Use of CFD modeling for estimating spatial representativeness of urban air pollution monitoring sites and suitability of their locations

Jose Luis SANTIAGO, Fernando MARTÍN
División de Contaminación Atmosférica
Dpto. Medio Ambiente. CIEMAT
Avda. Complutense 40, Ed. 3, 28040 Madrid
jl.santiago@ciemat.es

Received: 20/05/2015
Accepted: 04/08/2015

Abstract
A methodology to estimate the spatial representativeness of air pollution monitoring sites is applied to two urban districts. This methodology is based on high resolution maps of air pollution computed by using Computational Fluid Dynamics (CFD) modelling tools. Traffic-emitted NO_2 dispersion is simulated for several meteorological conditions taking into account the effect of the buildings on air flow and pollutant dispersion and using a steady state CFD-RANS approach. From these results, maps of average pollutant concentrations for January–May 2011 are computed as a combination of the simulated scenarios. Two urban districts of Madrid City were simulated. Spatial representativeness areas for 32 different sites within the same district (including the site of the operative air quality stations) have been estimated by computing the portion of the domains with average NO_2 concentration differing less than a 20% of the concentration at each candidate monitoring site. New parameters such as the ratio *AR* between the representativeness area and the whole domain area or the representativeness index (*IR*) has been proposed to discuss and compare the representativeness areas. Significant differences between the spatial representativeness of the candidate sites of both studied districts have been found. The sites of the Escuelas Aguirre district have generally smaller representativeness areas than those of the Plaza de Castilla. More stations are needed to cover the Escuelas Aguirre district than for the Plaza de Castilla one. The operative air quality station of the Escuelas Aguirre district is less representative than the station of the Plaza de Castilla district. The cause of these differences seems to be the differences in urban structure of both districts prompting different ventilation.
Key words: Air quality, CFD modelling, monitoring stations, spatial representativeness

Uso de modelos CFD para estimar la representatividad espacial de estaciones de medida de la contaminación atmosférica urbana y la idoneidad de sus ubicaciones

Resumen
Se ha aplicado una metodología para estimar la representatividad espacial de sitios de monitoreo de contaminación atmosférica en dos distritos urbanos. Esta metodología se basa en mapas de alta resolución de la contaminación del aire calculados mediante modelos CFD (Computational Fluid Dynamics). Se ha simulado la dispersión de NO_2 emitido por el tráfico para varias condiciones meteorológicas, teniendo en cuenta el efecto de los edificios en los flujos de aire y la dispersión de contaminantes y utilizando simulaciones CFD-RANS en estado estacionario. A partir de estos resultados se calcularon mapas de las concentraciones de contaminantes promedio para el periodo Enero-Mayo de 2011 como una combinación de los escenarios simulados. Se simularon dos distritos urbanos de la ciudad de Madrid. Se han estimado las zonas de representatividad espacial para 32 ubicaciones diferentes

ISSN: 0214-4557
http://dx.doi.org/10.5209/rev_FITE.2015.v27.51200

dentro del mismo distrito (incluyendo la ubicación de las estaciones de calidad del aire operativas) mediante el cálculo de la parte de ambos dominios en la que la concentración media de NO_2 difiere menos de un 20% de la concentración en cada ubicación. Se han propuesto nuevos parámetros como la relación entre el área representatividad y toda la zona de dominio (*AR*) o el índice de representatividad (*IR*) para discutir y comparar las áreas de representatividad. Se han encontrado diferencias significativas entre la representatividad espacial de los sitios candidatos de ambos distritos estudiados. Las ubicaciones del distrito Escuelas Aguirre tienen áreas de representatividad generalmente más pequeñas que las de la Plaza de Castilla. Se necesitan más estaciones para cubrir el distrito Escuelas Aguirre que para la Plaza de Castilla. La estación operativa la calidad del aire del distrito Escuelas Aguirre es menos representativa que la estación del barrio de la Plaza de Castilla. La causa de estas diferencias parece estar en la diferente ventilación en ambas zonas debido a diferencias en la estructura urbana de ambos distritos.
Palabras clave: Calidad del aire, modelización CFD, estaciones de medida, representatividad espacial

Contents:

Normalized reference
Santiago J.L., Martín F. (2015). Use of CFD modeling for estimating spatial representativeness of urban air pollution monitoring sites and suitability of their locations. *Física de la Tierra,* Vol. 27, 195-225

1. Introduction

In the streets and squares of a city the distribution of air pollutants is very heterogeneous and a large concentration difference between two nearby locations can be found. This is due to the interaction of the atmosphere with urban surfaces (buildings, trees, etc.) that cause complex airflows influencing the dispersion of pollutants. It is therefore important and difficult issue to find good location of urban air quality stations in order to make as representative as possible the recorded concentration data. This question is addressed even also by the air quality legislation. For example, according micro-implantation criteria of Directive 2008/50 / EC, Annex III, an urban station should be, as far as possible, representative of a street segment no less than 100 m. Then, to deploy a large amount of air quality monitoring stations can help to cover most of the urban territory. However, it should be extremely expensive. Hence, it is very important to analyze the spatial representativeness (SR, hereafter) of the air quality (AQ, hereafter) monitoring stations and to find out what could be the best locations studying what would happen in other possible locations around the monitoring station. This is an important question, which has been investigated and discussed intensively in the past. First works are from Ott and Elliasen (1973) who applied a survey technique sampling air CO concentrations in streets and urban areas in order to estimate whether an AQ station represents all the urban area of San José (California). Presently, some European expert groups such as AQUILA (Air Quality Reference Laboratories) (Geiger et al., 2014) and FAIRMODE (Forum for Air Quality Modelling in Europe) (Castell-Balaguer and Denby, 2012) is discussing this issue.

SR depends on how the air pollutant concentration is distributed around the monitoring site, which in turn depends on the topography or obstacles, air flows, distribution of pollution sources, averaging time and pollutant type. These factors can generate different areas of representativeness on different timescales and for different time periods. However, there is no agreement for setting up a reference procedure or methodology to objectively estimate the SR of air quality stations.

In the literature, we can find several methods for estimating the SR of stations. The common factor between them is all of these methods attempt to find out how the pollution is distributed around the site of a certain station and the area where pollutant concentrations do not differ by more than a certain percentage of the measured concentration at the station site (Spangl et al, 2007). They can be classified in three groups:

1) Use of data from experimental campaigns to obtain air pollution concentration measurements around the station. It can be done passive samplers deployed during periods of several days or weeks (Galán Madruga et al., 2001, Vardulakis et al, 2005) or with dense monitoring networks (Blanchard et al, 1999). More deployed samplers or more monitors, better information about the spatial distribution of pollutants. However, it can be very expensive.
2) Use of some surrogate indicators or proxies related to emission sources distribution, land use data, etc. With these data, knowledge of the pollutant spatial distribution is obtained in spite of the effect of transport and dispersion of pollutants is not directly computed (Janssen et al., 2012).
3) Use of air quality modelling to estimate detailed maps of the distribution of air pollutants around the monitoring site. This methodology has been applied to several types of stations using suitable models depending of the spatial scale, for example, for urban traffic-oriented stations with CFD models (Santiago et al, 2013) or for rural background stations using the WRF/CHIMERE models (Martín et al, 2014). The advantage of using validated models is that the effects of emission sources and atmospheric pollutant processes are both taken into account, providing a quite realistic picture of air pollutant concentration. The main disadvantage is the computational cost and uncertainties.

A very complete review of the criteria and methods for air quality classification and representativeness estimate was made by Spangl et al. (2007). The authors are also carrying out a feasibility study for an intercomparison exercise of methodologies for estimating SR in the framework of the FAIRMODE activities. A report including a deep state-of-the-art discussion will be released in the next months.

The aim of this paper is to show how to use the CFD (Computational Fluid Dynamics) modelling tools for estimating high resolution maps of air pollutants in urban districts to estimate the spatial representativeness of monitoring sites and help to select the optimal locations, which have a better spatial representativeness.

Two districts of the Madrid City has been studied, both having an air quality monitoring station, which belongs to the Air Quality Monitoring Network of the Madrid Council. CFD simulations of the pollutant dispersion have been carried out by applying the methodology developed by Santiago et al. (2011a,b) and Santiago et al. (2013). The SR of the monitoring stations has been estimated and also for other possible monitoring sites. The results of SR are discussed in order to determine other better locations for measuring the air pollutant concentrations. Some indexes have been proposed helping to compare the SR of the several potential monitoring sites.

2. Methodology

High resolution maps of air pollutant concentrations in the streets are needed to estimate the SR of an urban air quality station. Inside the street-canyons, complex air flows are generated as result of the perturbation produced by mainly the buildings and the street trees. These flows and the turbulence drive the dispersion of pollutants inside the street-canyon. CFD models are the best tool to simulate the air flows and dispersion of pollutants in these cases with a spatial resolution of the order of 1 m being able to catch the strong spatial heterogeneities in the distribution of pollutants inside the streets. It is described in Section 2.1.

When high resolution maps of pollutant concentration are obtained, criteria for defining the SR area of a monitoring site has to be set up. In addition, a methodology for comparing the obtained SR for several candidates monitoring sites in the modelled area has to be defined. This is described in section 2.2.

