santiago, J.L., Rivas, E., Sanchez, B., Buccolieri, R., Esposito, A., Martilli, A., Martín, F., 2022a. Impact of different combinations of green infrastructure elements on traffic-related pollutant concentrations in urban areas. Forests 13 (8), 1195.

Santiago, J.-L., Sanchez, B., Rivas, E., Vivanco, M.G., Theobald, M.R., Garrido, J.L., Gil, V., Martilli, A., Rodríguez-Sánchez, A., Buccolieri, R., et al., 2022b. High spatial resolution assessment of the effect of the Spanish National Air Pollution Control Programme on street-level NO2 concentrations in three neighborhoods of Madrid (Spain) using mesoscale and CFD modelling. Atmosphere 13, 248. https://doi.org/10.3390/atmos13020248.

Sousa, J., Gorlé, C., 2019. Computational urban flow predictions with Bayesian inference: validation with field data. Build. Environ. 154, 13–22.

Sousa, J., García-Sánchez, C., Gorlé, C., 2018. Improving urban flow predictions through data assimilation. Build. Environ. 132, 282–290.

Stedman, J.R., Kent, A.J., Grice, S., Bush, T.J., Derwent, R.G., 2007. A consistent method for modelling PM10 and PM2.5 concentrations across the United Kingdom in 2004 for air quality assessment. Atmos. Environ. 41, 161–172.

Su, L., Gao, C., Ren, X., Zhang, F., Cao, S., Zhang, S., Chen, T., Liu, M., Ni, B., Liu, M., 2022. Understanding the spatial representativeness of air quality monitoring network and its application to PM2.5 in the mainland China. Geosci. Front. 13, 101370. https://doi.org/10.1016/j.gsf.2022.101370.

Tarrasón, L., Hak, C., Soares, J., Vika, Røen H., Ødegård, R., Green, J., Marsteen, L., 2020. Assessing the spatial representativeness of air quality sampling points. Application of siting criteria and sampling point classification – Task 3 report. In: Service Request 5 under Framework Contract ENV.C.3/FRA/2017/0012. Specific Contract: 07.0203/2018/793545/SFRA/ENV.C.3. Final Report for European Commission -DG Environment. Ares, 4920320 (ED 11492 | Task 3 | Date 21/12/2020. Ricardo Confidential. 2020).

Trini, Castelli S., Armand, P., Tinarelli, G., Duchenne, C., Nibart, M., 2018. Validation of a Lagrangian particle dispersion model with wind tunnel and field experiments in urban environment. Atmos. Environ. 193, 273–289.

Van Aalst, R., Edwards, L., Pulles, T., De Saeger, E., Tombrou, M., Tønnesen, D., 1998. Guidance report on preliminary assessment under EC air quality directives. European Environmnet Agency Technical report 11/-1. https://www.eea.europa.eu/en/analysis/publications/tec11a.

Veratti, G., Fabbi, S., Bigi, A., Lupascu, A., Tinarelli, G., Teggi, S., Brusasca, G., Butler, T. M., Ghermandi, G., 2020. Towards the coupling of a chemical transport model with a micro-scale Lagrangian modelling system for evaluation of urban NOx levels in a European hotspot. Atmos. Environ. 223, 117285.

Villani, M.G., Russo, F., Adani, M., Piersanti, A., Vitali, L., Tinarelli, G., Ciancarella, L., Zanini, G., Donateo, A., Rinaldi, M., Carbone, C., Decesari, S., Sänger, P., 2021. Evaluating the impact of a wall-type green infrastructure on PM10 and NOx concentrations in an urban street environment. Atmosphere 12 (7).

Vitali, L., Morabito, A., Adani, M., Assennato, G., Ciancarella, L., Cremona, G., Giua, R., Pastore, T., Piersanti, A., Righini, G., Russo, F., Spagnolo, S., Tanzarella, A., Tinarelli, G., Zanini, G., 2016. A Lagrangian modelling approach to assess the representativeness area of an industrial air quality monitoring station. Atmos. Pollut. Res. 7, 990–1003. https://doi.org/10.1016/j.apr.2016.06.002.

Vivanco, M.G., Theobald, M., Garrido, J.L., Gil, V., Martín, F., 2018. Evaluación de la Calidad del Aire en España Utilizando Modelización Combinada con Mediciones; Preevaluación Año 2017; Acuerdo de Encomienda de Gestión 2014–2018 Entre el Ministerio de Agricultura, Alimentación y Medio Ambiente y CIEMAT; CIEMAT Report. CIEMAT, Madrid, Spain.

Vranckx, S., Vos, P., Maiheu, B., Janssen, S., 2015. Impact of trees on pollutant dispersion in street canyons: a numerical study of the annual average effects in Antwerp, Belgium. Sci. Total Environ. 532, 474–483.

Weydahl, T., Markelj, M., Høyem, H., 2023. Revidert tiltaksutredning for lokal luftkvalitet i Drammen. NILU rapport 9/2023. ISBN: 978-82-425-3119-3. https://hdl.handle.net/11250/3070408.

Weydahl, T., Walker, S.E., Markelj, M., 2024. Trafikk- og luftkvalitetsberegninger i forbindelse med ny tiltaksutredning for bedre luftkvalitet i Oslo og Bærum 2025–2030. Beregningsresultater for Oslo. NILU rapport 31/2024, ISBN: 978-82-425-3180-3. https://hdl.handle.net/11250/3170428.

Yatkin, S., Gerboles, M., Belis, C., Karagulian, F., Lagler, F., Barbiere, M., Borowiak, A., 2020. Representativeness of an air quality monitoring station for pm2.5 and source apportionment over a small urban domain. Atmos. Pollut. Res. 11, 225–233. https://doi.org/10.1016/j.apr.2019.10.004.

Yu, T., Wang, W., Ciren, P., Sun, R., 2018. An assessment of air-quality monitoring station locations based on satellite observations. Int. J. Remote Sens. 39, 6463–6478. https://doi.org/10.1080/01431161.2018.146050.

Zhong, J., Hood, C., Johnson, K., Stocker, J., Handley, J., Wolstencroft, M., Mazzeo, A., Cai, X., Bloss, W.J., 2021. Using task farming to optimise a street-scale resolution air quality model of the West Midlands (UK). Atmosphere 12 (8).

Zhong, J., Stocker, J., Cai, X., Harrison, R.M., Bloss, W.J., 2024. Street-scale air quality modelling over the west midlands, United Kingdom: effect of idealised traffic reduction scenarios. Urban Clim. 55. Article 101961.

Zhu, Y., Chen, J., Bi, X., Kuhlmann, G., Chan, K.L., Dietrich, F., Brunner, D., Ye, S., Wenig, M., 2020. Spatial and temporal representativeness of point measurements for nitrogen dioxide pollution levels in cities. Atmos. Chem. Phys. 20, 13241–13251. https://doi.org/10.5194/acp-20-13241-2020.