Research Report 236

Traffic-Related Air Pollution and Birth Weight: The Roles of Noise, Placental Function, Green Space, Physical Activity, and Socioeconomic Status (FRONTIER)

Payam Dadvand and Jordi Sunyer et al.

Additional Materials 2: Air Pollution Exposure Modelling of a Pregnant Group Control at Barcelona Metropolitan Area in 2018–2021

Additional Materials 2 was reviewed by the HEI Review Committee. It has not been fully edited or formatted by HEI.

Correspondence may be addressed to Dr. Payam Dadvand, Barcelona Institute for Global Health (ISGlobal), Doctor Aiguader 88, 08003 Barcelona, Spain; email: *payam.dadvand@isglobal.org*.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83998101 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and may not necessarily reflect the views of the Agency; thus, no official endorsement by it should be inferred. It also has not been reviewed by private-party institutions, including those that support the Health Effects Institute, and may not reflect the views or policies of these parties; thus, no endorsement by them should be inferred.

BARCELONA REGIONAL
AGÈNCIA DESENVOLUPAMENT URBÀ

AIR POLLUTION EXPOSURE MODELLING OF A PREGNANT GROUP CONTROL AT BARCELONA METROPOLITAN AREA IN 2018-2021

This document is aimed to present the main aspects of the modelling of the air pollution exposure of a pregnant group control in the Barcelona Metropolitan Area (AMB in Spanish) in 2018-2022, as such as the model implementation methodology and the main evaluation results of the air quality model and the commute data processing methodologies by pregnant women, in addition the main results of the modelling process for the control group.

The main goal of the modelling module is to provide an air quality model to evaluate the hourly exposure of a certain control group of pregnant to the pollutants NO_2, $PM_{2,5}$ and BC in the period 2018 - 2021 at the AMB.

1. AIR QUALITY MODEL IMPLEMENTATION

1.1. Methodology

Barcelona Regional implemented an air quality model for 2017 at nine municipalities of the Barcelona Metropolitan Area (AMB): Barcelona, Badalona, Sant Adrià de Besòs, Santa Coloma de Gramenet, l´Hospitalet de Llobregat, Esplugues de Llobregat, Sant Just Desvern, Prat de Llobregat and Cornellà de Llobregat (Barcelona Regional, 2020).

With this work as a starting point, the air quality model for 2017 has been extended to all other AMB municipalities (to cover 36 municipalities in total) and upgraded to the period between 2018 and 2021. The main work has been to update and expand the emission inventory to the entire modelling area to provide the most accurate air quality data for this present study.

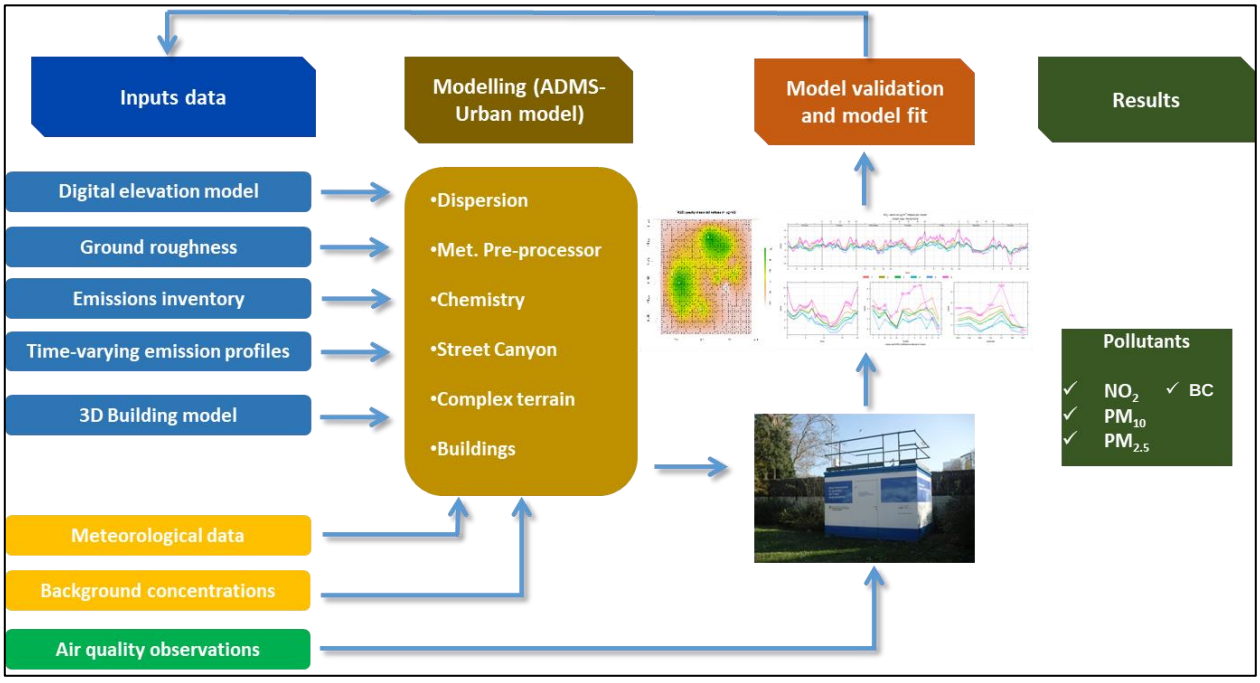
The dispersion model used is ADMS-Urban, from CERC (Cambridge Environmental Research Consultants) which is a Gaussian-type model designed for urban environments. The model calculates the pollutant concentration from various emissions sources, including vehicular traffic, household activities, industrial plants, airports, harbour, factories and commercial sectors, among others. It is also able to calculate hourly concentrations, daily and annual averages, as well as number of exceedances over legal limits.

The ADMS Urban model uses different algorithms to assess the chemical transformation and dispersion of pollutants. It calculates the effects of turbulence along the ground and from surrounding buildings in urban areas. The main characteristics of the used model are:

- It is a specific model for the analysis of the emission of pollutants in urban and metropolitan areas, at road level.
- It is based on a Gaussian model (mathematical Gaussian expressions).
- It incorporates a model of hourly meteorological pre-processes of fluxes and turbulence derived from the study of the ground of the whole modelled area. The module used is called FLOWSTAR[1], which incorporates a specific Danish OSPM[2] model, to evaluate the effect of street canyon[3] or urban canyon (Berkowicz, 2000).
- It considers the chemical reactions between several species present in the atmosphere, from different types of emission sources. It also considers the photochemical reactions derived from incident solar radiation, according to the GRS[4] mechanism (Venkatram et al, 1994). For sulphates chemistry, it uses the EMEP[5] model of mechanism (Tsyro, 2001).
- Intelligent virtual receptors system (smart grid) allow to spread a regular grid of points distributed within the territory. This system permits an automated assignment of denser measurement points along roads, improving detailed results for such points.

Regarding the input of data, there are structural data that are kept constant along at the study (same data for different years), such as digital elevation model, ground roughness and 3D Building model. However, there are other input year-dependent data such as the emission inventory, meteorological data and background pollutant concentrations. The next flowchart shows the main blocks of the followed methodology: input data preparation, modelling, model validation, model fit and exploitation of results.

Figure 1: Air quality model implemented at Barcelona Regional.



Source: Barcelona Regional.

[1] Model of airlfow over complex terrain.
[2] Operational Street Pollution Model.
[3] Created when a street is surrounded by tall buildings, blocking and difficulting dispersion of air pollutants.
[4] Generic Reaction Set
[5] European Monitoring and Evaluation Programme

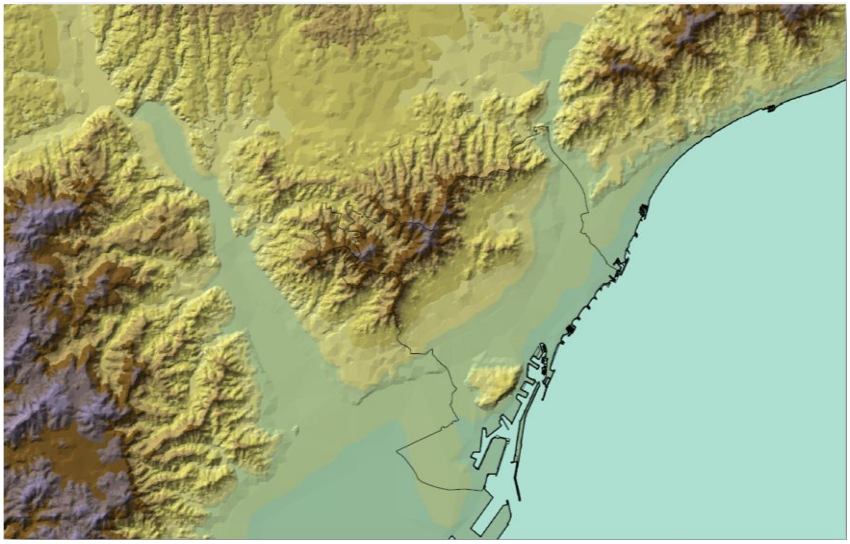
1.1.1. Structural data

This chapter shows all the structural data needed to carry out a correct air quality modelling.

1.1.1.1. Digital elevation model

A topographic data base from the “l'Institut Cartogràfic i Geològic de Catalunya” (ICGC, 2017) has been used in order to obtain level curves and altimetry elevations. In this way, a representative and uniform spatial set of points can be represented. Such representation can be seen in the figure below.

Figure 2. Terrain elevation map of the study area (mainly Barcelona Metropolitan Area).



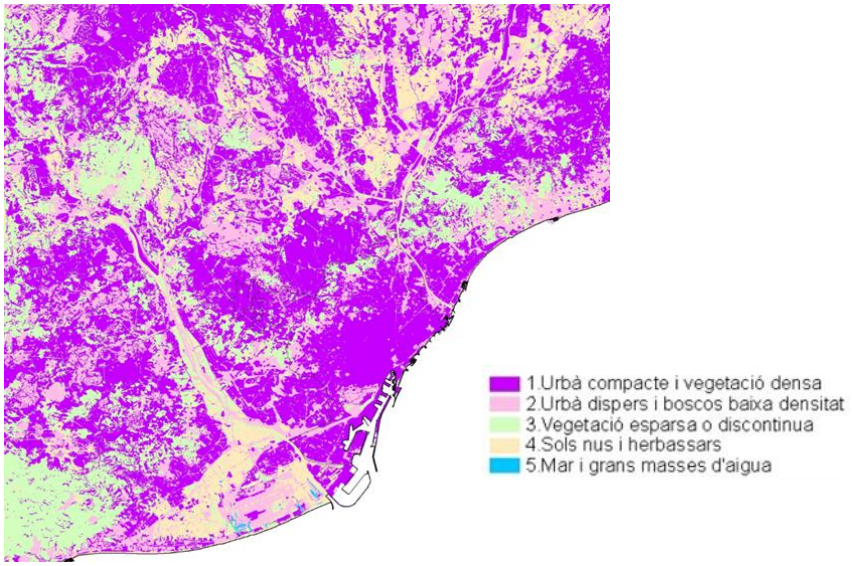
Source: Barcelona Regional with ICGC data, 2017.

1.1.1.2. Roughness

Roughness is a function of the spatial density, orientation and height of the obstacles encountered by the wind. Roughness plays an important role in how airflow interacts with the urban landscape (Salizzoni et al., 2011). Spatially-variable roughness creates horizontal variations in turbulence and the local mean flow, which can affect pollutant dispersion (Barnes et al., 2014). The roughness factor has been calculated by means of Land Cover Map of Catalonia (DVPDT[6]), which is shown in the figure below.

[6] Departament de la Vicepresidència i de Polítiques Digitals i Territori.

Figure 3. Land Cover Map of the model domain (mainly Barcelona Metropolitan Area).

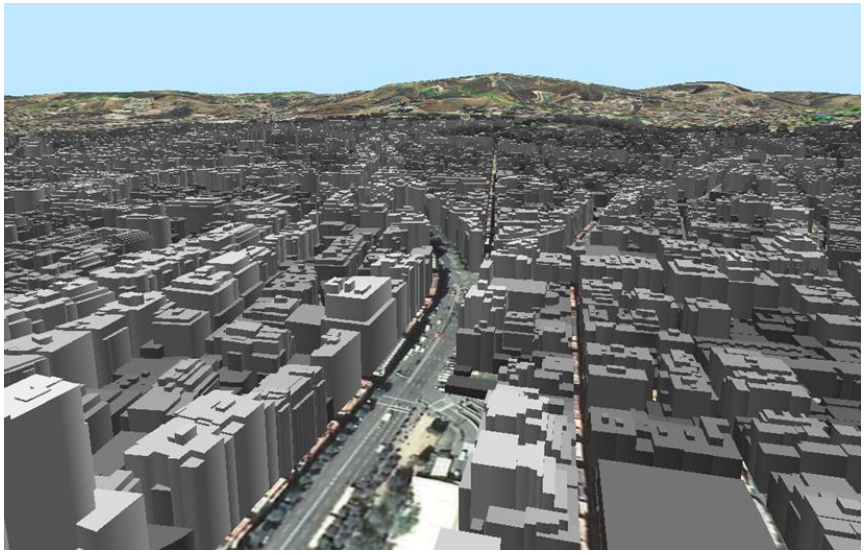


Source: Barcelona Regional from DVPDT.

1.1.1.3. 3D Building model

A three-dimensional model of buildings has been created to calculate the heights of buildings and street width from ICGC (2017), in order to include the street canyon effect (explained in the next lines). A partial representation of this 3D model is shown in the following figure.

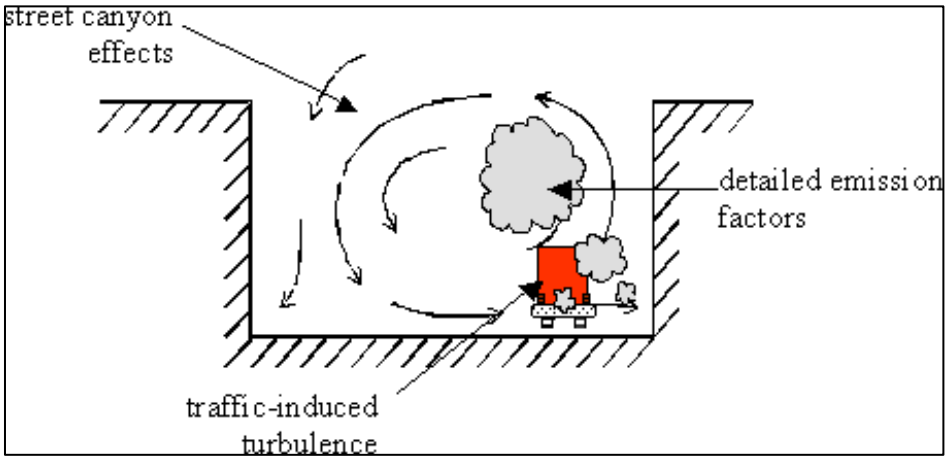
Figure 4. Three-dimensional Model of Barcelona buildings.



Source: Barcelona Regional from ICGC data (2017).

Detailed building model is needed to calculate the *street canyons effect*. This phenomenon occurs when tall buildings lie adjacent to roads with heavy traffic. In these geometrical conditions, atmospheric dispersion become poor. Therefore, pollutant concentrations tend to be rather uniform and significantly higher than those outside the urban areas affected by the canyon. The following figure represents such phenomenon in a urban street.

Figure 5. Street canyon effect representation.



Source: CERC, 2019.

1.1.2. Meteorological data

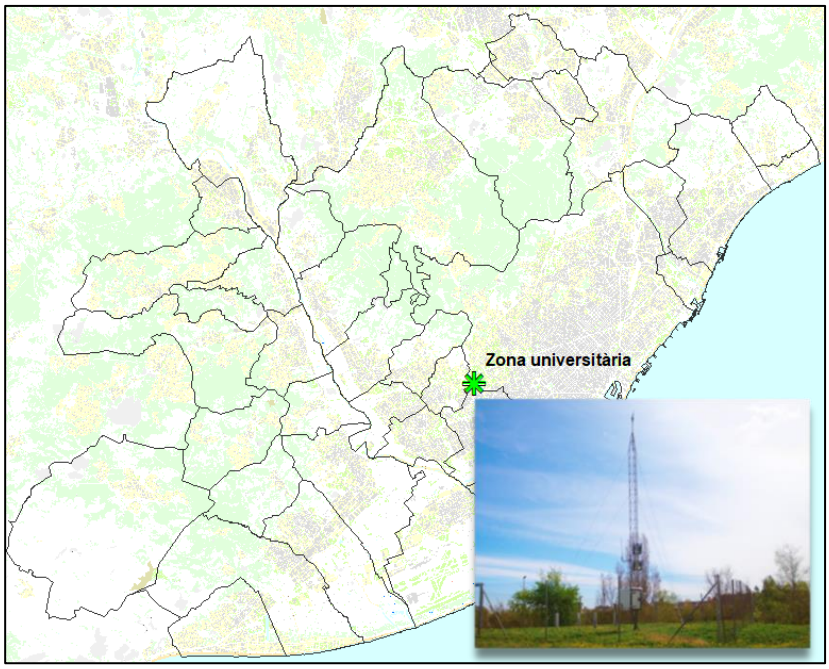
The meteorological data that have been used depends on the analysed period, which is between 2018 and 2020. The meteorological data input to the model is: wind speed and direction, air temperature, cloud cover level and solar radiation, relative humidity, precipitation rate and Planetary Boundary Layer Height (PBLH[7]).