2.1. Modelling approach

High resolution maps of average concentration over a large period of time (in these cases several months) are necessary to analyse SR of AQ stations. Due to computational loads, it is not possible (within a reasonable CPU time) to run an unsteady CFD simulation. The solution proposed (Parra et al., 2010; Santiago et al., 2011a,b; Santiago et al., 2013) is to run only a set of scenarios (16 inlet wind directions) using steady CFD-RANS simulations (STAR-CCM+ code from CD-Adapco). The emissions are modelled with a line source inside each street and several tracers (one for each street) are emitted in each simulation. The final map of average concentration is made by means of a combination of the simulated scenarios. Pollutant concentration is computed assuming non-reactive pollutants, thermal effects negligible in comparison with dynamical effects, emissions inside each street at a selected hour proportional to the number of cars at that hour and tracer concentration at a certain hour only depending on emissions and background at that hour. In addition, in these two districts of the Madrid, the dynamic effects of vegetation are included in CFD simulations assuming the trees as a porous medium. The background concentration at each hour from urban background stations is added to the results of the simulations as they only represent the local traffic impact. Finally, some modifications to the methodology of Parra et al. (2010) are introduced to take into account weak winds. For these cases, several assumptions are not fulfilled (e.g. thermal effects have more importance, concentration is not proportional to 1/v, etc).

Taking into account all of these assumptions, for wind speed higher than 2 m s^{-1} at 10 m-height a concentration proportional to the real one at time t is computed using CFD results and Eq. 1.

$$C_{\text{Real}}(t) - C_{background}(t) \propto C_{\text{computed}}(Sector(t)) = \sum_i C_i(Sector(t)) \cdot \frac{L_i}{Vsource_i} \cdot N_i(t) \cdot \frac{1}{v_{in}(t)} \quad (1)$$

where $C_{Real}(t)$ and $C_{background}(t)$ are the real and background concentration at time t, $Sector(t)$ is the wind direction sector at time t, i indicates the tracer emitted inside each street, $C_i(Sector(t))$ is the concentration computed for $Sector(t)$ for a given emission from street i and for a given inlet wind speed, L_i is the length of the street i, $Vsource_i$ is the volume of the row of computational cells where emission of the street i is located, $N_i(t)$ is the number of cars per unit time in street i and $v_{in}(t)$ is the inlet wind speed.

In the case of weak winds (lower than 2 m s^{-1}) the methodology was modified. For these cases, thermal effects are not negligible and we assume that the pollutant concentration is independent of wind direction and it is modelled depending on traffic intensity and mixing height along the day. This assumption allows taking into account indirectly the thermal effects. A simple parameterization of mixing height is considered. From 1h to 8h we considered that it was 100 m, from 15h to 18h was 1500m and for the other hours a linear change between these values was assumed. In future works these cases have to be investigated in depth.

The simulated period was January-May 2011. Meteorological data are from the AEMET (Spanish Meteorological Agency) of a near station are used. Comparison with observed data in the air quality stations were carried out for hourly data for NO_2 showing quite good agreement with the observations. Averaged NO_2 concentration maps for the time period of January-May 2011 were computed by applying weighted average for the simulations done for every wind sectors taking into account the wind speeds and directions frequencies occurred during the studied period and the dependence of the concentrations with the wind speed (see Eq. 1) for wind speed higher than 2 m s^{-1}. For weak wind speed a dependency proportional to traffic intensity and 1/mixing $height^{0.3}$ was assumed. The dependency with mixing height was fitted from experimental data. This assumption should be investigated in future works studying in depth urban thermal effects. More details about the methodology are shown in Parra et al. (2010) and Santiago et al. (2011a,b).

Validation exercises of the CFD-RANS simulations has been done by Parra et al. (2010) for an urban district of Pamplona using measurements from experimental campaigns using passive samplers and monitoring data, and by Santiago et al. (2011a,b) and Santiago et al. (2013) for the same Madrid districts using data from air quality monitoring stations. In all cases, the modelled concentrations fit well the observations with high correlation coefficients and without significant bias.

These studied urban districts are:

1. Escuelas Aguirre district (Figure 1)

The size of the modelling domains was 700 m x 800 m, approximately (Figure 1). Hourly traffic data inside each street was not available, but this information was estimated from the daily average traffic intensity (TI) inside each street (provided by the city council of Madrid) multiplied by the time evolution of traffic for different types of days (TF) (obtained from Palacios, 2001). In this location, 29 passive tracers were released (one for each street). An irregular mesh of $3 \cdot 10^6$ grid points with a resolution in horizontal and vertical directions of about 1 m- 3 m close to the buildings is used. The height of the domain is 540 m (6 times the height of the tallest building).

In this zone there are several streets and avenues with intense road traffic and also there is a large green urban area (El Retiro). Buildings with different heights are located in this area; most of them have a height between 18 m and 24 m, although the tallest building is 90 m high approximately. The traffic-oriented air quality station (EA) is located in a garden very close to the sidewalk in the intersection of the two more important avenues (Figure 1). Meteorological data were provided by AEMET (Spanish Meteorological Agency) from the meteorological station located in El Retiro (the urban park of location 2 at South of EA station) but outside of the numerical domain (Figure1). In this zone, there is an urban park that produces not negligible dynamical effects on pollutant dispersion. These dynamical effects of the vegetation are modelled considering trees as porous medium. More details can be found in Santiago et al. (2013), Santiago et al. (2014) and Krayenhoff et al. (2015). Inside the park, vegetation is present in the numerical cells located from the ground to 18 meters-height (trees of 18 m) but for the trees within the street only the cells located at the crown level have vegetation.

Figure 1. Escuelas Aguirre district. Left: Aerial picture of the domain (red line) including the location of the air quality station. Right: Modeled geometry of the domain including street and park vegetation.

2. Plaza de Castilla district (Figure 2)

This district corresponds to an area of high traffic intensity it is characterized by a more open area than Escuelas Aguirre district but where there are several very high buildings. The size of the modelling domain is 1 km x 1 km, approximately. In this location, 31 passive tracers were released (one for each street). An irregular mesh of $2 \cdot 10^6$ grid points with a resolution in horizontal and vertical directions of about 1 m-3 m close to the buildings is used. The height of the domain is 570 m (more than 5 times the height of the tallest building).

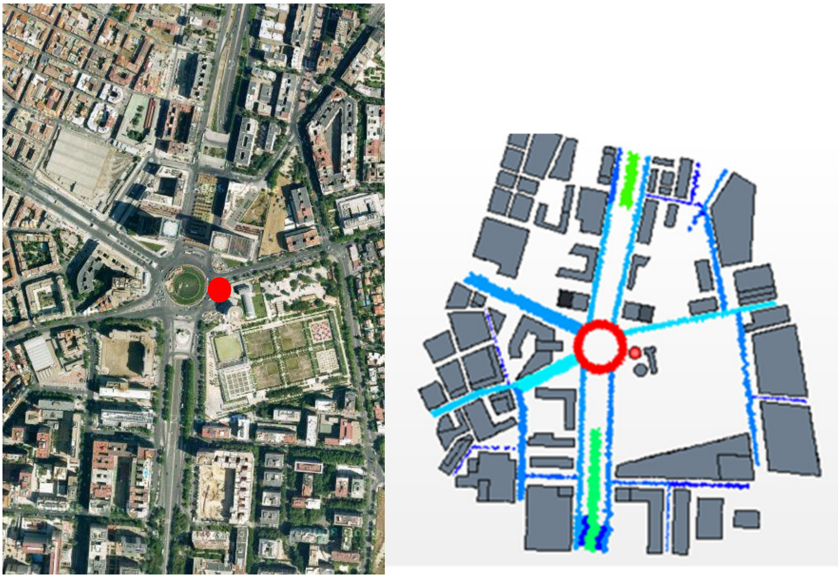

Figure 2. Plaza de Castilla district. Left: Aerial picture of the domain (red line) including the location of the air quality station. Right: Modeled geometry of the domain including traffic emissions.

2.2. Criteria for spatial representativeness estimates and monitoring site selection

In this study the criterion for determining spatial representativeness (SR) area is the portion of territory with a concentration value does not vary more than 20% of that at the station site. The value of 20% is related to the allowed average experimental error of the monitoring devices for several urban pollutants (see Directive 2008/50/EU).

The above referred directive referring to the human health protection said that the monitoring sites has to be representative of "the areas within zones and agglomerations where the highest concentrations occur to which the population is likely to be directly or indirectly exposed for a period which is significant in relation to the averaging period of the limit value(s)", so it is important that the station capture as much as possible areas of high concentrations but avoiding "the very small micro-environments in their immediate vicinity are measured".