Meteorological data has been obtained from “Zona Universitària” (located in Barcelona) meteorological station, from “Servei Meteorològic de Catalunya” info (2018 to 2021). On the other hand, the PBLH data has been taken from the Air Quality Forecast Barcelona model (meteorological data have been provided by Meteosim).

The following figure shows the location of this meteorological station on the map, as well as a picture of it.

[7] The PBL is the layer above soil and/or vegetation where vertical transports by turbulence play a dominant role in transfers of momentum, heat, moisture, CO_2, and other gases. The height of the PBL ranges from **100 to 3000 m** and varies with time, location, and weather conditions.

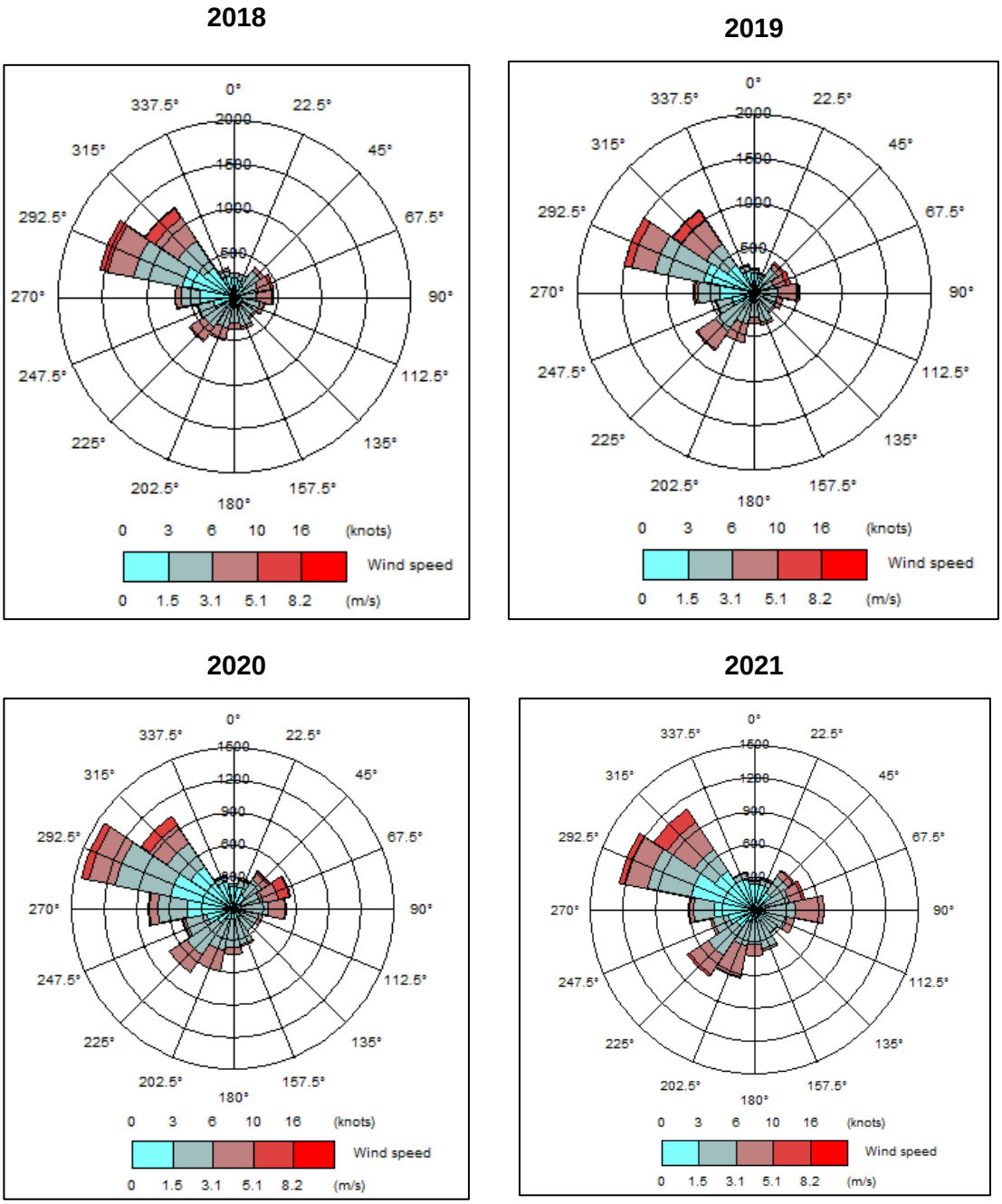
Figure 6: Zona Universitària meteorological station (Barcelona).



Sources: map: Barcelona Regional, picture: SMC (Servei Meteorològic de Catalunya)

Wind roses for period between 2018 and 2021 are shown in the following figure. They all show a predominant direction from NW[8] (292,5º – 315º) for all years.

Figure 7: Wind rose at Zona Universitària meteo station, from 2018 to 2021.



Source: Barcelona Regional with data from SMC.

[8] North-West

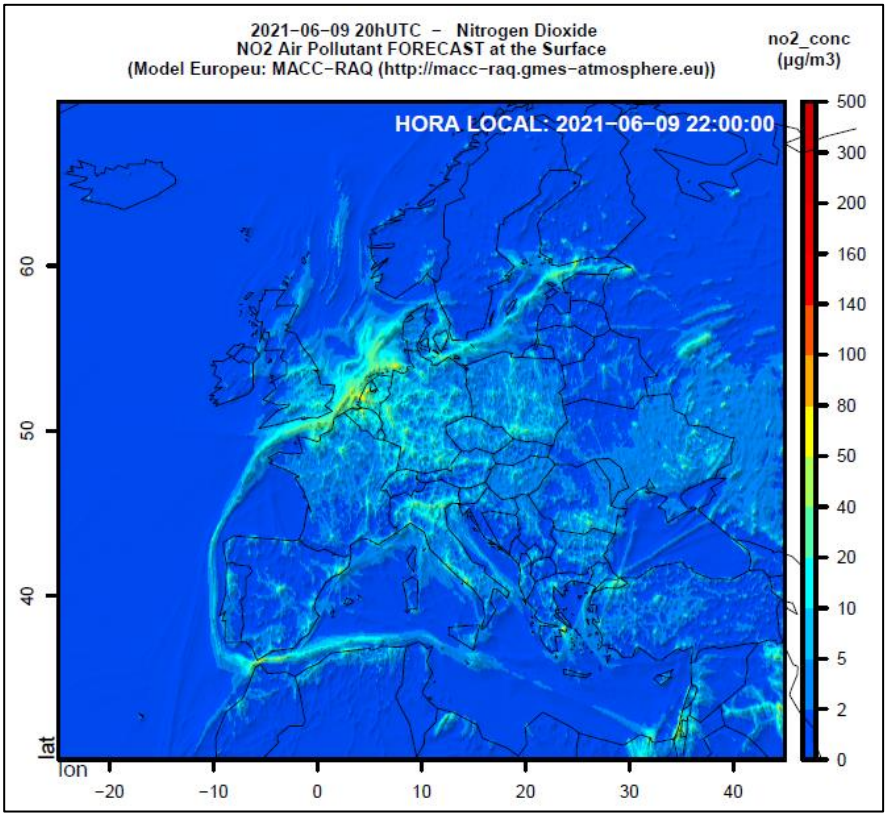
1.1.3. Background pollution levels

The pollution contribution from the contours of the simulation domain, also known as background pollution, represents the pollution coming from outside of the considered domain of the emissions inventory.

The background pollution can have different origin such as the regional accumulation of pollutants and emissions transport from exterior sources. It can also come from other latitudes such as large industries or cities, transport of Saharan dust, ashes from forest fires or volcanic eruptions, among others.

In order to incorporate this information into the model, data from the main European air quality project have been used: Copernicus Atmosphere Monitoring Service Information. Background pollution data has been extracted from the CAMS *reanalyses* ensemble model for 2018. In relation to 2019, 2020 and 2021, CAMS ensemble f*orecast* model has been used.

Figure 8: Europe air quality forecast model (CAMS Ensemble model).



Source: Generated using Copernicus Atmosphere Monitoring Service Information.

1.1.4. Emissions Inventory

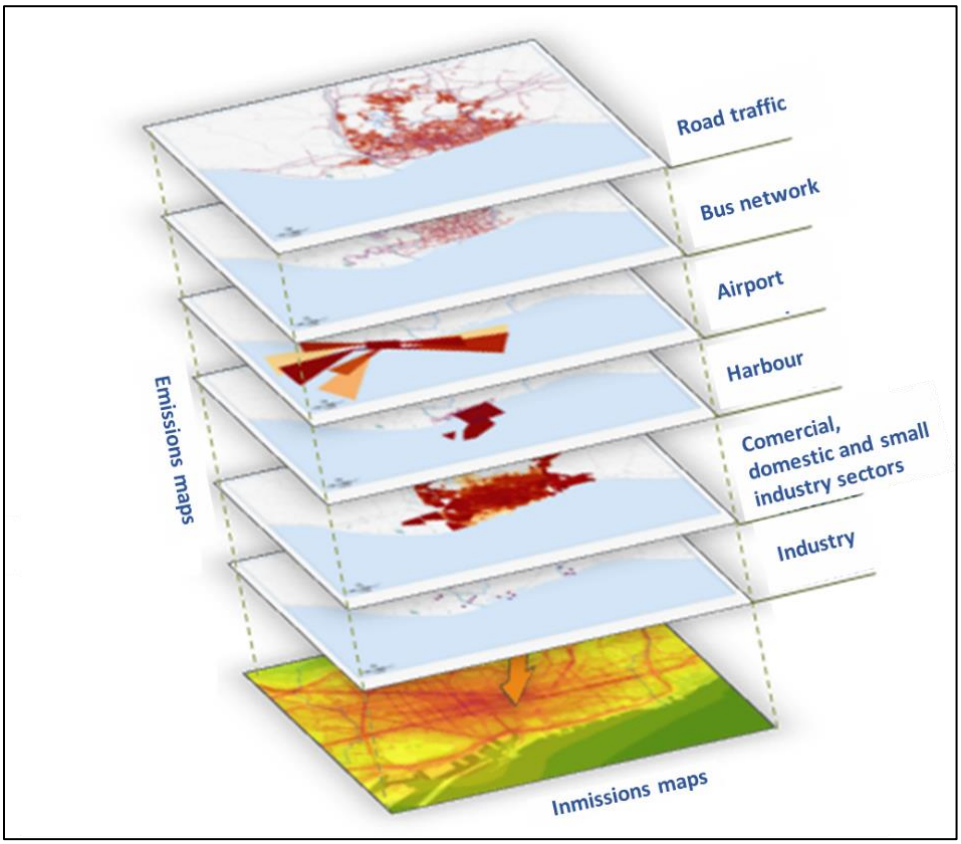
Related to the emissions inventory, EMEP/EEA2019 (EEA, 2019) methodologies have been used. For diffuse emissions (domestic, small industry, commercial and services sectors) a “top-down” approach has been utilized, based on fuel consumption from CORES[9] and ICAEN[10]. The fuel consumption has been spatially distributed according to the municipal cadastre (ICGC).

A "bottom-up" approach has been used for the road transport emissions, harbour and large industry, based on detailed data from these activities (Port de Barcelona, AENA[11], DVPDT[12] and EPER-PRTR[13]).

In relation to the airport emissions, an extrapolation from the Barcelona emissions inventory has been carried out (from the 2017 version). Dust emissions from parks and construction works (such as Plaça de les Glòries and La Sagrera) have been included also using EMEP/EEA methodologies (from 2019).

The following figure shows the main emission layers included in the emissions inventory, and how these interactions contribute to obtain an emissions map.

Figure 9: sectorial emissions layers.



[9] Corporación de Reservas Estratégicas de Productos Petrolíferos.
[10] Institut Català d’Energia.
[11] Aeropuertos Españoles y Navegación Aérea.
[12] Departament de la Vicepresidència i de Polítiques Digitals i Territori.
[13] EPER: European Pollutant Emission Register. PRTR: Pollutant Release and Transfer Registers.

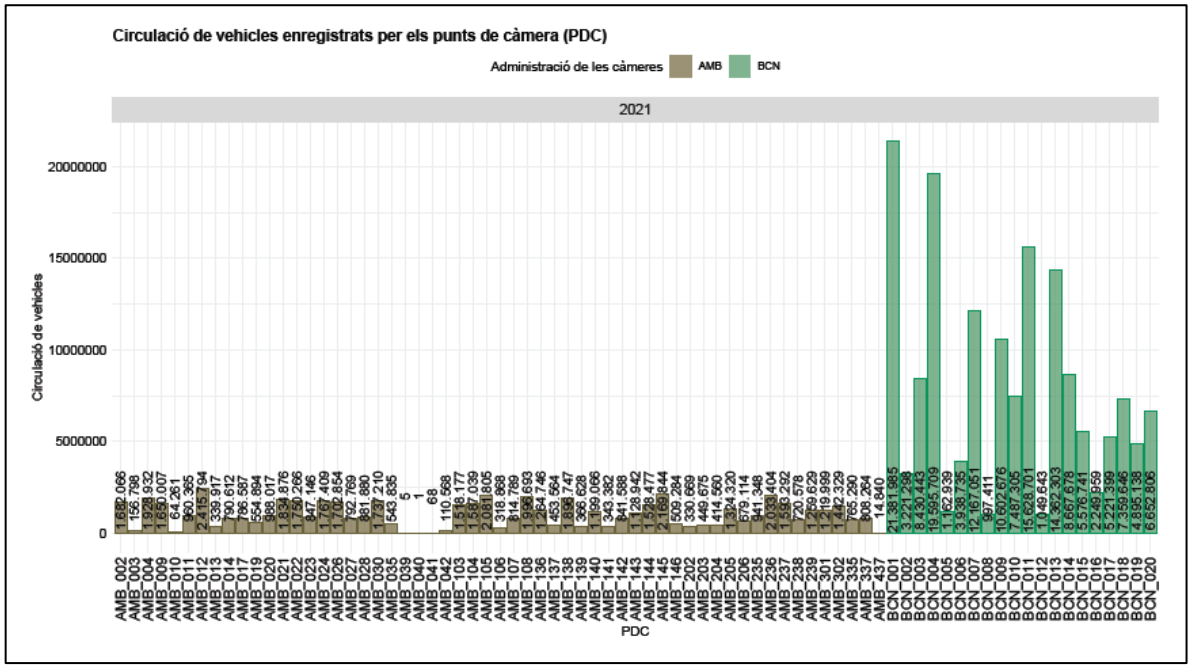
Source: Barcelona Regional

Regarding the road transport, it is important to mention the role of the LEZ control cameras. Due to the implementation of this low-emission zone (LEZ) within the Barcelona Metropolitan Area in 2020, there were installed several control cameras among AMB urban areas. Currently, there are about 80 cameras (20 in Barcelona) that have made possible to obtain, since 2020, real data from vehicle plates. The information acquired from such plates include, by means of the DGT[14] data base:

- Vehicle registration date
- Registration municipality of the vehicle
- Vehicle type
- Technical specs of the vehicle (size, age, energy type for traction, EURO standard…).

The following figures show, as an example, the number of vehicle registrations during 2021, for all the cameras installed within the AMB.

Figure 10. Number of vehicle registrations during 2021, for all the cameras installed within the AMB.

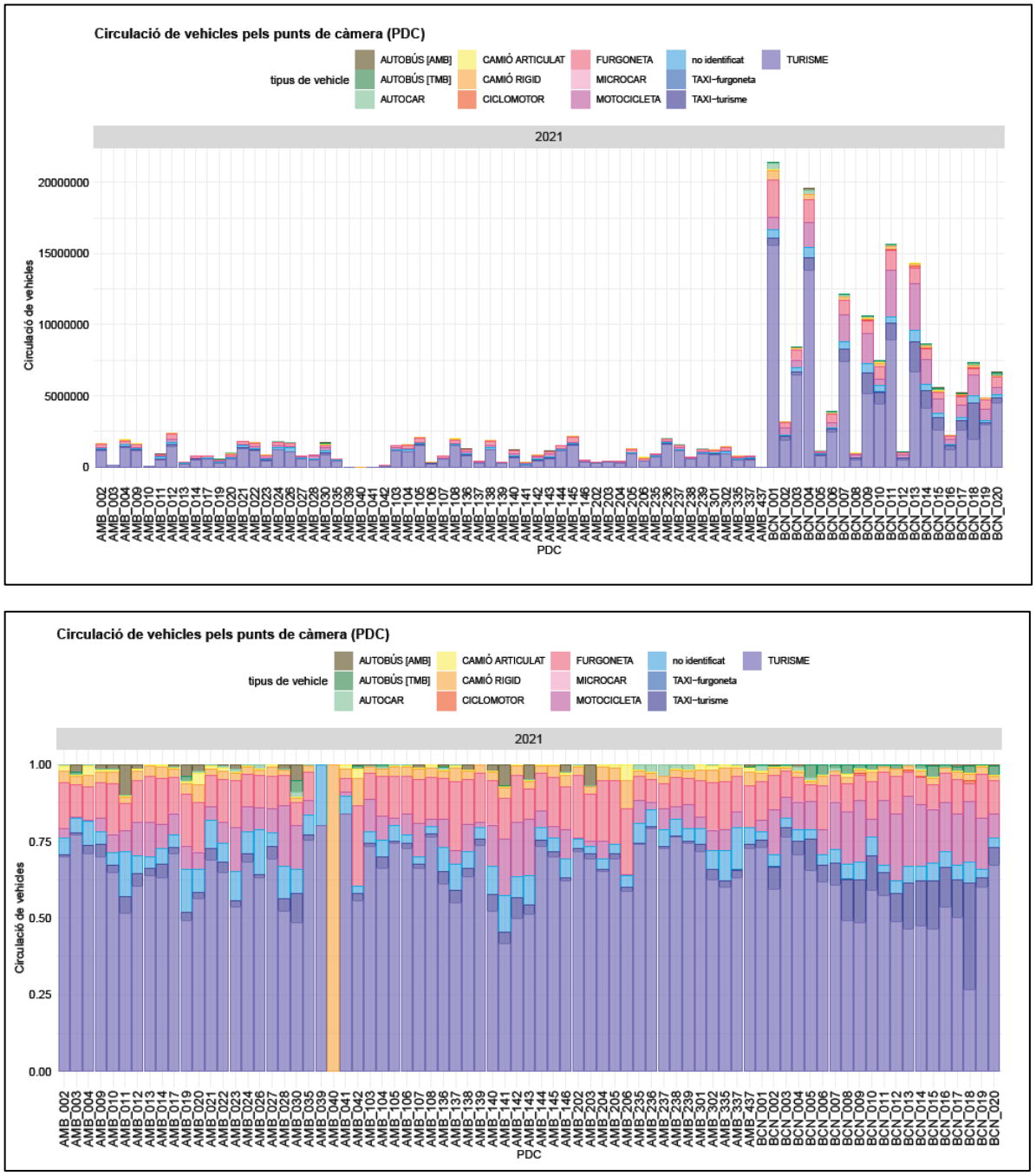


Source: Barcelona Regional with data (whole 2021) from l'Ajuntament de Barcelona, AMB. PDC: camera id.

[14] Dirección General de Tráfico.

In fact, such observations can be aggregated on vehicle type. They are show in the figure below (absolute observations and their distribution).

Figure 11. Vehicle type from camera observations (cameras installed within AMB). Above, in absolute value. Below, as percentatge of total observations.



Source: Barcelona Regional with data (whole 2021) from Ajuntament de Barcelona, AMB and DGT. PDC: camera id.

The vehicle data can also be aggregated even in more detail. Such level of detail has been used to capture a realistic and detailed fleet characterization that lead to a more accurate emissions estimation. From the model point of view, the emission profiles have played an important role (to

model time-variation emissions of road transport). This fleet characterization has been extended to Barcelona and other four municipalities more (that also have control cameras): l'Hospitalet de Llobregat, Cornellà de Llobregat, Esplugues de Llobregat and Sant Adrià de Besòs.

This realistic characterization of the vehicle fleet composition has been especially important during 2020, when lockdown periods affected not only the traffic volume but also the vehicle fleet composition.

In order to model the particle matter properly (PM_{10}, $PM_{2,5}$ and Black Carbon), the road transport emission factor has been split into source of emissions (exhaust emission and non-exhaust emission). The effect of re-emission has also been included, known as resuspension. The distribution of each type of emissions is applied to the total emission factor of each road section (which depends on its velocity and fleet distribution). Such distribution is shown in the table below.

Table 1. Distribution of road transport emissions by type of emission source on the total particle matter and black carbon [%].

Type of emission	PM_{10}	$PM_{2,5}$	BC
Road Exhaust	8,3%	13,7%	28,5%
Road Non-exhaust	34%	29,4%	7,8%
Road resuspension	57,7%	56,9%	63,7%
Total vehicle emission	**100%**	**100%**	**100%**

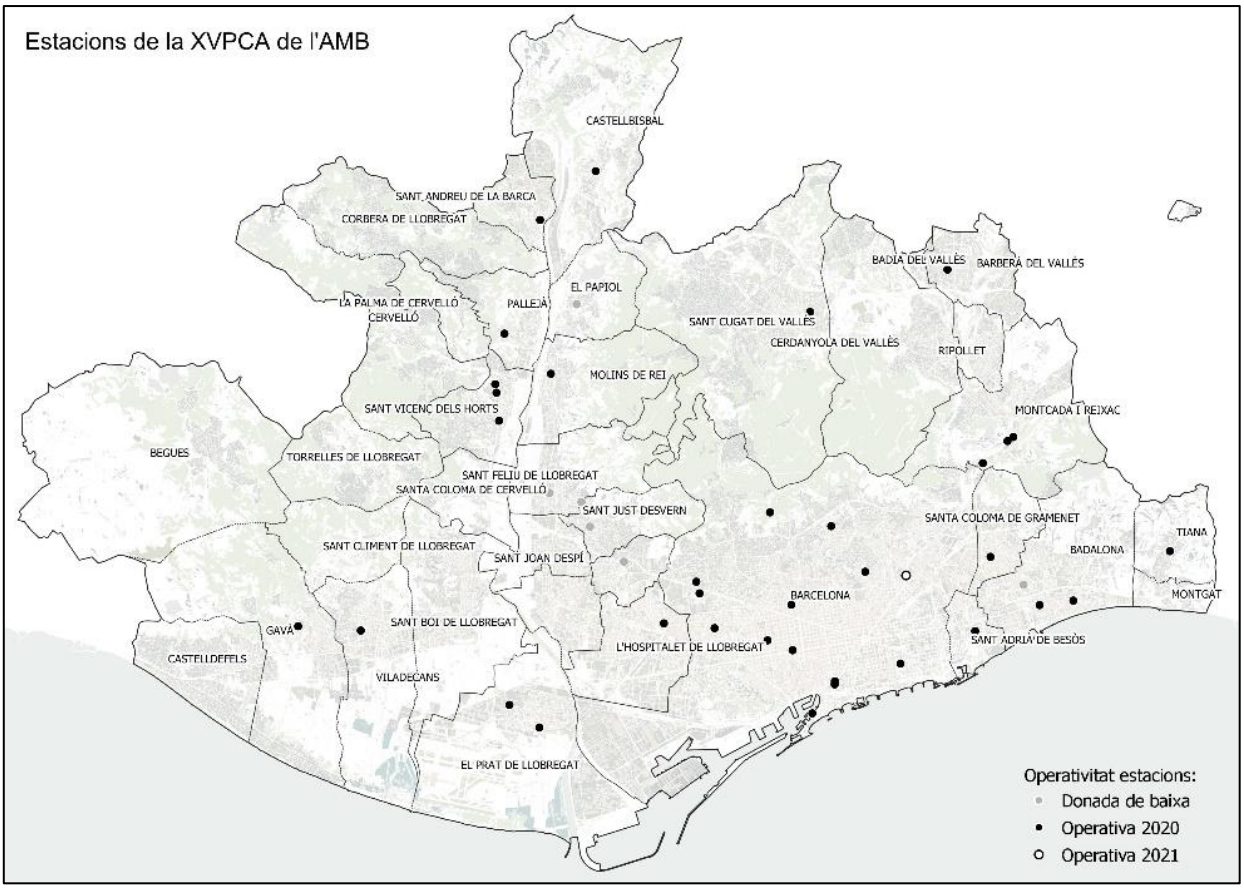
Source: Barcelona Regional

1.1.5. Model validation and adjustment

Once the input data has been introduced to the ADMS Urban model, the calculation of NO_2, $PM_{2,5}$ and BC concentration levels has been carried out. The concentration has been calculated at several locations of XVPCA[15] stations, in order to compare the yielded results from the model to the real values measured by the air quality stations. The closer these values are, the better is the air quality modelling. The XVPCA stations located throughout the AMB (and used to validate the model) are shown in the following figure.

[15] Xarxa de Vigilància i Previsió de la Contaminació Atmosfèrica de Catalunya: Catalan air quality monitoring stations network.

Figure 12: XVPCA network distribution (only in AMB) for 2021.



Source: Barcelona Regional.

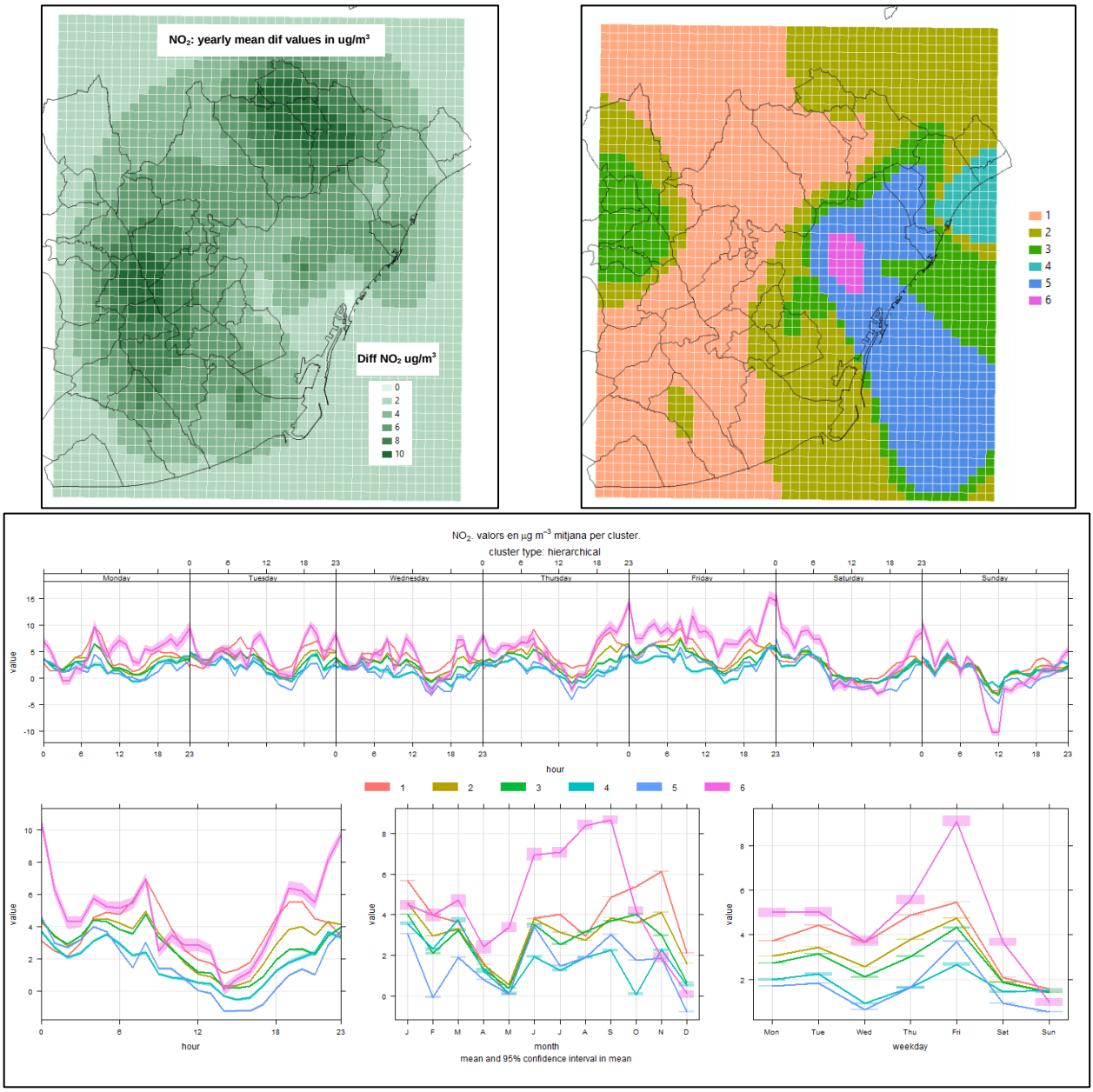
As said before, model results have been evaluated at the considered locations of XVPCA stations, just to evaluate the model performance at such emplacements. Then, an adjustment of the model (known as calibration) has been performed to fit the model results to the data measured by each XVPCA station.

Bias-adjustment techniques have been used to fit the model data to the observations, through hourly bias spatial interpolation and generation of hourly bias maps. After that, data are processed by cluster analysis techniques in order to obtain a mean cluster map. Each bias map by pollutant is associated to a cluster map class and an hourly-varying profile.

Finally, these layers of pollutants have been introduced into the model as a new emission sector, using a certain cluster time profile. The emissions are multiplied by a correction factor to convert emissions (tons/year) to immission ($\mu g/m^3$). Then, the model runs and evaluates iteratively until a proper correction factor is found (based of error tolerances).

The following figure shows, for 2019, an example of the bias interpolation map, the cluster map and temporal profiles for NO_2.

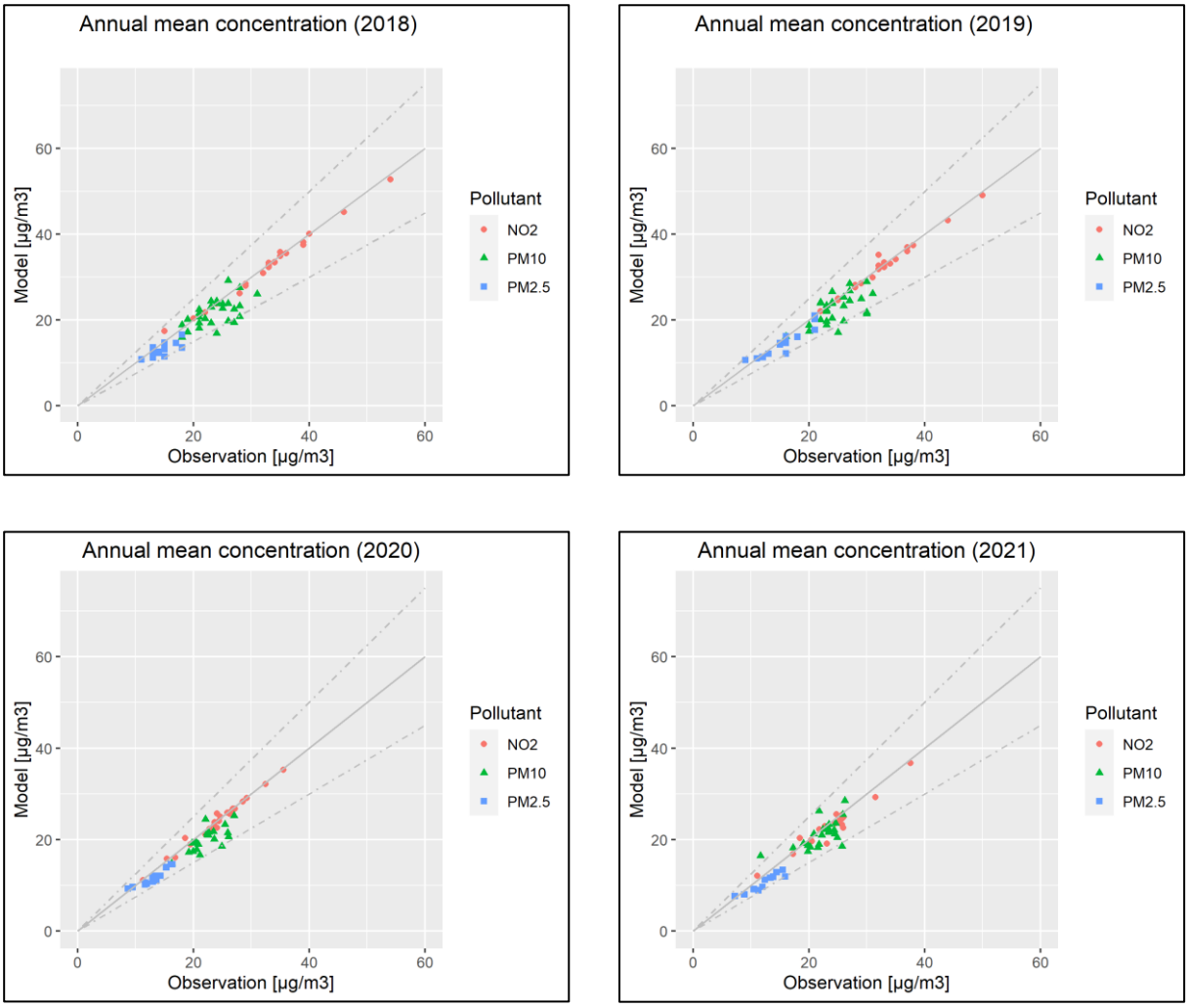
Figure 13: Example of bias and cluster maps for NO_2 (above). Temporal cluster NO_2 profiles for 2019 (below).