Hence, respect to the discussion about the suitability of a monitoring site, besides the size of the SR area, other aspects should also be taken into account such as the exposure of the population. However, it is a very difficult issue because to know how much pollutant is inhaled by the population, it would be necessary to know the residence time of the individuals at each location throughout the day. However, at least it is necessary to take into account the concentration values in the discussion of the suitability of the monitoring site. According to this, an index of representation (*IR*) has been defined as a Gaussian function that takes into account both the SR area and the portion of the zone of higher concentrations that is out of the SR area.

$$IR = e^{-\left[\frac{(AR-100)^2}{2\sigma_x}+\frac{(ANRmayor)^2}{2\sigma_y}\right]} \quad (2)$$

where *AR* is percentage of the SR area respect to the total modeled domain, *ANRmayor* is the percentage of non-representativeness area with concentrations higher than that at the monitoring site. *ANRmayor* equals to 0 means that the SR area of the monitoring site includes the highest concentrations of the studied area. Higher *ANRmayor*, more zones with high concentrations are out of the SR area of the monitoring site. σ_x and σ_y are weights in the equation. For this study, $\sigma_x = \sigma_y = 42.48$ in order to get circular isolines and when $AR = 50\%$ and $ANRmayor = 0\%$ then $IR = 0.5$ (see Figure 3). *IR* ranges between 0 and 1. IR = 1 is the best value (SR area cover all the studied domain). Other interpretation of the IR index is that a monitoring site, which SR area includes the highest pollutant concentrations is equivalent to other monitoring site with a larger SR area but without including inside the highest pollutant concentrations (e.g., the case of *AR* = 50% and *ANRmayor* = 0% has the same *IR* than the case of *AR* ≈ 65% and *ANRmayor* ≈ 35%).

The IR index has been computed for the monitoring sites corresponding to the present AQ stations and other candidate locations in both urban districts using the results of the CFD modelling. Maps of NO_2 concentration at 3 m above surface level have been obtained and used for computing *AR, ANRmayor* and *IR* for every present and candidate location.

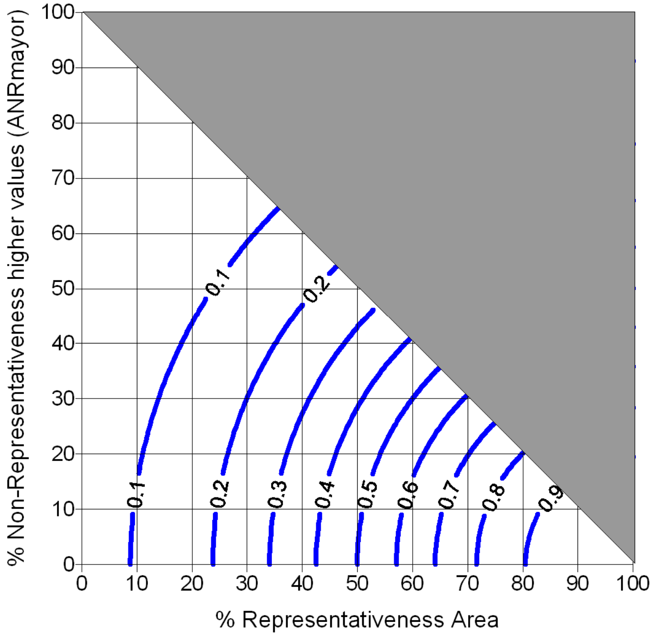

Figure 3. Values of *IR* as function of the percentage of SR area (x-axis) and the percentage of non-representative area with pollutant concentrations higher than at monitoring site (y-axis).

3. Results and discussion

Different candidate locations for air quality monitoring stations within the two modelled urban districts were analyzed and their SR area and IR index was calculated. Then we estimated what would be the best locations for sitting air quality stations in order to get the best representation of the air quality in each district. In each of them, a rectangle containing most of the simulated domain where the SR will be analyzed was defined. Note that the area occupied by buildings was excluded for the calculations.

Escuelas Aguirre district

To characterize this area, 32 different sites which theoretically an AQ station could be located have been selected. Figure 4 shows the map with the average NO_2 concentration computed from the CFD-RANS simulations and the locations of the 32 candidate sites. Site 1 is the location of a presently operative AQ station. Figure 5 shows the spatial representativeness area (in white) for site 1 in the Escuelas Aguirre district and the normalized average concentration respect to the monitoring average concentration.

For each candidate site, the percentage of the SR area respect to the total area inside the rectangle was calculated. In addition we have computed the percentage of non-representative area which includes higher concentrations and lower concentrations than that in the candidate monitoring site (*ANRmayor* and *ANRmenor*, respectively). The value of the proposed index (*IR*) and the modelled concentration is also calculated. These values are shown in Table 1.

Figures 4 and 5 show the high heterogeneity in pollutant concentrations throughout the area with strong gradients inside the streets. The concentration in one sidewalk of a street can be more than twice the concentration in the opposite one. This strong spatial variability makes difficult that the pollutant concentration recorded at a monitoring site can suitably represent the air pollution of a wide area of the district. It is confirmed by the data of Table 1. Any site represents more than 40% of the area of the zone. According to the results of modelling the current station location (site 1) represents 33%, which is within the range of values of the area of maximum representation for all candidate sites (32 to 39%).

The sites corresponding to high pollutant concentrations have a significantly lower representation (*AR* < 20%) because of they are generally very local maxima. This is the case of sites P2, P10, P18 and P24 (see Figure 6), whose *AR* values are lower than 15% while the *ANRmayor* and *ANRmenor* values are 0% and higher than 80%, respectively. The *IR* values are also very low. These sites are located in streets or crossing with high traffic intensity.

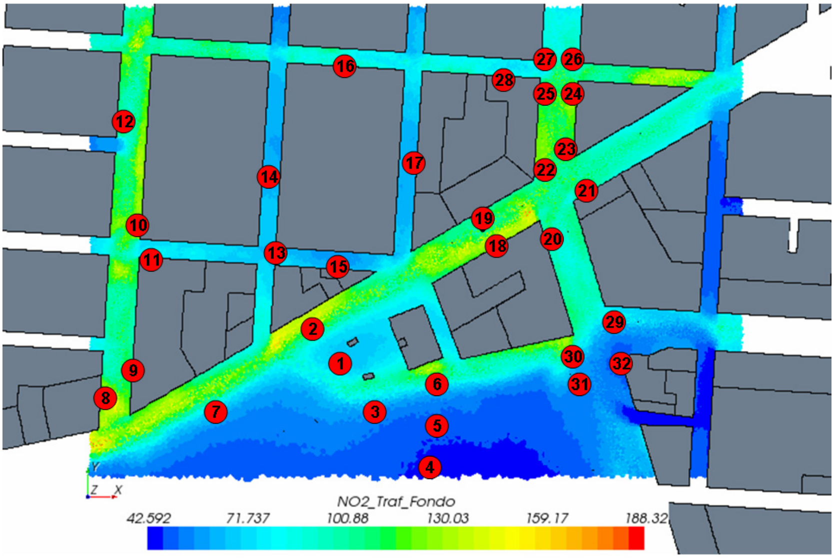

Figure 4. Map with the average NO_2 concentration computed from the CFD-RANS simulations and the locations of the 32 candidates sites for the Escuelas Aguirre district. Location 1 corresponds to the present station. The rectangle used for calculations corresponds to the colored area.

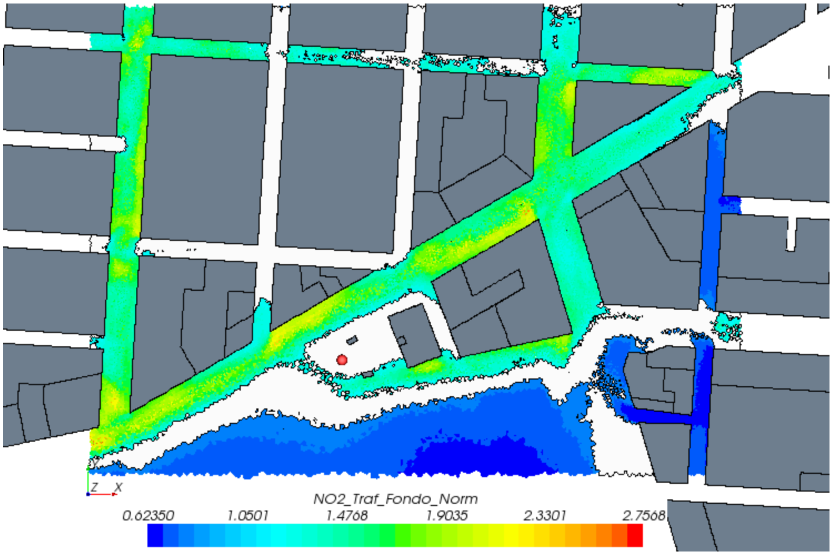

Figure 5. Spatial representativeness area (in white) for the present location (red dot) of the air quality monitoring station (site P1) in the Escuelas Aguirre district and normalized average concentration respect to the monitoring average concentration

Table 1. Modeled NO_2 concentration and values of *AR*, *ANRmayor, ANRmenor* and *IR* calculated for every candidate location for air quality monitoring station. Location 1 corresponds to the present station in the Escuelas Aguirre district.

Location	**Concentration ($\mu g\ m^{-3}$)**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
1	68.3	33	42	25	0.17
2	137.8	14	0	86	0.13
3	53.6	39	61	0	0.13
4	46.7	28	72	0	0.06
5	50.3	33	67	0	0.09
6	134.9	16	0	84	0.14
7	65.9	34	45	21	0.17
8	130.4	19	0	81	0.16
9	108	35	3	62	0.31
10	130.9	19	0	81	0.16
11	79.4	33	29	38	0.22
12	103.6	36	5	59	0.32
13	69.4	32	41	27	0.17
14	63.3	37	47	16	0.18
15	63.2	37	48	15	0.18
16	81.8	33	26	41	0.24
17	62.9	38	48	14	0.18
18	138.2	14	0	86	0.13
19	103.4	36	5	59	0.32
20	95.5	36	11	53	0.31
21	106.7	35	4	61	0.31
22	127.2	22	0	78	0.19
23	115	32	1	67	0.27
24	140.4	13	0	88	0.12
25	112.4	33	2	65	0.29
26	96.4	36	10	54	0.31
27	93.7	36	13	52	0.31
28	75.9	33	34	34	0.21
29	82	33	26	41	0.24
30	108.3	35	3	62	0.31
31	53.8	39	60	0	0.13
32	49.4	32	68	0	0.08

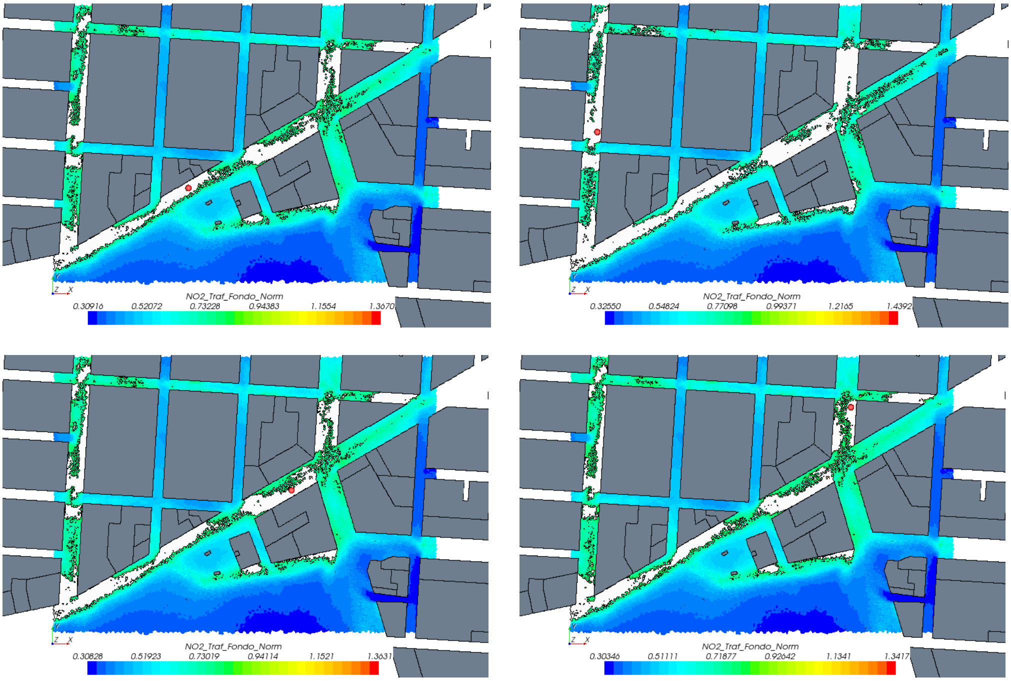

Figure 6. SR areas (white color) corresponding to the sites P2 (upper left), P10 (upper right), P18 (lower left) and P24 (lower right) for the Escuelas Aguirre district. Red dots show the location of the candidate sites. Concentrations are normalized respect to the estimated averaged concentration at the candidate sites.

In an opposite way, sites with low pollutant concentrations have wider representativeness areas. However, site P4 with the lowest concentration has an area of representativeness < 30% (28%), but we have to keep in mind that P4 is within the Retiro Park and then is just representative of the park. Sites located inside or close to the park (P3, P4, P5, P31 and P32) have high values of *AR* (the maximum *AR* is 39% for P4) but also represents areas with relatively low concentrations having high values of *ANRmayor* (> 60%) and low values of *ANRmenor* (=0%) (see Figure 7).

Now, we sort the suitability of the selected locations following different criteria: 1) *AR* (Table 2), 2) *ANRmayor* (Table 3), 3) Representativeness Index (*IR*) (table 4). The site ranking by the *AR* values (Table 2) shows how the 75% of the stations are in a small range of values (32% - 39%) having a relatively larger area. The top of the ranking is for sites P31 and P3, which represent the areas of relatively low concentrations in the district. Analyzing the site ranking according the *ANRmayor* values (Table 3), the best locations classified (*ANRmayor* = 0%) generally have a small *AR* so this classification also seems to be inadequate and insufficient. However, if the *IR* index is considered, which deals with both *AR* and *ANRmayor* trying to maximize the *AR* and minimize the *ANRmayor* (that is, wide representativeness areas including high concentrations), the resulted site ranking (Table 4) changes radically.

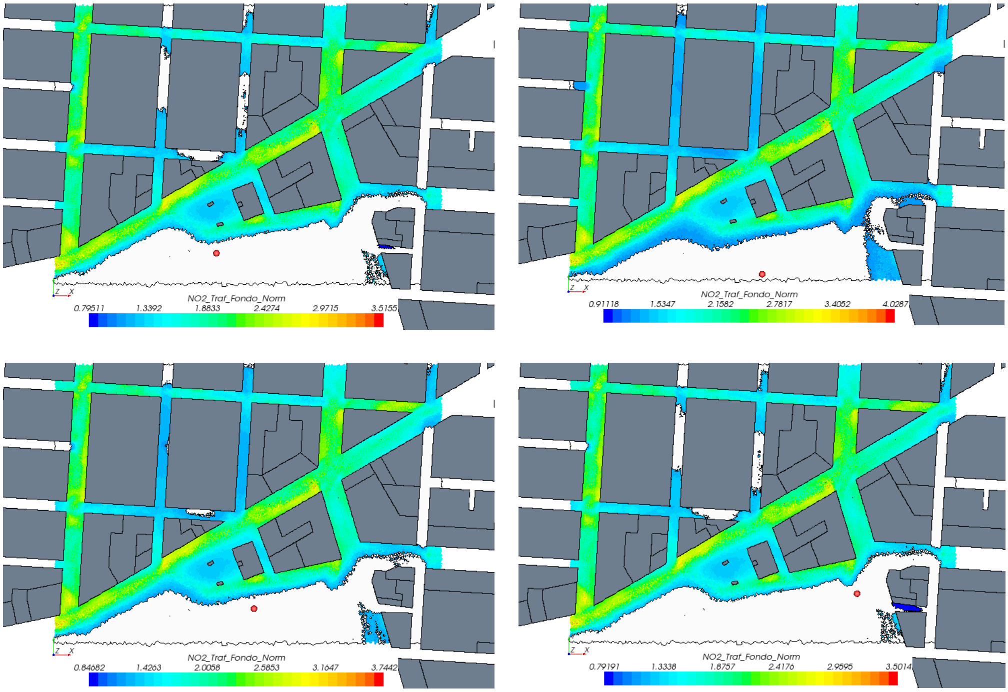

Figure 7. SR areas (white color) corresponding to the sites P3 (upper left), P4 (upper right), P5 (lower left) and P31 (lower right) for the Escuelas Aguirre district. Red dots show the location of the candidate sites. Concentrations are normalized respect to the estimated averaged concentration at the candidate sites.

Comparing Tables 2 and 4, site P31 goes from the first place to be ranked 25th and site P3 from the second position to be 27th, along with *IR* values well below those of the first places. This is because both locations are representative of only low concentrations areas and have similar *AR* to other candidate sites representing higher values of concentration. This is what was sought to define IR: for locations with similar *AR*, those representing the highest concentrations (low *ANRmayor*) within the domain prevail. The *IR* ranking (Table 4) shows how the top positions are occupied by sites providing a medium to high concentration (P12, P19, P26, P21, P20, P9, P27, P30) values. For example, P21 would be a very suitable location because it is in the upper zone of the *IR* ranking and around it seems that the concentration is more homogeneous than in the other locations (see Figure 8). Current location of the operative air quality station (site P1) ranks 20th considering the *IR* index. Although it is noted that the *AR* is very similar to those of the best locations classified and *ANRmayor* and the *IR* values are in middle of the range. In summary, it would be representative of areas of medium-low concentration in the area (Figure 5).

Table 2. Site ranking in the Escuelas Aguirre district sorted by *AR*

Location	**Concentración ($\mu g\ m^{-3}$)**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
31	53.8	39	60	0	0.13
3	53.6	39	61	0	0.13
17	62.9	38	48	14	0.18
15	63.2	37	48	15	0.18
14	63.3	37	47	16	0.18
26	96.4	36	10	54	0.31
12	103.6	36	5	59	0.32
19	103.4	36	5	59	0.32
20	95.5	36	11	53	0.31
27	93.7	36	13	52	0.31
21	106.7	35	4	61	0.31
9	108	35	3	62	0.31
30	108.3	35	3	62	0.31
7	65.9	34	45	21	0.17
5	50.3	33	67	0	0.09
29	82	33	26	41	0.24
16	81.8	33	26	41	0.24
25	112.4	33	2	65	0.29
11	79.4	33	29	38	0.22
28	75.9	33	34	34	0.21
1	68.3	33	42	25	0.17
32	49.4	32	68	0	0.08
13	69.4	32	41	27	0.17
23	115	32	1	67	0.27
4	46.7	28	72	0	0.06
22	127.2	22	0	78	0.19
8	130.4	19	0	81	0.16
10	130.9	19	0	81	0.16
6	134.9	16	0	84	0.14
2	137.8	14	0	86	0.13
18	138.2	14	0	86	0.13
24	140.4	13	0	87	0.12

.