Source: Barcelona Regional.

The following graphics and tables show some of the main results of the model evaluation, such as a scatterplot of annual mean concentrations for pollutants NO_2, PM_{10} and $PM_{2,5}$ and for 2018 to 2021.

Figure 14. Scatterplots of the annual mean concentration (from 2018 to 2021 from all stations) for NO_2 (red), PM_{10} (green) and $PM_{2,5}$ (blue). Solid middle line is the 1:1 representation (perfect fit), whereas the dotted upper and lower lines represent 1,25:1 and 0,75:1 respectively.



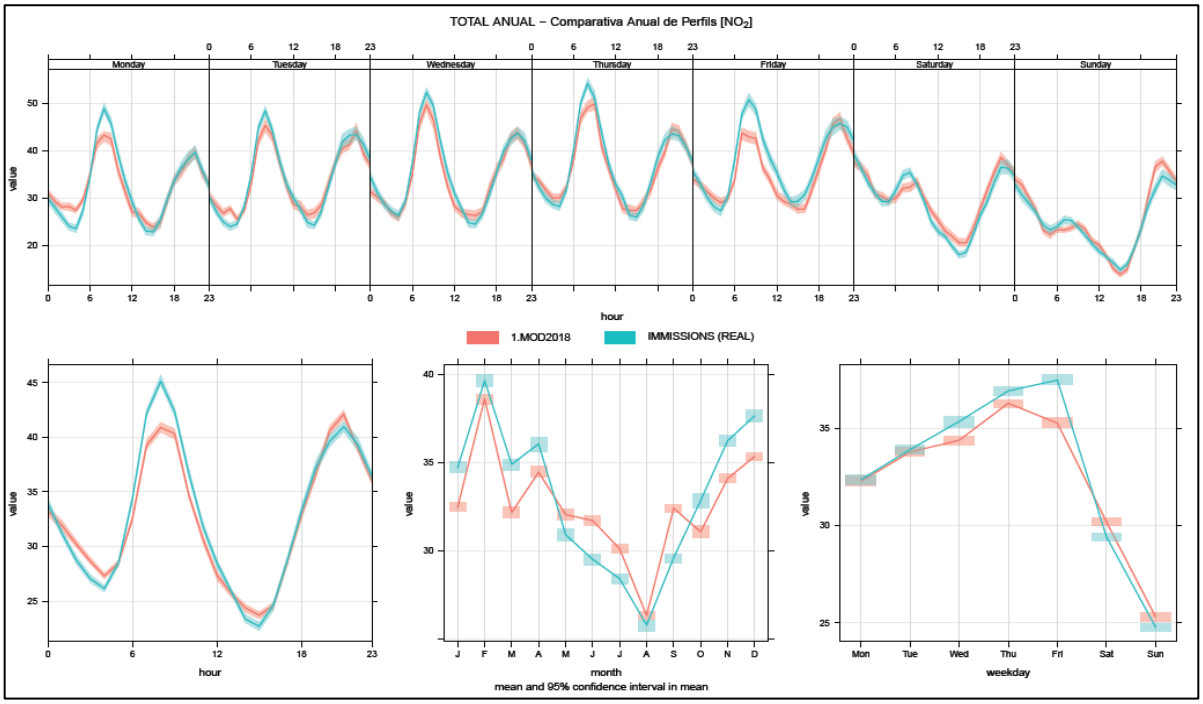
Source: Barcelona Regional.

As can be seen in the last figure, most of the points lie within these lines (±25% of perfect fit), showing a satisfactory behaviour of the model for the main three pollutants for the annual average, especially for NO_2. For the pollutant PM_{10}, there is more dispersion on results for all years. Anyway, yearly, there are only one or two stations when the mean annual concentrations are out of this range (±25%). However, model performance is more than acceptable when it comes to PM_{10}.

It can also be noticed how the model reproduces in a good way the evolution of air quality over the four years analysed. The first two years (2018 and 2019) report higher NO_2 values compared to the last two (2020 and 2021). Note how values have lowered a bit since COVID-19 pandemic started in 2020. For particulate matter, the variation seems not as strong as the NO_2. Anyway, it can be seen for 2020 and 2021 how the values are not as scattered as 2018 and 2019. This is due to the contribution of road traffic reduction, leading to an increase of the background pollution in the total pollution share.

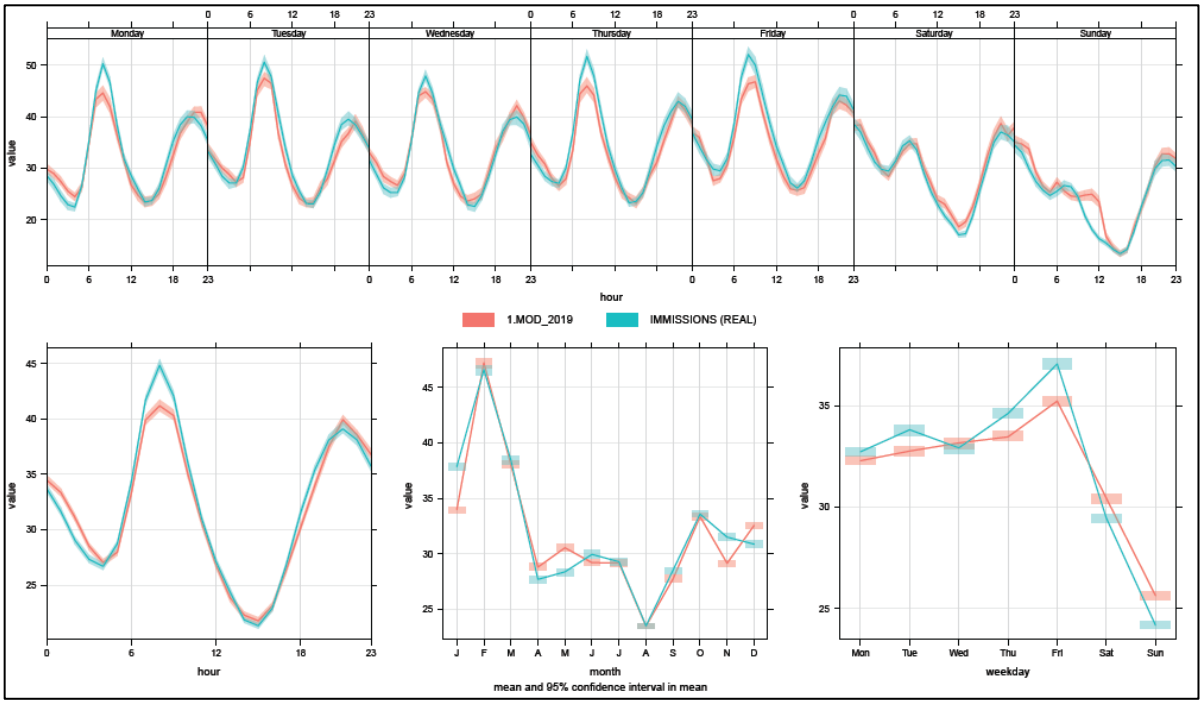
Concerning the NO_2 time-variation profiles, the following figures show the NO_2 time-variation profiles for the average of all stations in hourly, daily, weekly and monthly mean for the year 2018 (top) and 2019 (down).

Figure 15. Average NO_2 time-variation mean profiles for all stations. Hourly, daily, weekly and monthly basis (2018).



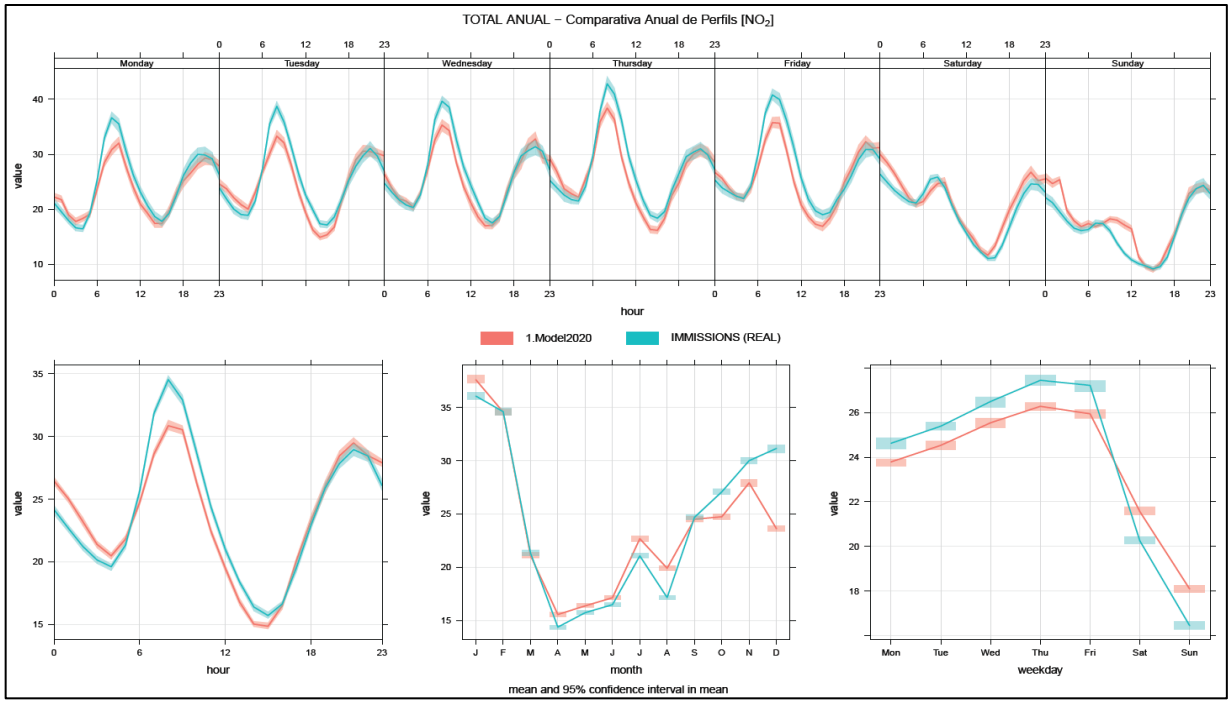
Source: Barcelona Regional

Figure 16. Average NO_2 time-variation mean profiles for all stations. Hourly, daily, weekly and monthly basis (2019).



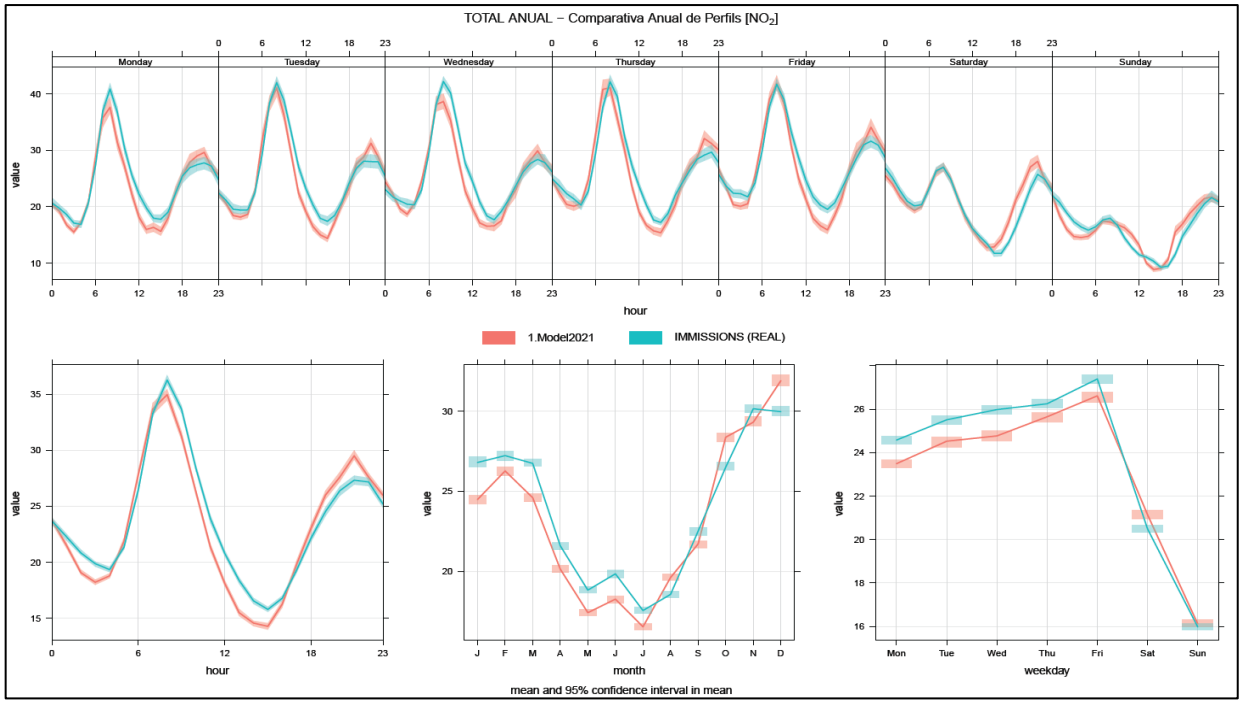
Source: Barcelona Regional

Figure 17. Average NO_2 time-variation mean profiles for all stations. Hourly, daily, weekly and monthly basis (2020).



Source: Barcelona Regional

Figure 18. Average NO_2 time-variation mean profiles for all stations. Hourly, daily, weekly and monthly basis (2021).