Table 3. Site ranking in the Escuelas Aguirre district sorted by *ANRmayor* in ascending order. Zero indicates that areas with the highest concentration are within the area of representation. The higher *ANRmayor*, more areas of high concentrations fall outside the area of representation.

Location	**Concentración ($\mu g\ m^{-3}$)**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
22	127.2	22	0	78	0.19
8	130.4	19	0	81	0.16
10	130.9	19	0	81	0.16
6	134.9	16	0	84	0.14
2	137.8	14	0	86	0.13
18	138.2	14	0	86	0.13
24	140.4	13	0	88	0.12
23	115	32	1	67	0.27
25	112.4	33	2	65	0.29
30	108.3	35	3	62	0.31
9	108	35	3	62	0.31
21	106.7	35	4	61	0.31
19	103.4	36	5	59	0.32
12	103.6	36	5	59	0.32
26	96.4	36	10	54	0.31
20	95.5	36	11	53	0.31
27	93.7	36	13	52	0.31
29	82	33	26	41	0.24
16	81.8	33	26	41	0.24
11	79.4	33	29	38	0.22
28	75.9	33	34	34	0.21
13	69.4	32	41	27	0.17
1	68.3	33	42	25	0.17
7	65.9	34	45	21	0.17
14	63.3	37	47	16	0.18
15	63.2	37	48	15	0.18
17	62.9	38	48	14	0.18
31	53.8	39	60	0	0.13
3	53.6	39	61	0	0.13
5	50.3	33	67	0	0.09
32	49.4	32	68	0	0.08
4	46.7	28	72	0	0.06

Table 4. Site ranking in the Escuelas Aguirre district sorted by *IR*.

Location	**Concentración ($\mu g\ m^{-3}$)**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
19	103.4	36	5	59	0.32
12	103.6	36	5	19	0.32
26	96.4	36	10	54	0.31
21	106.7	35	4	61	0.31
20	95.5	36	11	53	0.31
9	108	35	3	62	0.31
27	93.7	36	13	52	0.31
30	108.3	35	3	62	0.31
25	112.4	33	2	65	0.29
23	115	32	1	67	0.27
29	82	33	26	41	0.24
16	81.8	33	26	41	0.24
11	79.4	33	29	38	0.22
28	75.9	33	34	34	0.21
22	127.2	22	0	78	0.19
17	62.9	38	48	14	0.18
15	63.2	37	48	15	0.18
14	63.3	37	47	16	0.18
13	69.4	32	41	27	0.17
7	65.9	34	45	21	0.17
1	68.3	33	42	25	0.17
8	130.4	19	0	81	0.16
10	130.9	19	0	81	0.16
6	134.9	16	0	84	0.14
31	53.8	39	60	0	0.13
2	137.8	14	0	86	0.13
3	53.6	39	61	0	0.13
18	138.2	14	0	86	0.13
24	140.4	13	0	88	0.12
5	50.3	33	67	0	0.09
32	49.4	32	68	0	0.08
4	46.7	28	72	0	0.06

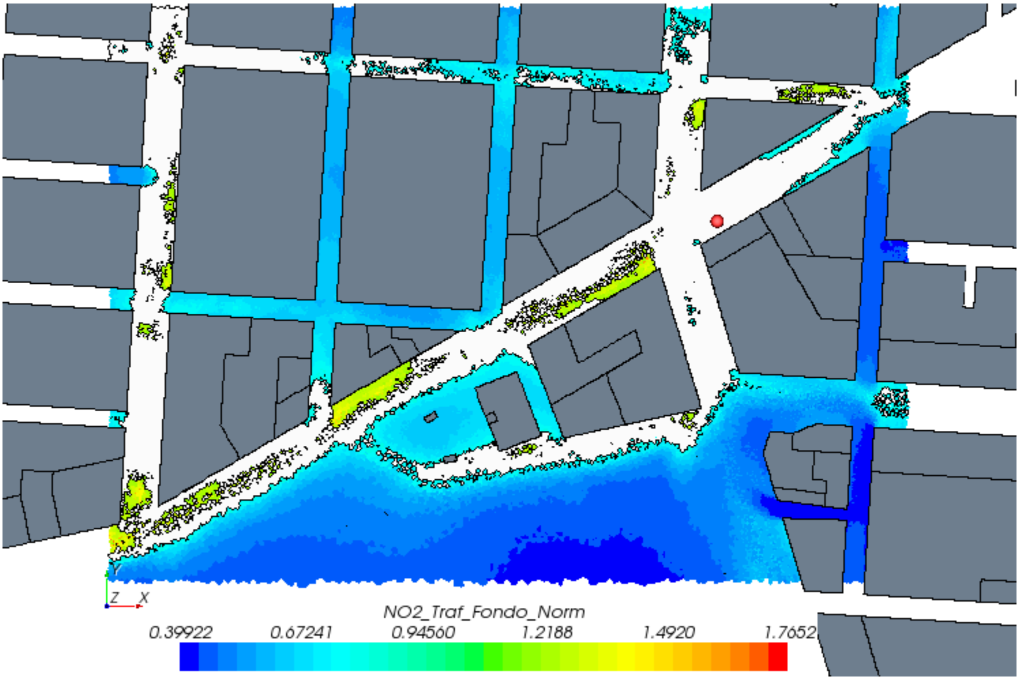

Figure 8. Spatial representativeness area (in white) for the site P21 (red dot) in the Escuelas Aguirre district. Concentrations are normalized respect to the estimated averaged concentration at the candidate sites.

To fully cover the studied domain by adding the spatial representativeness of a minimum number of air quality stations will need at least 4 stations, for example, located at sites P4, P1, P21 and P18. With this idea, the domain can be divided in four sub-areas:

1. low concentrations (next to the park areas) do not have a high *AR*;
2. medium-low concentrations (P1 or positions in narrow streets with not high traffic intensity) having an *AR* within the range of maximum values;
3. medium-high concentrations (in some street sidewalks with high traffic intensity or crossing such streets) having an *AR* within the range of maximum values;
4. high concentrations (local maxima usually on streets with high traffic intensity) have a smaller *AR*.

Plaza de Castilla district

The Plaza de Castilla has been studied following the same methodology as for the Escuela Aguirre district. 32 candidates sites for sitting air quality monitoring stations were analyzed (Figure 9).

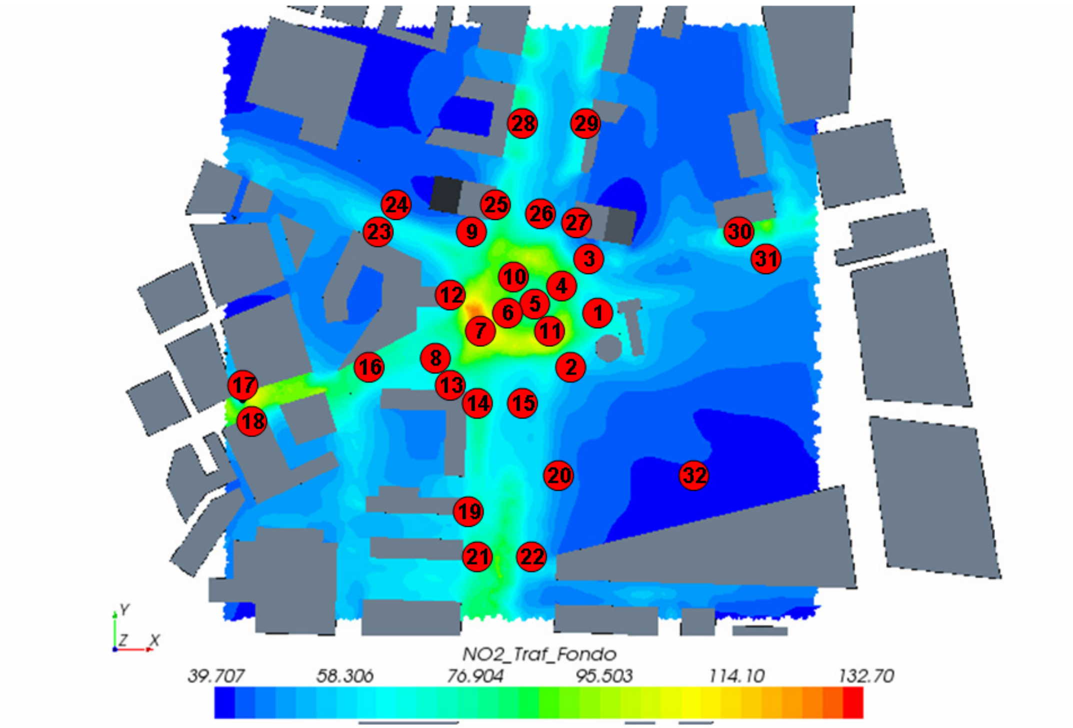

Figure 9. Map with the average NO_2 concentration computed from the CFD-RANS simulations and the locations of the 32 candidates sites for the Plaza de Castilla district. Location 1 corresponds to the present station. The rectangle used for calculations corresponds to the colored area.