Source: Barcelona Regional

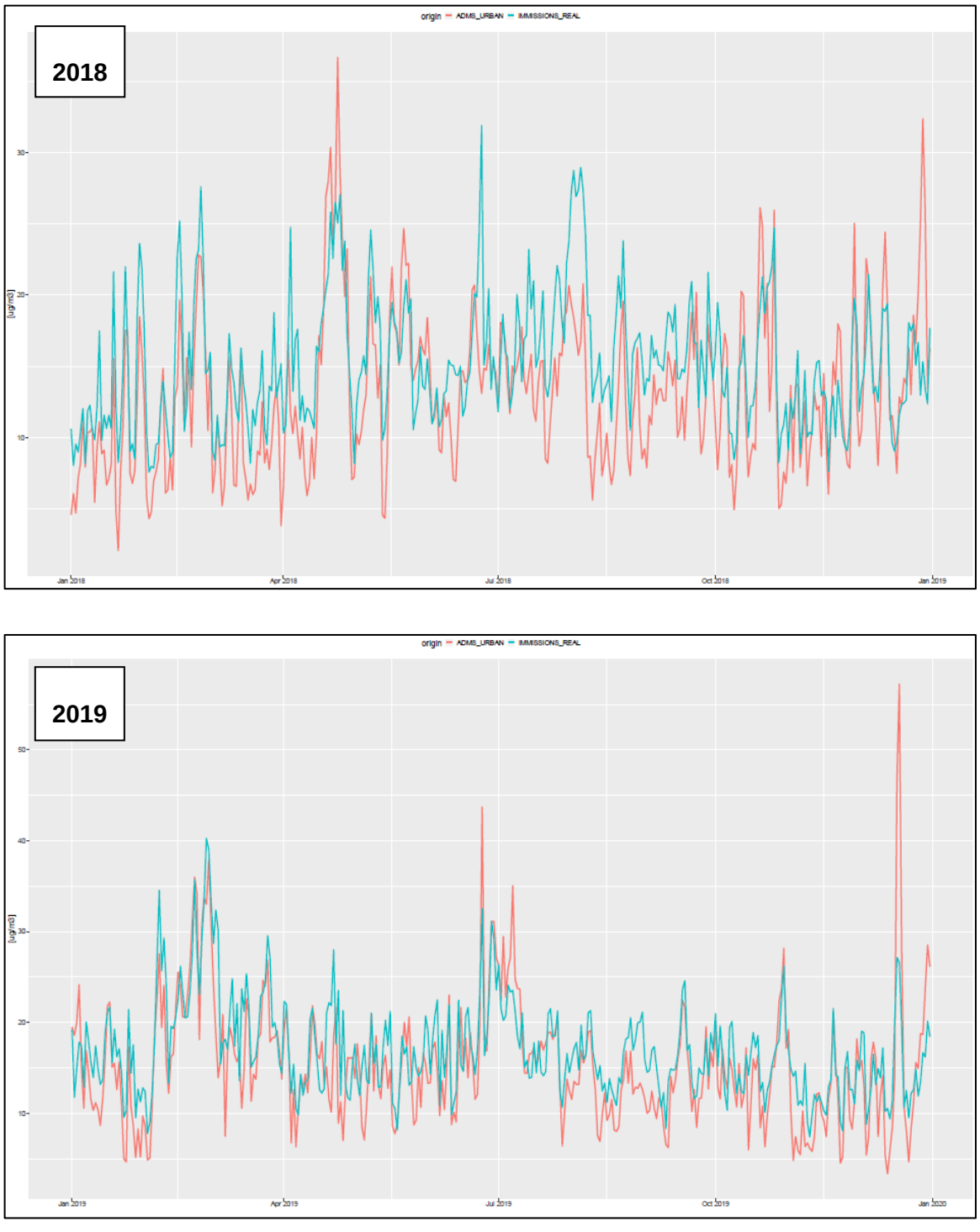
The last figures show how the modelled data fits the observed data really well for all time-variation profiles, for all years. Usually, monthly profiles show NO_2 concentration is the highest in winter months (December, January, and February). On the other hand, the NO_2 concentration seems to be the lowest in summer months (especially in August, except for 2020).

The weekly profiles exhibit the highest levels on Friday and the lowest, on Sunday. The hourly mean profiles reveal the highest concentrations are between 8:00 am and 9:00 am and between 9:00 pm and 10:00 pm. The lowest concentration seems to be at 2:00 pm.

Concerning the $PM_{2,5}$, XVPCA stations only measure them on daily basis. Therefore, the time-variation profiles below show only the daily mean data from the model and the XVPCA observations from 2018 to 2021 (complete).

In relation to the Black Carbon, the time-variation profiles show only the daily mean data from the model and the XVPCA observations from 2020 to 2021.

Figure 19: $PM_{2,5}$ daily mean data from the model and XVPCA observations for 2018 (top) and for 2019 (below).



Source: Barcelona Regional

Figure 20. $PM_{2,5}$ daily mean data from the model and XVPCA observations for 2020 (top) and for 2021 (below).

Source: Barcelona Regional

Figure 21: BC daily mean data from the model and XVPCA observations for 2020 (top) and for 2021 (below).

Once the annual means and temporal profiles have been checked and considered adequate, the performance of the model has been evaluated.

1.2. Total emissions

The total emissions amount for each year (for the considered pollutants) for modelling the air quality are shown in the following table. Each amount include strictly the emissions considered for the model, which include the resuspension and the calibration layer.

Table 2. Total emissions (all the municipalities included) [tons/year].

Year	NO_2	NO_X	PM_{10}	$PM_{2,5}$	SO_2	NH_3	BC
2018	4.048	24.930	2.260	1.379	1.822	122	338
2019	3.750	23.690	2.087	1.257	2.432	95	233
2020	2.731	17.370	1.931	1.079	1.770	99	129
2021	2.796	18.000	1.820	990	1.902	80	223

Source: Barcelona Regional.

1.3. Model Evaluation

Since the target of this work is the evaluation of NO_2 and $PM_{2,5}$, only detailed evaluation of such pollutants are presented.

The air quality model has been assessed in hourly basis for NO_2 and daily basis for $PM_{2,5}$ to ensure routes tracking to represent pollutant exposure as real as possible. Thus, some typical statistical parameters have been used, such as:

- **n.** It is the percentage of observed-modelled pairs of data used (Carslaw, 2019) [%].
- **FAC2.** The fraction of modelled values within a factor of two of the observed values are the fraction of model predictions.
- **MB.** The mean bias provides a good indication of the mean over or underestimate of predictions [µg/m^3].
- **MGE.** The mean gross error provides a good indication of the mean error regardless of whether it is an over or underestimate [µg/m^3].
- **NMB.** The normalised mean bias is useful for comparing results that cover different concentration scales [%].
- **RMSE.** It is a commonly used statistic that provides a good overall measure of how close modelled values are to predicted values [µg/m^3].
- **r.** The Pearson correlation coefficient is a measure of the strength of the linear relationship between two variables.

1.3.1. NO_2 Evaluation

The following tables summarize the main statistics evaluated for NO_2 (hourly basis) for 2018.

Table 3. Summary of the main statistics evaluated for NO_2 and 2018 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

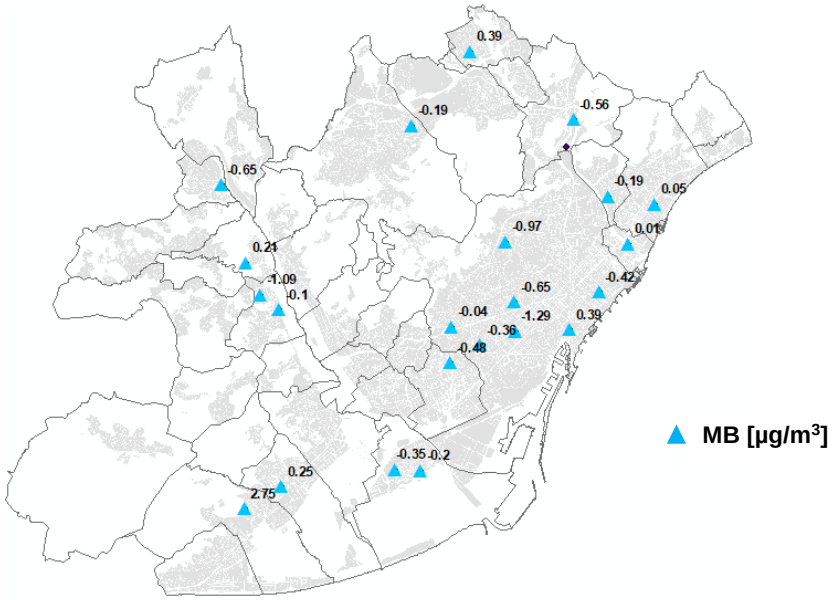
Receptor name	n (%)	FAC2	MB (µg/m³)	MGE (µg/m³)	NMB (%)	RMSE (µg/m³)	r
BCN_CIUTADELLA	94,9%	0,75	0,39	16,3	1,11%	22,6	0,39
BCN_EIXAMPLE	95,5%	0,86	-1,29	18,7	-2,39%	24,1	0,53
BCN_GRACIA	93,5%	0,85	-0,65	15,9	-1,41%	21,4	0,59
BCN_PALAU_REIAL	84,2%	0,75	-0,04	12,9	-0,15%	18,1	0,55
BCN_POBLENOU	95,5%	0,88	-0,42	11,8	-1,08%	16,0	0,71
BCN_SANTS	93,8%	0,87	-0,36	11,2	-1,10%	15,3	0,69
BCN_VALL_HEBRON	93,7%	0,77	-0,97	12,9	-3,35%	18,1	0,50
BD_AUSIAS_MARC	93,8%	0,86	0,05	10,8	0,13%	15,1	0,75
BV_MORAGUES_MONTSERRAT	93,1%	0,91	0,39	8,9	1,17%	12,4	0,80
G_PARC_MILENI	94,9%	0,70	2,75	8,4	18,6%	12,2	0,50
HOS_TORRENT_GOR	89,4%	0,91	-0,48	10,6	-1,34%	14,5	0,75
MONTC_COMPANYS	90,1%	0,83	-0,56	11,4	-1,77%	15,5	0,61
P_ROCA_VILANA	89,2%	0,72	0,21	9,7	1,03%	13,6	0,52
PLL_CEM_SAGNIER	93,8%	0,85	-0,35	11,3	-1,03%	14,9	0,65
PR_JARDINS	92,7%	0,88	-0,20	10,2	-0,60%	13,9	0,72
SA_CANAL_OLIMPIC	95,5%	0,87	0,01	11,9	0,02%	16,2	0,77
SAB_CEIP_JOSEP_PLA	94,5%	0,88	-0,65	12,1	-1,67%	17,4	0,65
SCG_BALLDOVINA	95,1%	0,76	-0,19	13,7	-0,57%	18,2	0,52
SCV_PARC_S_FRANCESC	94,5%	0,77	-0,19	10,4	-0,80%	14,2	0,54
SVH_ALABA	92,1%	0,79	-1,09	10,9	-3,95%	14,9	0,60
SVH_RIBOT_S_MIQUEL	94,4%	0,85	-0,10	11,5	-0,31%	15,6	0,66
VILADECANS_ATRIUM	94,9%	0,73	0,25	10,3	1,13%	14,6	0,51
Desired value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional

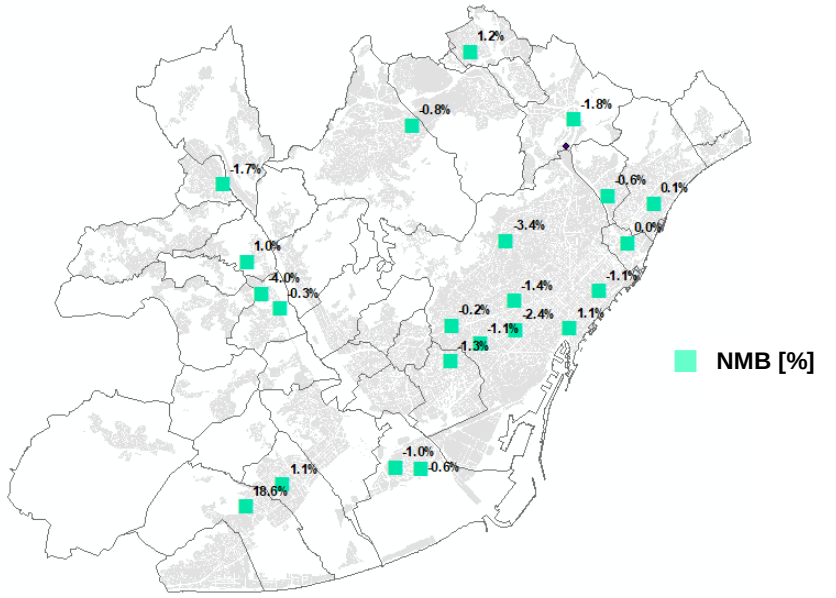
The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2018.

Figure 22. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for NO_2 during 2018.

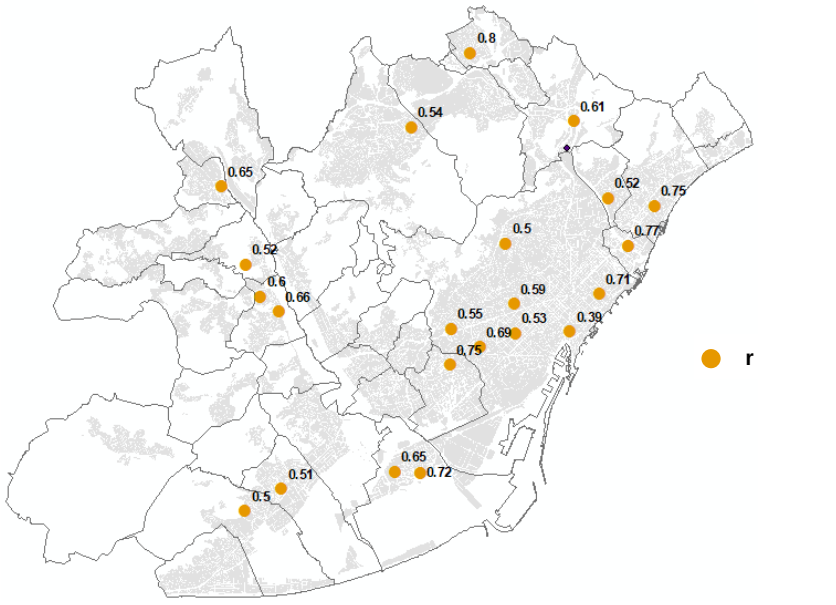
a)



b)



c)



Source: Barcelona Regional

The map a) of last figure shows that *MB* values lie between -1,3 and 2,75 µg/m^3.

The map b) point *NMB* values range between -3,95% and 18,6%.

In the case of map c), most the *r* values are above 0,6 (which means more than acceptable correlation). As it can be seen, these maps show that most of the air quality stations have a good performance, except for Parc Mil·leni station (located in Gavà) and Ciutadella (Barcelona), where the model shows a modest, but acceptable performance for the main statistics.

The following tables summarize the main statistics evaluated for NO_2 (hourly basis) for 2019.

Table 1: Summary of the main statistics evaluated for NO_2 and 2019 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

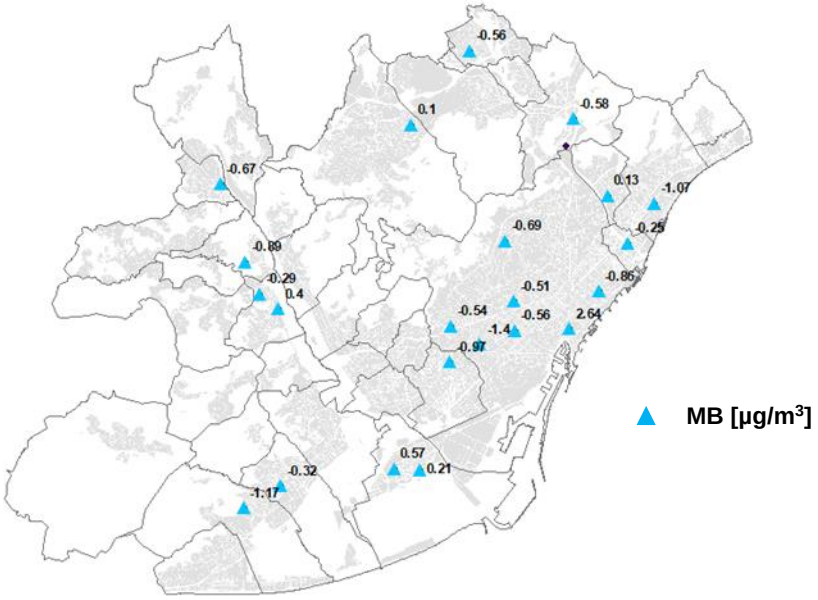
Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
BCN_CIUTADELLA	96,7%	0,75	2,64	15,2	8,1%	21,7	0,48
BCN_EIXAMPLE	95,8%	0,85	-0,56	16,9	-1,1%	21,8	0,59
BCN_GRACIA	96,8%	0,87	-0,51	13,8	-1,2%	18,8	0,72
BCN_PALAU_REIAL	91,7%	0,76	-0,54	13,0	-1,9%	18,3	0,56
BCN_POBLENOU	97,6%	0,84	-0,86	12,1	-2,3%	16,5	0,69
BCN_SANTS	97,2%	0,78	-1,40	12,9	-4,5%	17,9	0,58
BCN_VALL_HEBRON	96,4%	0,81	-0,69	12,0	-2,4%	17,0	0,63
BD_AUSIAS_MARC	96,8%	0,81	-1,07	12,0	-3,2%	16,2	0,69
BV_MORAGUES_MONTSERRAT	96,0%	0,83	-0,56	12,0	-1,6%	16,4	0,69
G_PARC_MILENI	95,8%	0,61	-1,17	9,0	-7,5%	13,0	0,35
HOS_TORRENT_GOR	97,3%	0,81	-0,97	13,1	-2,9%	18,2	0,61
MONTC_COMPANYS	95,1%	0,78	-0,58	13,3	-1,7%	17,5	0,59
P_ROCA_VILANA	95,7%	0,69	-0,89	11,5	-3,9%	16,1	0,52
PLL_CEM_SAGNIER	96,8%	0,83	0,15	11,5	0,5%	15,3	0,63
PR_JARDINS	96,9%	0,84	-0,43	11,0	-1,3%	14,7	0,71
SA_CANAL_OLIMPIC	97,6%	0,80	-0,25	13,4	-0,7%	18,0	0,69
SAB_CEIP_JOSEP_PLA	97,3%	0,73	-0,67	17,2	-1,8%	22,7	0,52
SCG_BALLDOVINA	96,5%	0,75	0,13	14,0	0,4%	18,7	0,57
SCV_PARC_S_FRANCESC	95,2%	0,75	0,10	10,4	0,4%	14,1	0,61
SVH_ALABA	94,9%	0,75	-0,29	12,6	-1,0%	17,6	0,58
SVH_RIBOT_S_MIQUEL	97,4%	0,79	0,40	12,7	1,3%	17,2	0,64
VILADECANS_ATRIUM	95,4%	0,69	-0,32	13,0	-1,3%	17,9	0,43
Desired value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional

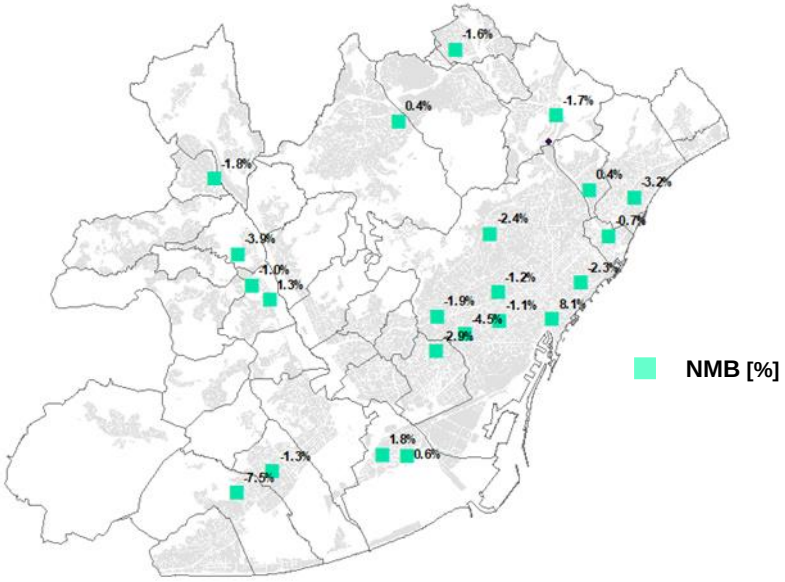
The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2019.

Figure 23. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for NO_2 during 2019.

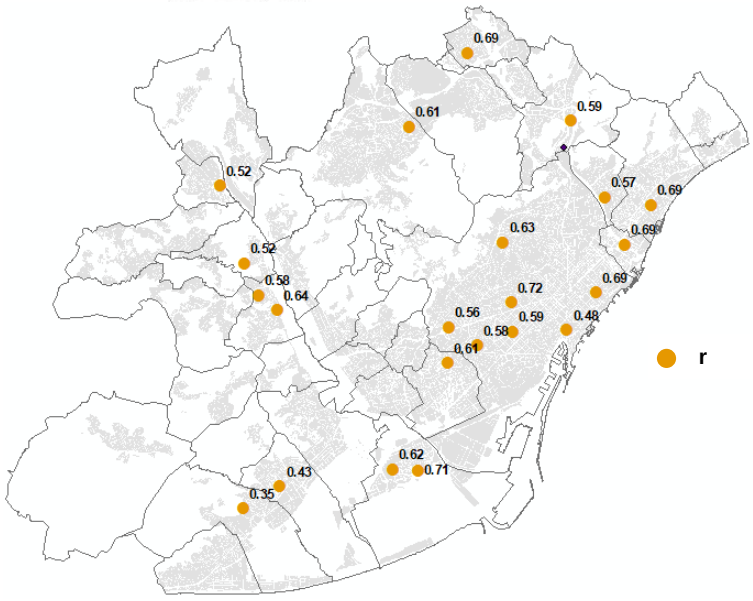
a)



b)



c)



Source: Barcelona Regional

The map a) shows how MB values vary between -1,4 and 2,6 µg/m^3.

In the case of map b), the NMB fraction is less than ±10% at all XVPCA stations, ranging between -7,5 and 8,1 µg/m^3.

In the case of map c), most of the *r* values are above 0,6 (which means more than acceptable correlation).

Results, reveal, in general, a good behaviour of NO_2 in Barcelona and in its surrounding municipalities. However, a modest performance has been observed in Parc del Mil·leni (Gavà) and Ciutadella (Barcelona) air quality stations (but within satisfactory values). This modest behaviour may be due to local meteorological effects, since these locations are protected in Gavà by Serra de les Ferreres mountain range and Montjuïc mountain in Barcelona, in the windward side. Therefore, a microclimate is created near the location of both air quality stations. The following tables summarize the main statistics evaluated for NO_2 (hourly basis) for 2020.

Table 4. Summary of the main statistics evaluated for NO_2 and 2020 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

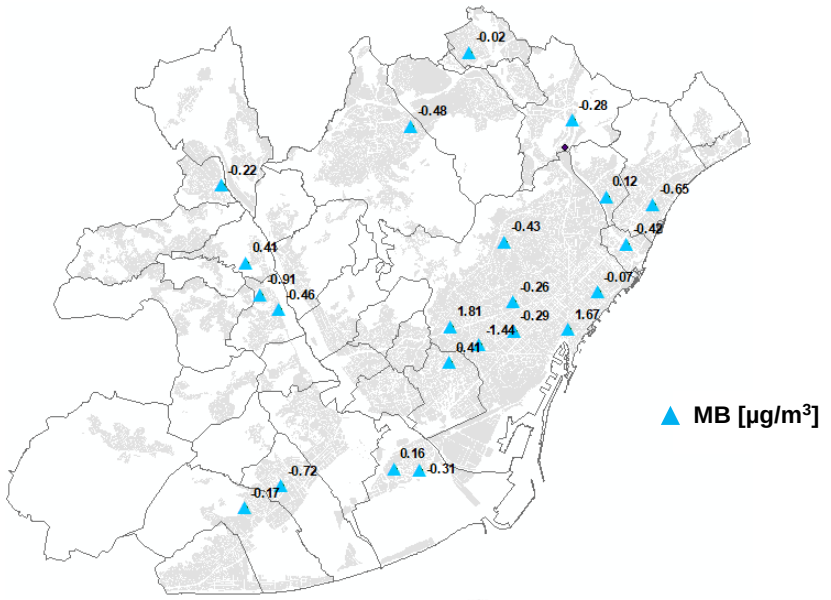
Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
BCN_CIUTADELLA	93,7%	0,71	1,67	11,68	6,93%	16,7	0,58
BCN_EIXAMPLE	94,4%	0,81	-0,29	12,64	-0,80%	16,4	0,68
BCN_GRACIA	95,3%	0,83	-0,26	10,94	-0,79%	14,7	0,75
BCN_PALAU_REIAL	89,4%	0,70	1,81	9,44	9,72%	13,4	0,63
BCN_POBLENOU	95,3%	0,84	-0,07	9,45	-0,23%	12,8	0,78
BCN_SANTS	95,4%	0,77	-1,44	10,11	-5,98%	14,1	0,66
BCN_VALL_HEBRON	87,4%	0,80	-0,43	9,16	-1,91%	13,1	0,70
BD_AUSIAS_MARC	93,9%	0,76	-0,65	10,10	-2,48%	13,8	0,69
BV_MORAGUES_MONTSERRAT	93,7%	0,90	-0,02	6,30	-0,08%	9,3	0,88
G_PARC_MILENI	94,1%	0,60	-0,17	6,96	-1,48%	10,2	0,37
HOS_TORRENT_GOR	95,1%	0,78	0,41	10,00	1,65%	13,8	0,68
MONTC_COMPANYS	94,6%	0,78	-0,28	9,93	-1,08%	13,5	0,69
P_ROCA_VILANA	94,4%	0,70	0,41	7,65	2,64%	10,8	0,56
PLL_CEM_SAGNIER	94,4%	0,78	0,16	9,20	0,67%	12,4	0,69
PR_JARDINS	95,2%	0,80	-0,31	8,63	-1,27%	11,8	0,76
SA_CANAL_OLIMPIC	93,9%	0,74	-0,42	10,72	-1,54%	14,7	0,71
SAB_CEIP_JOSEP_PLA	93,1%	0,86	-0,22	8,84	-0,78%	12,9	0,74
SCG_BALLDOVINA	95,1%	0,74	0,12	10,90	0,46%	14,6	0,63
SCV_PARC_S_FRANCESC	93,0%	0,74	-0,48	8,37	-2,47%	11,6	0,64
SVH_ALABA	92,2%	0,77	-0,91	8,74	-4,11%	12,1	0,69
SVH_RIBOT_S_MIQUEL	95,0%	0,80	-0,46	8,72	-1,94%	12,1	0,73
VILADECANS_ATRIUM	92,3%	0,68	-0,72	8,61	-4,26%	12,4	0,54
Desired value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional

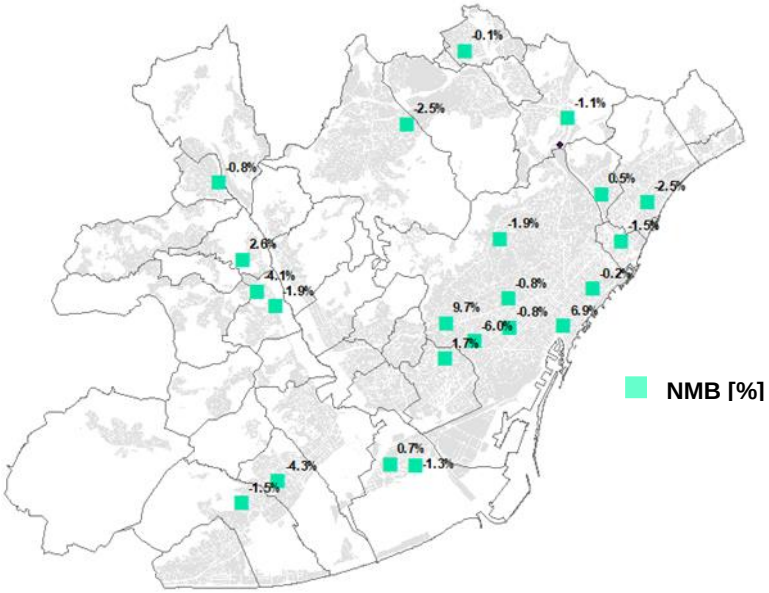
The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2020.

Figure 24. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for NO_2 during 2020.

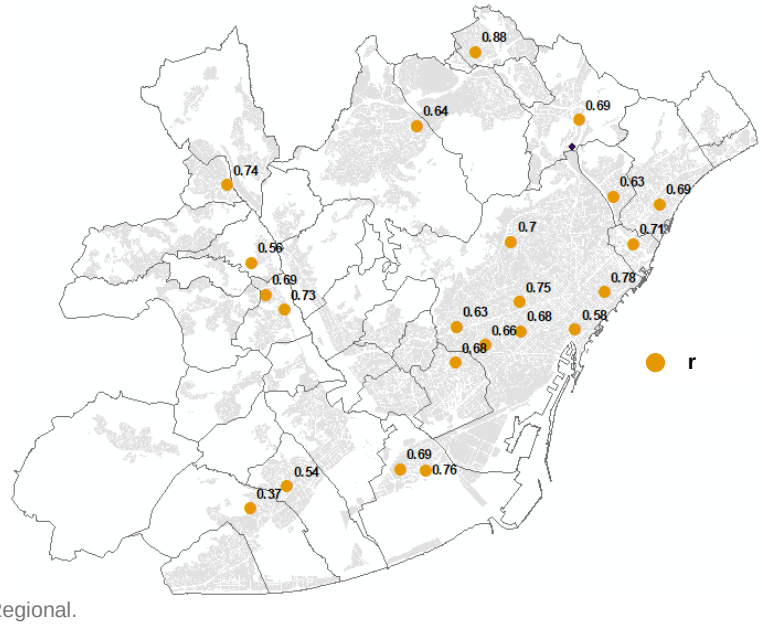
a)



b)



c)



Source: Barcelona Regional.

Data shows for MB to lie between -1,44 and 1,81 µg/m^3.

For NMB, values range between –5,98% and 9,72%.

r parameter cover values between 0,37 for Gavà (Parc del Mil·leni) station again, and 0,88 for Barberà del Vallès (Moragues - Montserrat).

The following tables summarize the main statistics evaluated for NO_2 (hourly basis) for 2021.

Table 5. Summary of the main statistics evaluated for NO_2 and 2021 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

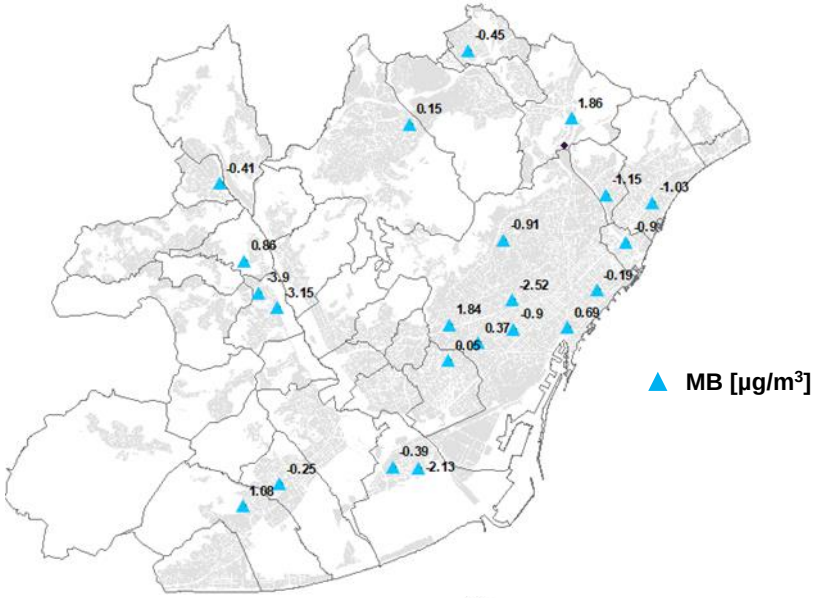
Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
BCN_CIUTADELLA	97%	0,68	0,69	13,05	2,79%	18,21	0,45
BCN_EIXAMPLE	96%	0,83	-0,90	13,25	-2,38%	17,12	0,60
BCN_GRACIA	96%	0,79	-2,52	11,99	-8,00%	15,92	0,59
BCN_PALAU_REIAL	96%	0,73	1,84	9,17	9,99%	12,80	0,57
BCN_POBLENOU	97%	0,76	-0,19	10,70	-0,74%	14,55	0,60
BCN_SANTS	97%	0,76	0,37	9,85	1,70%	13,75	0,58
BCN_VALL_HEBRON	97%	0,76	-0,91	9,27	-4,43%	13,37	0,56
BD_AUSIAS_MARC	96%	0,73	-1,03	10,30	-4,22%	13,99	0,64
BV_MORAGUES_MONTSERRAT	96%	0,73	-0,45	11,68	-1,82%	17,40	0,58
G_PARC_MILENI	92%	0,62	1,08	6,87	9,91%	10,16	0,38
HOS_TORRENT_GOR	95%	0,80	0,05	9,30	0,23%	13,04	0,64
MONTC_COMPANYS	97%	0,66	1,86	16,30	6,48%	26,47	0,48
P_ROCA_VILANA	95%	0,63	0,86	9,31	5,28%	13,56	0,51
PLL_CEM_SAGNIER	81%	0,71	-0,39	10,21	-1,76%	13,44	0,53
PR_JARDINS	97%	0,79	-2,13	9,77	-8,36%	13,40	0,68
SA_CANAL_OLIMPIC	96%	0,71	-0,90	11,11	-3,51%	15,45	0,67
SAB_CEIP_JOSEP_PLA	95%	0,66	-0,41	14,23	-1,47%	19,80	0,53
SCG_BALLDOVINA	96%	0,73	-1,15	11,05	-4,45%	14,99	0,60
SCV_PARC_S_FRANCESC	94%	0,69	0,15	9,93	0,75%	14,56	0,55
SVH_ALABA	93%	0,62	-3,90	11,87	-17,2%	17,11	0,52
SVH_RIBOT_S_MIQUEL	95%	0,66	-3,15	12,56	-12,2%	18,54	0,57
VILADECANS_ATRIUM	83%	0,72	-0,25	8,53	-1,45%	12,34	0,53
Ideal value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional

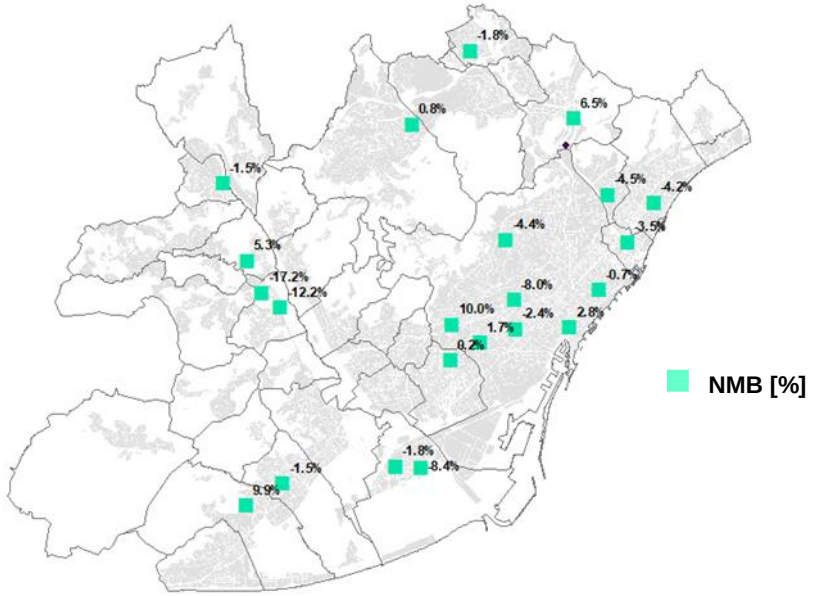
The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2021.

Figure 25. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for NO_2 during 2021.

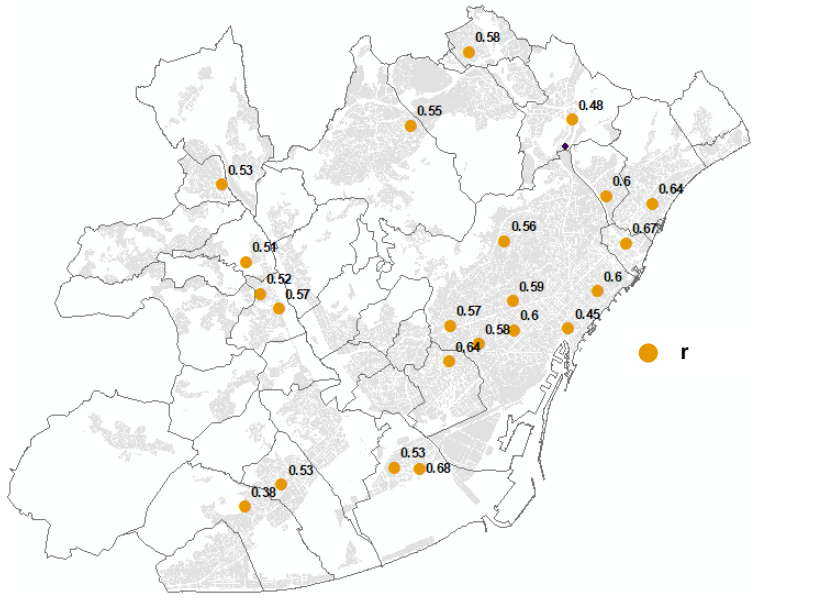
a)



b)



c)



Source: Barcelona Regional.

Data from last figures illustrate MB values range between -3,9 and 1,86 µg/m^3.

In relation to NMB, values are between –17,2% and 10%.

r parameter ranges between 0,38 for Gavà (Parc del Mil·leni) station again, and 0,68 for El Prat de LLobregat (Jardins)

1.3.2. $PM_{2,5}$ Evaluation

As said before, the $PM_{2,5}$ data is only available on daily basis (since the measurement stations are manual). Therefore, data from such stations are used for $PM_{2,5}$ evaluation. These stations are used by local authority in order to evaluate this pollutant.

The following tables summarize the main statistics evaluated for $PM_{2,5}$ (daily basis) for 2018.

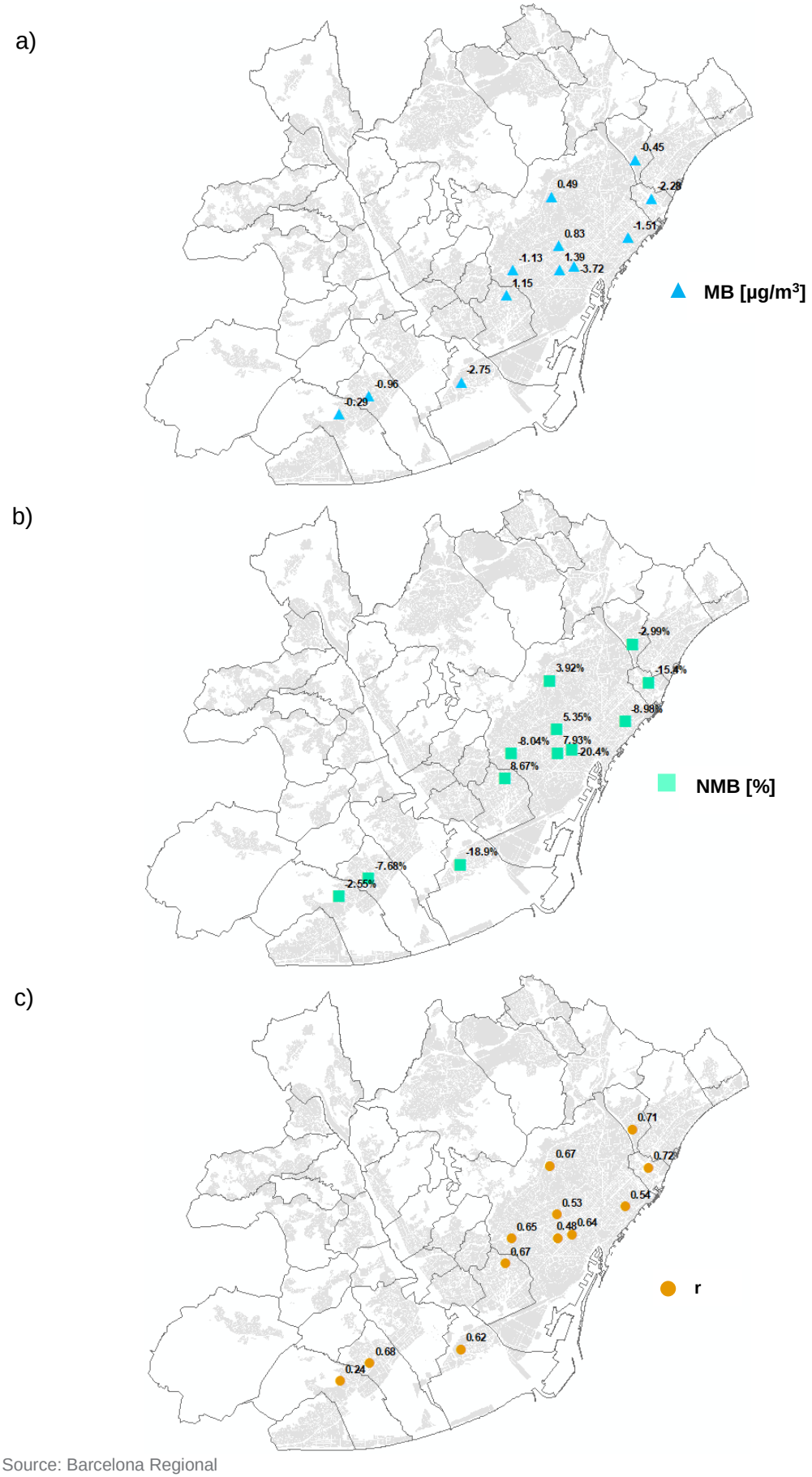
Table 6. Summary of the main statistics evaluated for $PM_{2,5}$ and 2018 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
BCN_EIXAMPLE	95,3%	0,96	1,39	4,84	7,93%	6,49	0,48
BCN_GRACIA	95,1%	0,95	0,83	4,12	5,35%	5,98	0,53
BCN_PLAÇA_UNIV	93,7%	0,92	-3,72	4,80	-20,4%	5,98	0,64
BCN_POBLENOU	96,7%	0,95	-1,51	4,20	-8,98%	6,06	0,54
BCN_VALL_HEBRON	90,7%	0,97	0,49	3,11	3,92%	4,09	0,67
BCN_Z_UNIV	97,3%	0,94	-1,13	3,29	-8,04%	4,23	0,65
G_PARC_MILENI	45,2%	0,90	-0,29	3,67	-2,55%	7,01	0,24
HOS_TORRENT_GOR	44,7%	0,98	1,15	3,10	8,67%	4,75	0,67
PLL_CEM_SAGNIER	44,7%	0,90	-2,75	4,08	-18,9%	5,11	0,62
SA_CANAL_OLIMPIC	44,9%	0,97	-2,28	3,46	-15,4%	4,75	0,72
SCG_BALLDOVINA	42,2%	0,98	-0,45	3,20	-2,99%	4,40	0,71
VILADECANS_ATRIUM	45,2%	0,93	-0,96	3,00	-7,68%	3,94	0,68
Desired value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional

The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2018.

Figure 26. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for $PM_{2,5}$ during 2018.



Source: Barcelona Regional

As can be seen in last figures and Table 6, the number of stations with $PM_{2,5}$ observations is way fewer than NO_2 ones. These stations are located in Barcelona city and its surrounding municipalities.

The map a) shows MB values range between –3,72 and 1,39 µg/m^3.

The map b) reveal NMB values are between -20,4% and 8,7%.

The last map (map c)), show how values of correlation lie between 0,24 and 0,72.

In general, statistics show good model performance when it comes to $PM_{2,5}$. However, the model tends to underestimate the $PM_{2,5}$ concentration. Modest behaviour has been found at Gavà (Parc del Milenni).

The following tables summarize the main statistics evaluated for $PM_{2,5}$ (daily basis) for 2019.

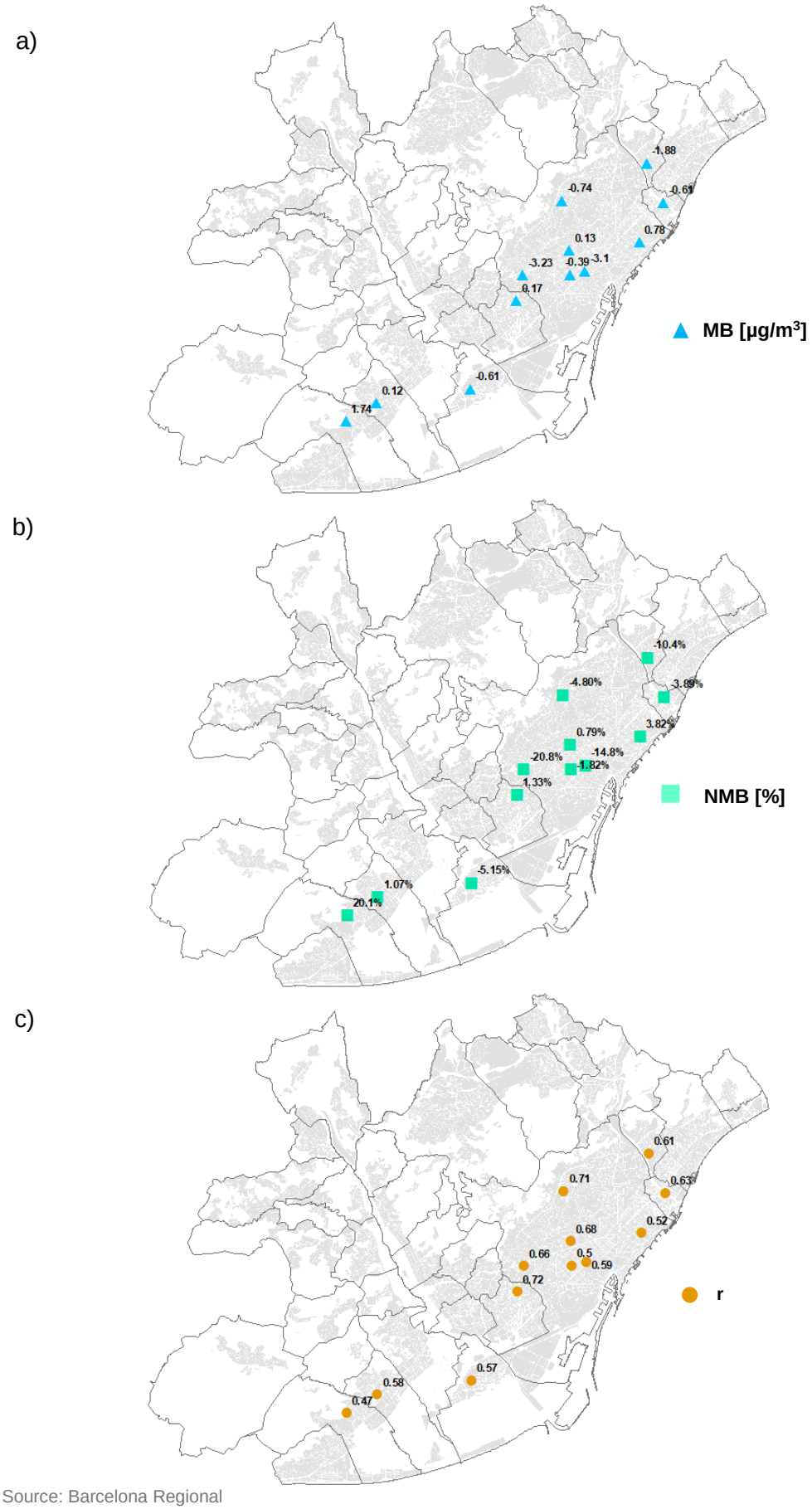
Table 7. Summary of the main statistics evaluated for $PM_{2,5}$ and 2019 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
BCN_EIXAMPLE	86,8%	0,95	-0,39	4,97	-1,82%	8,34	0,50
BCN_GRACIA	93,7%	0,95	0,13	3,82	0,79%	5,06	0,68
BCN_PLAÇA_UNIV	90,7%	0,91	-3,10	5,51	-14,8%	8,38	0,59
BCN_POBLENOU	92,3%	0,95	0,78	5,55	3,82%	10,72	0,52
BCN_VALL_HEBRON	92,1%	0,96	-0,74	3,41	-4,80%	4,72	0,71
BCN_Z_UNIV	93,7%	0,87	-3,23	4,28	-20,8%	5,45	0,66
G_PARC_MILENI	45,5%	0,92	1,74	3,66	20,1%	5,44	0,47
HOS_TORRENT_GOR	49,3%	0,97	0,17	3,04	1,33%	4,23	0,72
PLL_CEM_SAGNIER	44,1%	0,95	-0,61	3,45	-5,15%	5,00	0,57
SA_CANAL_OLIMPIC	48,8%	0,95	-0,61	4,18	-3,89%	5,84	0,63
SCG_BALLDOVINA	46,6%	0,94	-1,88	4,71	-10,4%	7,27	0,61
VILADECANS_ATRIUM	34,5%	0,90	0,12	3,79	1,07%	4,96	0,58
Desired value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional.

The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2019.

Figure 27. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for $PM_{2,5}$ during 2019.



Source: Barcelona Regional

The map a) shows MB values range between –3,23 and 1,74 µg/m^3.

The map b) reveal NMB values are between -20,8% and 20,1%.

The last map (map c)), show how values of correlation lie between 0,47 and 0,72.

The following tables summarize the main statistics evaluated for $PM_{2,5}$ (daily basis) for 2020.