The pollutant concentration field estimated from the CFD simulations seems to be less heterogeneous than in the case of the Escuelas Aguirre district because of this is an area with wider avenues, streets and squares favoring ventilation (Figures 9 and 10). High pollutant concentrations are mainly in the square and in the main avenues. The sites of these areas have low values of *AR* and *ANRmayor* (Table 5) corresponding to small spatial representativeness areas (areas of local concentration maxima) as can be seen at sites P7, P11, P17 and P18 (Figure 11).

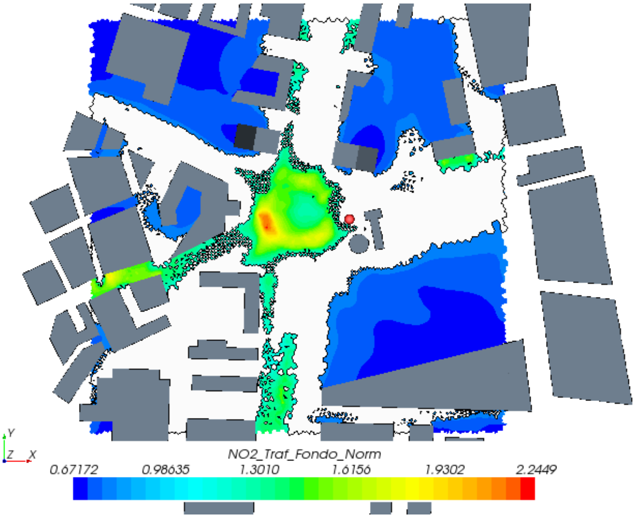

Figure 10. Spatial representativeness area (in white) for the present location (red dot) of the air quality monitoring station (site P1) in the Plaza de Castilla district and normalized average concentration respect to the monitoring average concentration.

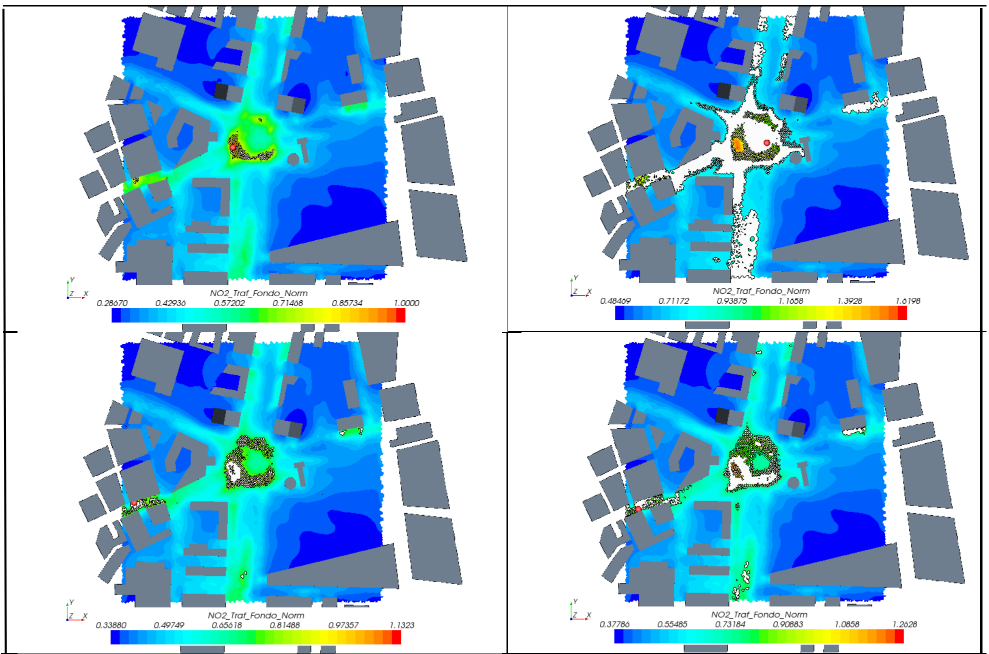

Figure 11. SR areas (white color) corresponding to the sites P7 (upper left), P11 (upper right), P17 (lower left) and P18 (lower right) for the Plaza de Castilla district. Red dots show the location of the candidate sites. Concentrations are normalized respect to the estimated averaged concentration at the candidate sites.

Table 5. Modeled NO_2 concentration and values of *AR*, *ANRmayor, ANRmenor* and *IR* calculated for every candidate location for air quality monitoring station. Location 1 corresponds to the present station in the Plaza de Castilla district.

Location	**Concentración ($\mu g\ m^{-3}$)**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
1	59.1	49	12	39	0.47
2	62.4	39	10	51	0.35
3	47.4	70	30	0	0.61
4	78.9	17	2	81	0.15
5	80.4	16	2	82	0.14
6	85.4	13	1	86	0.12
7	138.5	0	0	100	0.06
8	64.9	34	8	58	0.30
9	50.4	77	23	0	0.75
10	80.6	16	2	82	0.14
11	81.9	15	2	83	0.14
12	62.6	38	9	52	0.34
13	63.9	36	9	56	0.31
14	77.5	18	2	79	0.16
15	57	53	14	33	0.52
16	81.8	15	2	83	0.14
17	117.2	2	0	98	0.07
18	105	5	0	95	0.08
19	70.1	27	5	68	0.23
20	47.8	71	29	0	0.63
21	73.8	22	3	75	0.18
22	47.6	71	29	0	0.62
23	58.9	50	12	38	0.48
24	44	60	40	0	0.41
25	67.6	30	6	64	0.26
26	57.7	52	13	35	0.50
27	52.8	74	19	7	0.75
28	66.2	33	7	61	0.28
29	56.4	55	15	31	0.53
30	85	13	1	86	0.12
31	58.9	50	12	38	0.48
32	42.4	55	45	0	0.33

There are wide areas with low pollutant concentrations and with wide spatial representativeness areas such as the sites P3, P9, P20 and P32, whose values of *AR* and *ANRmayor* generally high (see Figure 12 and Table 5).

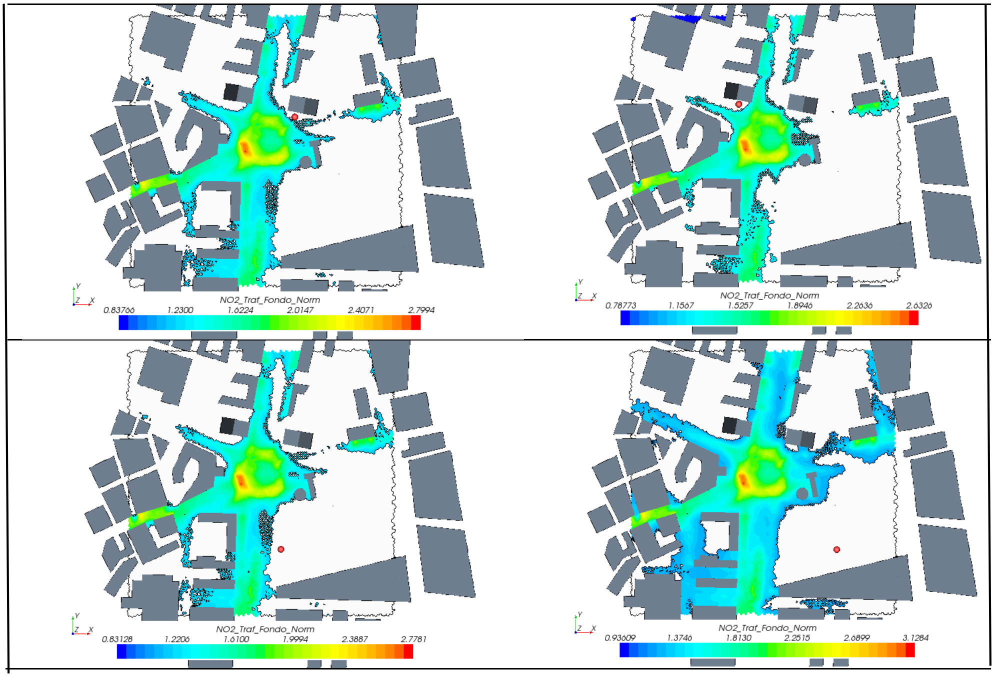

Figure 12. SR areas (white color) corresponding to the sites P3 (upper left), P9 (upper right), P20 (lower left) and P32 (lower right) for the Plaza de Castilla district. Red dots show the location of the candidate sites. Concentrations are normalized respect to the estimated averaged concentration at the candidate sites.

As for the Escuelas Aguirre district, we have sorted the locations depending on the values of *AR* (Table 6), *ANRmayor* (Table 7) and *IR* (Table 8). The representativeness areas of the candidate sites of the Plaza de Castilla district are much larger than those for the Escuelas Aguirre district. The five higher *AR* values range from 70 to 77% (site P1 is the top of the ranking), most of the sites have *AR* values higher than 30% and only 3 sites have *AR* values lower than 10% being site P7 the last of ranking (Table 6). The top position of site ranking according the *ANRmayor* values (Table 7) corresponds to site with very low *AR* showing a very small spatial representativeness. Analyzing the site ranking for *IR*, the best values (0.77) correspond to sites P27 and P9. Both have large representativeness areas with AR values of 74 and 77% and with not very high values of *ANRmayor* (19 and 23%). The estimated concentration at these sites is relatively low (slightly higher than 50 µg m^{-3}). Both sites seem to be good choices for sitting air quality monitoring stations (Figure 13).

The current location of the operative station in this district has relatively high *AR* value (49%), low *ANRmayor* value (12%) and quite high *IR* value (0.47), being a more representative location than that of the Escuelas Aguirre district, whose *AR*, *ANRmayor* and *IR* values were 33%, 42% and 0.17, respectively. In comparison with P27 and P9, this location have lower representativeness area (*AR*) but a wider zone of higher concentrations is covered (*ANRmayor* = 12%).

For the Plaza de Castilla district, we would need 3 stations at sites P20, P19 and P18 to cover all the studied domain, even also a 95% of the domain is covered by adding the representativeness area of only two sites such as P19 and P20. This result is quite different to the case of the Escuelas Aguirre, which would needs 4 stations to cover the whole domain. It is due to the high variability of the concentration field within the Escuelas Aguirre district respect to that obtained within the Plaza de Castilla domain.

The different features of both districts are highlighted when the plots of the *AR* vs concentration and *IR* vs concentration from each district are compared each other (Figure 14). For the Escuelas Aguirre district, there is a wide range (40-120 μg m^{-3}) of concentration values corresponding to *AR* values very similar (30-40%) and the IR maximum is less than 0.4 corresponding to concentration values of around 100 μgm^{-3}. However, for the Plaza de Castilla district, both plots show a sharp peak corresponding to concentrations of 45 μg m^{-3} approximately and with maxima of AR and IR close to 80% and 0.8, respectively. These differences are related to the different features of the concentration fields in both districts. The Plaza de Castilla concentration field is much smoother but with a local and small area of high concentrations in the square, while that of the Escuelas Aguirre district has stronger spatial variation of the pollutant concentration at very short scale with a relatively wide representative area for high (and also intermediate and low) concentration zones. It is much related to the fact that the Plaza de Castilla district has a more open urban structure with wider squares and streets allowing a better ventilation than the Escuelas Aguirre district.

Table 6. Site ranking in the Plaza de Castilla district sorted by *AR*.

Location	**Concentración (µg m^{-3})**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
9	50.4	77	23	0	0.75
27	52.8	74	19	7	0.75
20	47.8	71	29	0	0.63
22	47.6	71	29	0	0.62
3	47.4	70	30	0	0.61
24	44	60	40	0	0.41
32	42.4	55	45	0	0.33
29	56.4	55	15	31	0.53
15	57	53	14	33	0.52
26	57.7	52	13	35	0.50
23	58.9	50	12	38	0.48
31	58.9	50	12	38	0.48
1	59.1	49	12	39	0.47
2	62.4	39	10	51	0.35
12	62.6	38	9	52	0.34
13	63.9	36	9	56	0.31
8	64.9	34	8	58	0.30
28	66.2	33	7	61	0.28
25	67.6	30	6	64	0.26
19	70.1	27	5	68	0.23
21	73.8	22	3	75	0.18
14	77.5	18	2	79	0.16
4	78.9	17	2	81	0.15
5	80.4	16	2	82	0.14
10	80.6	16	2	82	0.14
16	81.8	15	2	83	0.14
11	81.9	15	2	83	0.14
30	85	13	1	86	0.12
6	85.4	13	1	86	0.12
18	105	5	0	95	0.08
17	117.2	2	0	98	0.07
7	138.5	0	0	100	0.06

Table 7. Site ranking in the Plaza de Castilla district sorted by *ANRmayor* in ascending order. Zero indicates that areas with the highest concentration are within the area of representation. The higher *ANRmayor*, more areas of high concentrations fall outside the area of representation.

Location	**Concentración ($\mu g\ m^{-3}$)**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
17	117.2	2	0	98	0.07
7	138.5	0	0	100	0.06
18	105	5	0	95	0.08
6	85.4	13	1	86	0.12
30	85	13	1	86	0.12
16	81.8	15	2	83	0.14
11	81.9	15	2	83	0.14
5	80.4	16	2	82	0.14
10	80.6	16	2	82	0.14
4	78.9	17	2	81	0.15
14	77.5	18	2	79	0.16
21	73.8	22	3	75	0.18
19	70.1	27	5	68	0.23
25	67.6	30	6	64	0.26
28	66.2	33	7	61	0.28
8	64.9	34	8	58	0.30
13	63.9	36	9	56	0.31
12	62.6	38	9	52	0.34
2	62.4	39	10	51	0.35
1	59.1	49	12	39	0.47
23	58.9	50	12	38	0.48
31	58.9	50	12	38	0.48
26	57.7	52	13	35	0.50
15	57	53	14	33	0.52
29	56.4	55	15	31	0.53
27	52.8	74	19	7	0.75
9	50.4	77	23	0	0.75
20	47.8	71	29	0	0.63
22	47.6	71	29	0	0.62
3	47.4	70	30	0	0.61
24	44	60	40	0	0.41
32	42.4	55	45	0	0.33

Table 8. Site ranking in the Plaza de Castilla district sorted by *IR*.

Location	**Concentración ($\mu g\ m^{-3}$)**	***AR* (%)**	***ANRmayor* (%)**	***ANRmenor* (%)**	***IR***
27	52.8	74	19	7	0.75
9	50.4	77	23	0	0.75
20	47.8	71	29	0	0.63
22	47.6	71	29	0	0.62
3	47.4	70	30	0	0.61
29	56.4	55	15	31	0.53
15	57	53	14	33	0.52
26	57.7	52	13	35	0.50
23	58.9	50	12	38	0.48
31	58.9	50	12	38	0.48
1	59.1	49	12	39	0.47
24	44	60	40	0	0.41
2	62.4	39	10	51	0.35
12	62.6	38	9	52	0.34
32	42.4	55	45	0	0.33
13	63.9	36	9	56	0.31
8	64.9	34	8	58	0.30
28	66.2	33	7	61	0.28
25	67.6	30	6	64	0.26
19	70.1	27	5	68	0.23
21	73.8	22	3	75	0.18
14	77.5	18	2	79	0.16
4	78.9	17	2	81	0.15
5	80.4	16	2	82	0.14
10	80.6	16	2	82	0.14
16	81.8	15	2	83	0.14
11	81.9	15	2	83	0.14
30	85	13	1	86	0.12
6	85.4	13	1	86	0.12
18	105	5	0	95	0.08
17	117.2	2	0	98	0.07
7	138.5	0	0	100	0.06

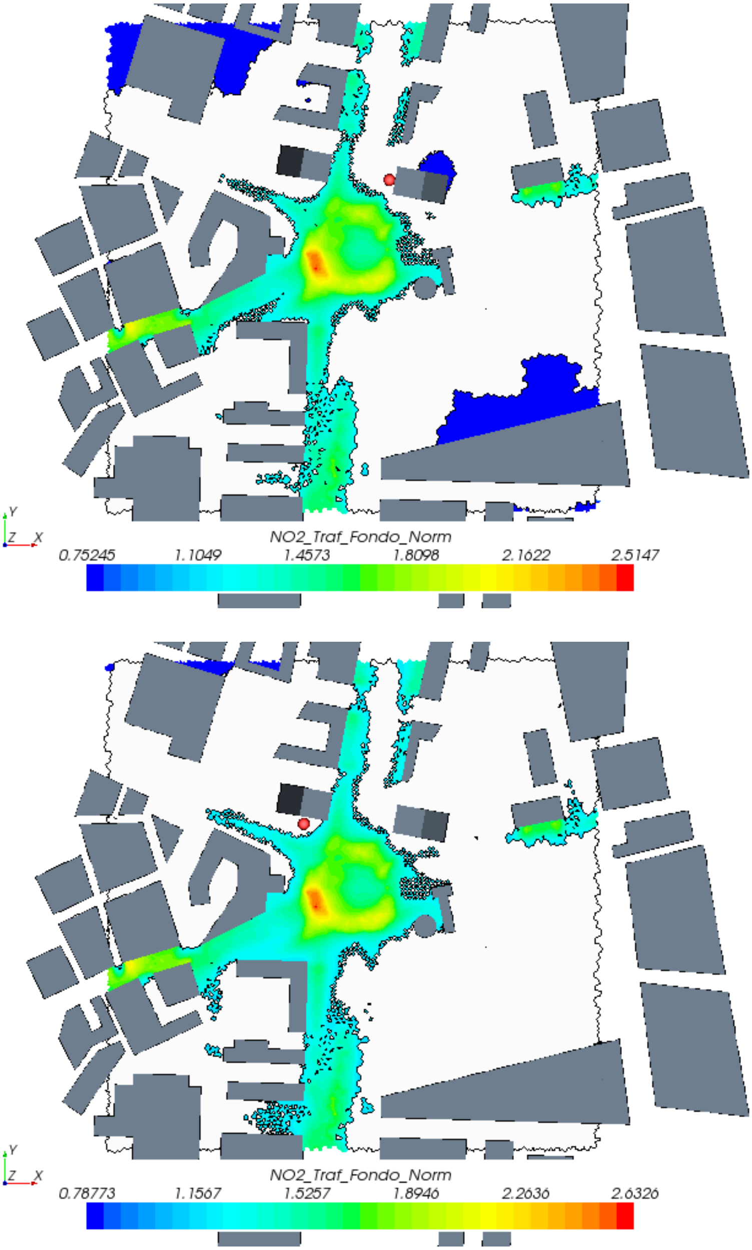

Figure 13. SR areas (white color) corresponding to the sites P27 (upper) and P9 (lower right) for the Plaza de Castilla district. Red dots show the location of the candidate sites. Concentrations are normalized respect to the estimated averaged concentration at the candidate sites.

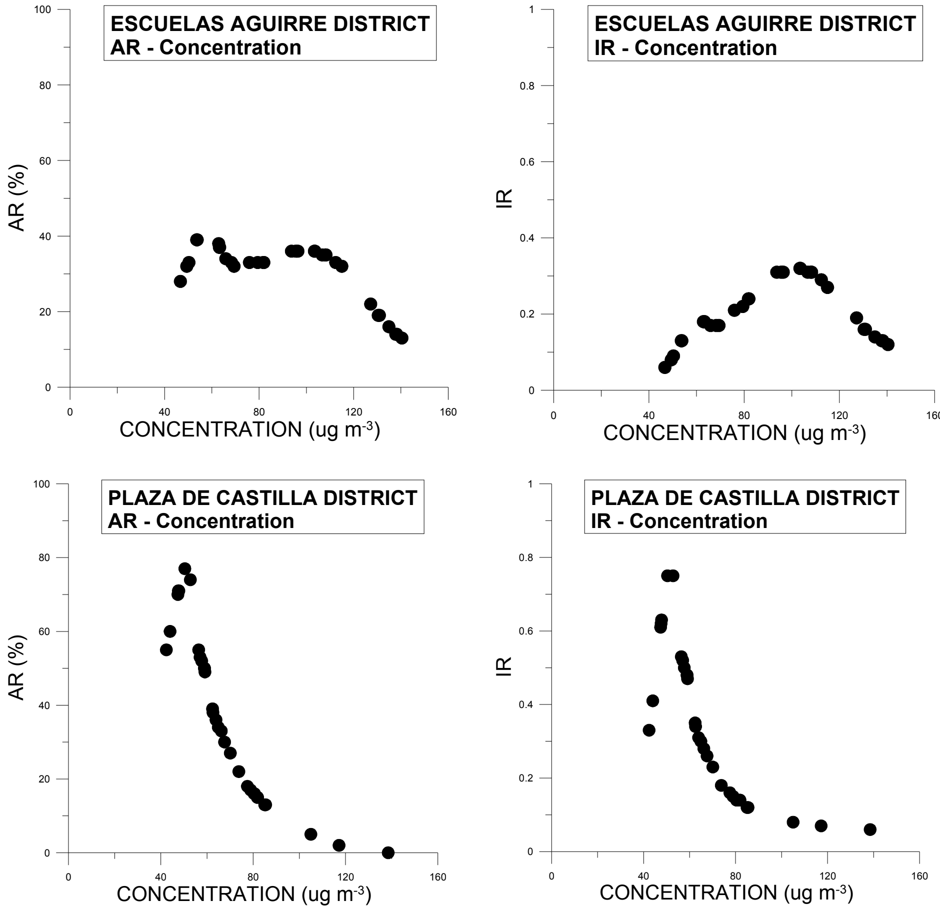

Figure 14. Plots of representativeness area (*AR*) (left) and representativeness index (*IR*) (right) versus pollutant concentrations at the candidate sites at the Escuelas Aguirre (above) and Plaza de Castilla (below) districts.

4. Summary and conclusions

This paper shows a methodology to estimate the spatial representativeness of air pollution monitoring sites in urban zones based on high resolution maps of air pollution computed by using Computational Fluid Dynamics (CFD) modelling tools. For several meteorological conditions, traffic-emitted pollutant dispersion is simulated taking into account the effect of the buildings on air flow and pollutant dispersion using steady state CFD-RANS approach. The maps of average pollutant concentrations over periods of several months are computed by using the combination of simulated scenarios. In this study, NO_2 dispersion within two urban districts of Madrid City (Escuelas Aguirre and Plaza de Castilla) was simulated. Each district has an operative air quality station. For each district, 32 candidates monitoring sites (including the present sites of the air quality stations) have been distributed throughout the domains and the spatial representativeness of every site was estimated. The spatial representativeness area of a monitoring site is defined as the area where concentrations are within an interval of ±20% of the concentration at the monitoring site. The resulted representativeness areas has been discussed by analysing the values of several proposed indexes such as the ratio *AR* between the representativeness area and the whole domain area or the representativeness index (*IR*) in order to determine what is the most representative site in every studied district. *IR* is designed to favour high values of *AR* avoiding to leave out large areas with high pollutant concentrations.

There are very important differences between the spatial representativeness of the candidate sites of both studied districts. The sites of the Escuelas Aguirre district have smaller representativeness areas ($AR < 40\%$ and $IR < 0.4$) than those of the Plaza de Castilla ($AR < 80\%$ and $IR < 0.8$). The operative air quality station of the Escuelas Aguirre district is less representative than the station of the Plaza de Castilla district. On the other hand, four stations would have to be deployed to cover the whole domain of the Escuelas Aguirre district, while three stations are enough for the Plaza de Castilla district. It is due to the high variability of the concentration field within the Escuelas Aguirre district respect to that obtained within the Plaza de Castilla domain, whose pollutant concentration field is smoother. It is due to the good ventilation in the Plaza de Castilla district because of the urban structure is more open with wider squares and streets than in the Escuelas Aguirre district.

Acknowlegdements

This work has been financially supported by the Spanish Ministry of Agriculture, Food and Environment.

References

BLANCHARD, C.L., CARR, E.L., COLLINS, J.F., SMITH, T.B., LEHRMAN, D.E., MICHAELS, H.M. (1999). Spatial representativeness and scales of transport during the 1995 integrated monitoring study in California's San Joaquin Valley. *Atmospheric Environment*, **33**, 4775-4786.

CASTELL-BALAGUER, N., DENBY, B. (2012). A survey to elicit expert opinion on the spatial representativeness of ground based monitoring data. A FAIRMODE activity for WG2-SG1: Combining models and monitoring. FARIMODE, Joint Research Center.

GALÁN MADRUGA D., FERNÁNDEZ PATIER R., DIAZ RAMIRO E., HERCE GARRALETA M. (2001). Study of the superficial ozone concentrations in the atmosphere of Comunidad de Madrid using passive samplers. *Rev Salud Ambient*, **1**, 20–9.

GEIGER J. AND 14 AUTHORS MORE (2014). Assessment on siting criteria, classification and representativeness of air quality monitoring stations. JRC-AQUILA Position Paper.

JANSSEN, S., DUMONT, G., FIERENS, F., DEUTSCH, F., MAIHEU, B., CELIS, D., AND MENSINK, C. (2012). Land use to characterize spatial representativeness of air quality monitoring stations and its relevance for model validation. *Atmospheric Environment*, **59**, 492-500.

KRAYENHOFF E.S., SANTIAGO J.L., MARTILLI A., CHRISTEN A., OKE T.R. (2015). Parametrization of drag and turbulence for urban neighbourhoods with trees. *Boundary-Layer Meteorology*, **156**, 157-189.

MARTIN, F., FILENI, L., PALOMINO, I., VIVANCO, M. G., GARRIDO, J. L. (2014). Spatial representativeness of rural background monitoring stations in Spain. *Atmospheric Pollution Research*, **5**, 779-788.

OTT, W., ELIASSEN, R. (1973). A survey technique for determining the representativeness of urban air monitoring stations with respect to carbon monoxide. *Journal of the Air Pollution Control Association*, **23**, 685-690.

PALACIOS M. Influencia del tráfico rodado en la generación de la contaminación atmosférica.Aplicación de un modelo de dispersión al área de influencia de la Comunidad de Madrid (in Spanish). Doctoral Thesis; 2001.

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. *Atmospheric Environment*, **44**, 2089-2097.

SANTIAGO J.L., MARTÍN F., MARTILLI A. (2011a) Representativeness of urban monitoring stations. 14th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Kos (Greece) 2-6 October.

SANTIAGO J.L., MARTÍN F., MARTILLI A. (2011b). Informe de 2011 de la Tarea 1.d: Estudio de representatividad espacial de estaciones urbanas de calidad del aire. Encomienda de Gestión 2010-2014 entre el Ministerio de Medio Ambiente y el CIEMAT para trabajos en materia de calidad del aire, energía y evaluación ambiental. Actuación 2. Tarea 1.d. CIEMAT. Ref: 31/2011. (in Spanish)

SANTIAGO J.L., MARTÍN F., MARTILLI A. (2013) A computational fluid dynamic modelling approach to assess the representativeness of urban monitoring stations. *Science of the Total Environment*, **454-455**, 61-72.

SANTIAGO J.L., MARTILLI A., MARTÍN F. (2014). Estimating the impact of street vegetation on air quality: a simple case with different types and position of vegetation. In: 16th International Conference on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, Varna (Bulgaria), 8-11 September 2014.

SPANGL W. SCHNEIDER J. MOOSMANN L., NAGL C. (2007). Representativeness and Classification of Air Quality Monitoring Stations. Final Report. Umweltbundesamt Wien.

VARDOULAKIS, S., GONZALEZ-FLESCA, N., FISHER, B. E. A., PERICLEOUS, K. (2005). Spatial variability of air pollution in the vincinity of a permanent monitoring station in central Paris. *Atmospheric Environment*, **39**, 2725-2736.