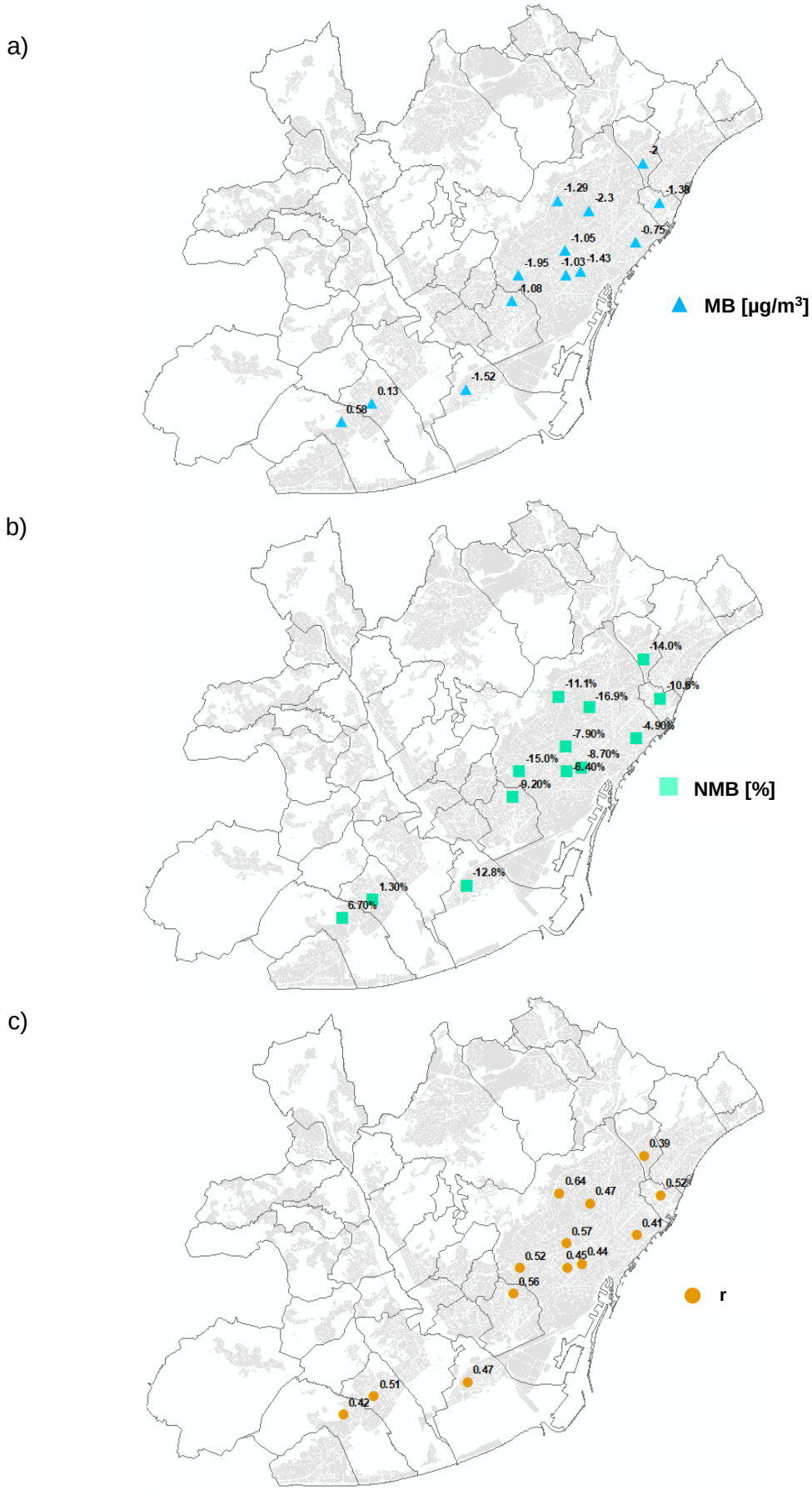
Table 8. Summary of the main statistics evaluated for $PM_{2,5}$ and 2020 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
BCN_EIXAMPLE	85,0%	0,94	-1,03	4,37	-6,4%	6,05	0,45
BCN_GRACIA	92,0%	0,95	-1,05	3,24	-7,9%	4,51	0,57
BCN_IES_GOYA	35,0%	0,91	-2,30	4,27	-16,9%	6,48	0,47
BCN_PLAÇA_UNIV	83,0%	0,90	-1,43	5,09	-8,7%	7,30	0,44
BCN_POBLENOU	84,0%	0,91	-0,75	4,49	-4,9%	6,75	0,41
BCN_VALL_HEBRON	89,0%	0,96	-1,29	2,62	-11,1%	3,59	0,64
BCN_Z_UNIV	81,0%	0,93	-1,95	3,47	-15,0%	5,33	0,52
G_PARC_MILENI	38,0%	0,93	0,58	3,19	6,7%	4,86	0,42
HOS_TORRENT_GOR	51,0%	0,93	-1,08	3,17	-9,2%	5,03	0,56
PLL_CEM_SAGNIER	37,0%	0,89	-1,52	4,06	-12,8%	5,75	0,47
SA_CANAL_OLIMPIC	47,0%	0,90	-1,38	4,12	-10,6%	5,63	0,52
SCG_BALLDOVINA	46,0%	0,91	-2,00	4,87	-14,0%	8,87	0,39
VILADECANS_ATRIUM	37,0%	0,94	0,13	3,20	1,3%	4,61	0,51
Desired value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional.

The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2020.

Figure 28. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for $PM_{2,5}$ during 2020.



Source: Barcelona Regional.

Last figures report MB values from -2,3 and 0,58 µg/m^3 (map a)). In relation to NMB, its values range between –16,9% and 6,7% (map b)), while the values of r parameter lie between 0,39 and 0,64 (map c)).

The following tables summarize the main statistics evaluated for $PM_{2,5}$ (daily basis) for 2021.

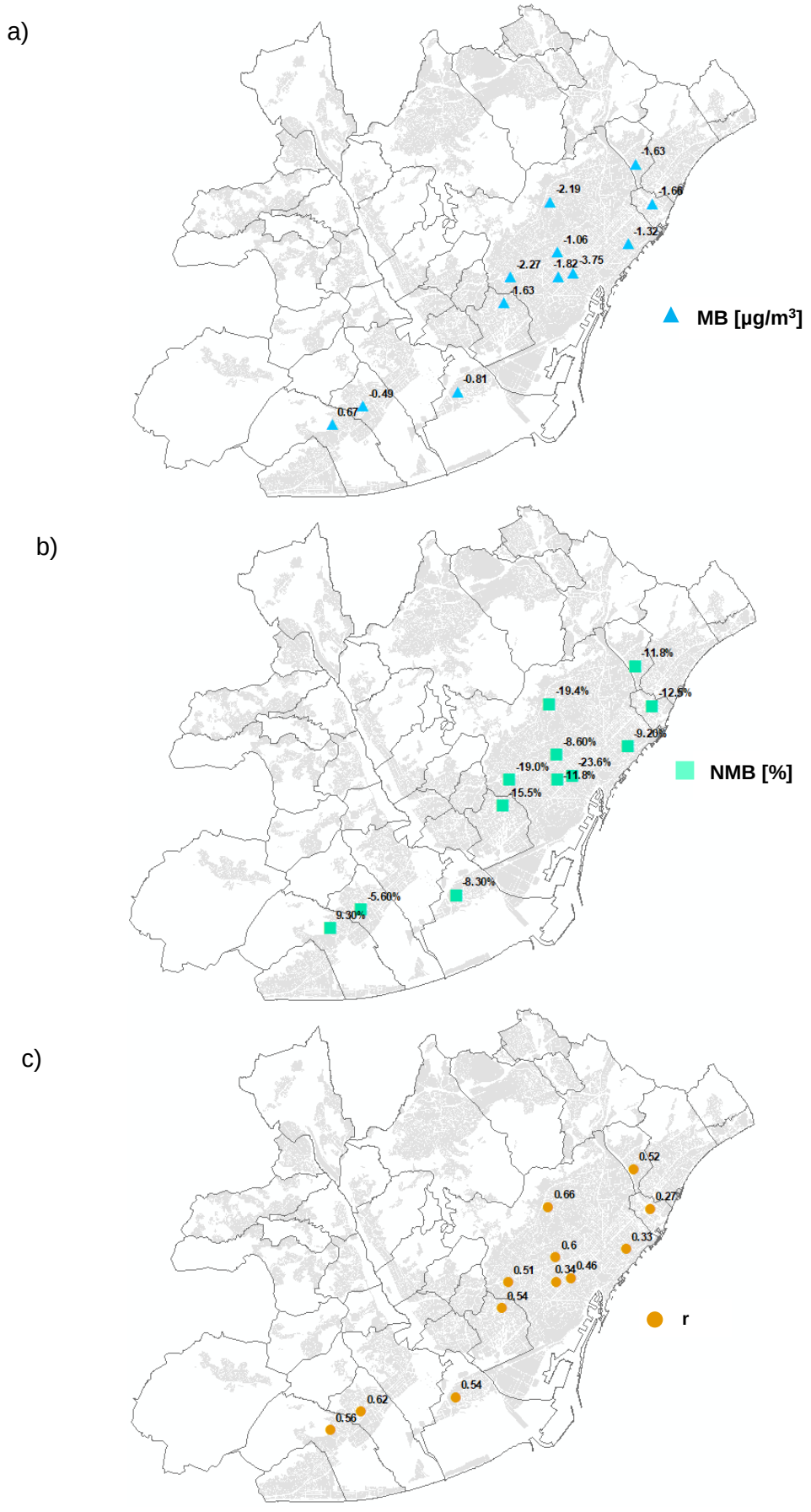
Figure 29. Summary of the main statistics evaluated for $PM_{2,5}$ and 2021 for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
BCN_EIXAMPLE	77,8%	0,90	-1,82	4,45	-11,8%	6,42	0,34
BCN_GRACIA	86,6%	0,94	-1,06	2,97	-8,60%	4,03	0,60
BCN_PLAÇA_UNIV	78,4%	0,85	-3,75	4,93	-23,6%	6,72	0,46
BCN_POBLENOU	87,4%	0,87	-1,32	4,89	-9,20%	6,60	0,33
BCN_VALL_HEBRON	82,2%	0,89	-2,19	3,01	-19,4%	3,89	0,66
BCN_Z_UNIV	83,8%	0,89	-2,27	3,36	-19,0%	4,86	0,51
G_PARC_MILENI	40,0%	0,91	0,67	2,37	9,30%	3,32	0,56
HOS_TORRENT_GOR	64,1%	0,91	-1,63	2,87	-15,5%	3,97	0,54
PLL_CEM_SAGNIER	38,4%	0,91	-0,81	3,10	-8,30%	4,03	0,54
SA_CANAL_OLIMPIC	63,8%	0,90	-1,66	4,03	-12,5%	7,92	0,27
SCG_BALLDOVINA	55,3%	0,91	-1,63	4,12	-11,8%	5,88	0,52
VILADECANS_ATRIUM	40,8%	0,95	-0,49	2,45	-5,60%	3,34	0,62
Desired value	**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

Source: Barcelona Regional

The following figures plot, on an AMB map, the spatial distribution of the values of the *MB*, *NMB* and *r* for 2021.

Figure 30. Spatial distribution of MB (triangular blue), NMB (green) and r (brown rounded) for $PM_{2,5}$ during 2021.



Source: Barcelona Regional

Last figures report MB values from -3,75 and 0,67 µg/m^3 (map a)). In relation to NMB, its values range between –23,6% and 9,3% (map b)), while the values of r parameter lie between 0,27 and 0,66 (map c)). This time the lowest r value was a l'Hospitalet de Llobregat,

1.3.3. Black Carbon Evaluation

The data availability for Black Carbon measurements is lower than the other evaluated pollutants. In fact, there is only available data for 2020 and 2021 for l'Eixample air quality station, while there's only 2021 data for Parc de la Vall d'Hebron air quality station.

Table 9. Summary of the main statistics evaluated for BC (2020 and 2021) for each station (XVPCA). The desired (ideal) value for each variable appears at the end of each column.

Year	Receptor name	n (%)	FAC2	MB (µg/m^3)	MGE (µg/m^3)	NMB (%)	RMSE (µg/m^3)	r
2020	BCN_EIXAMPLE	66,4%	0,926	-0,565	0,595	-30,7%	0,829	0,820
2021	BCN_EIXAMPLE	98,1%	0,994	-0,016	0,333	-1,0%	0,482	0,776
2021	BCN_VALL_HEBRON	67,9%	0,948	-0,203	0,251	-21,6%	0,354	0,830
Desired value		**100%**	**0,5-2,0**	**0**	**0**	**0**	**0**	**1**

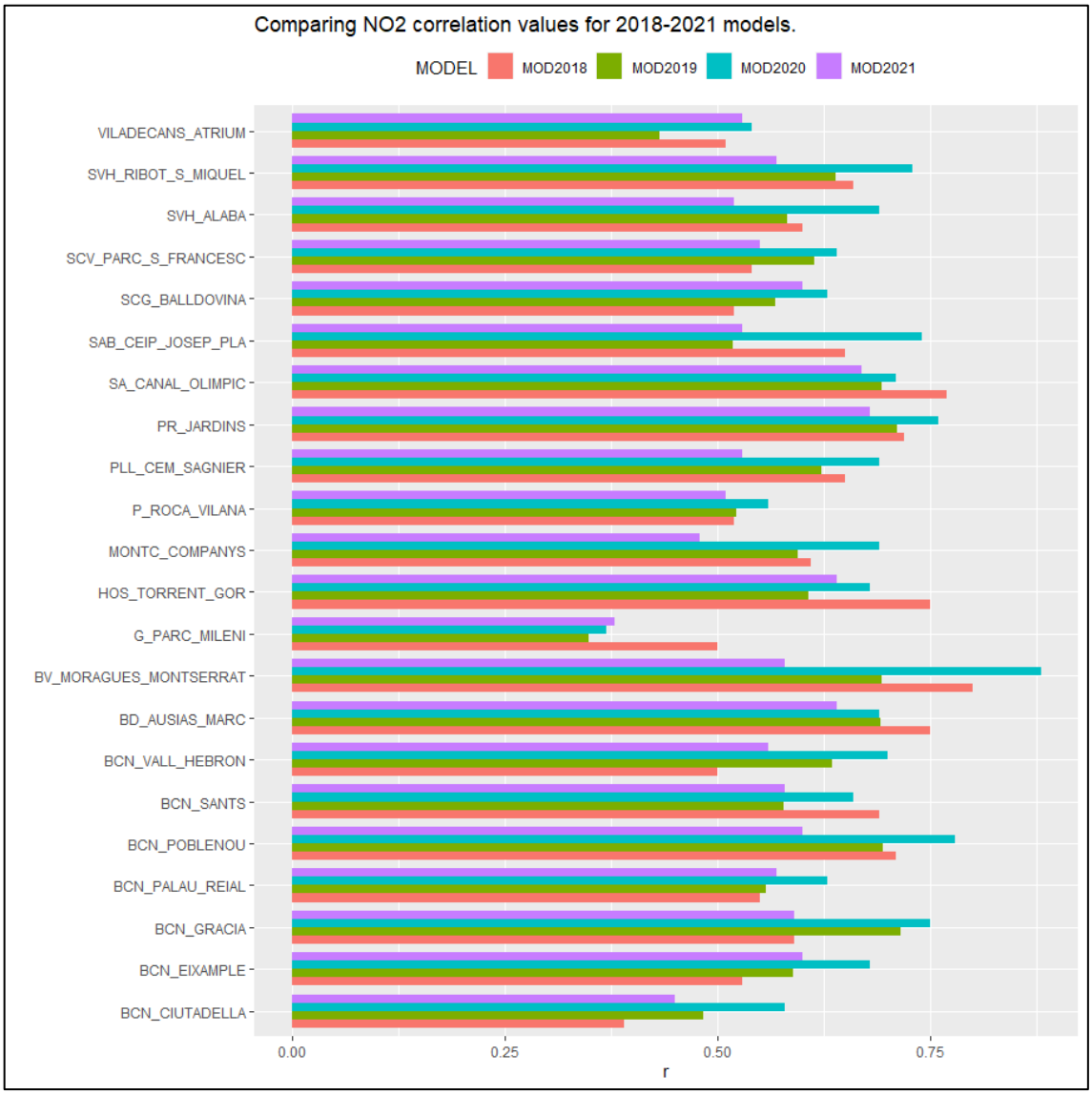
Source: Barcelona Regional.

1.3.4. Annual results comparison for NO_2 and $PM_{2,5}$

NO_2 performance evaluation

The following figure shows, for NO_2, the correlation level (r parameter) between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 31. NO_2 correlation level (r parameter) between the results of the model and the real observations of the XVPCA stations, for each year of analysis (2018 to 2021).



Source: Barcelona Regional

As it can be seen in Figure 31, the hourly correlation in most of the stations has a *r* parameter greater than 0,5 (while the desired value is 1), except for the following stations: Parc del Mil·leni (Gavà) and Ciutadella (Barcelona). This could due to the proximity of a mountain range (Serra de les Ferreres) located windward of Parc del Mil·leni station (Gavà) and the Montjuïc mountain affecting Ciutadella

station (Barcelona). A microclimate is created near both stations and it is not well represented by the model.

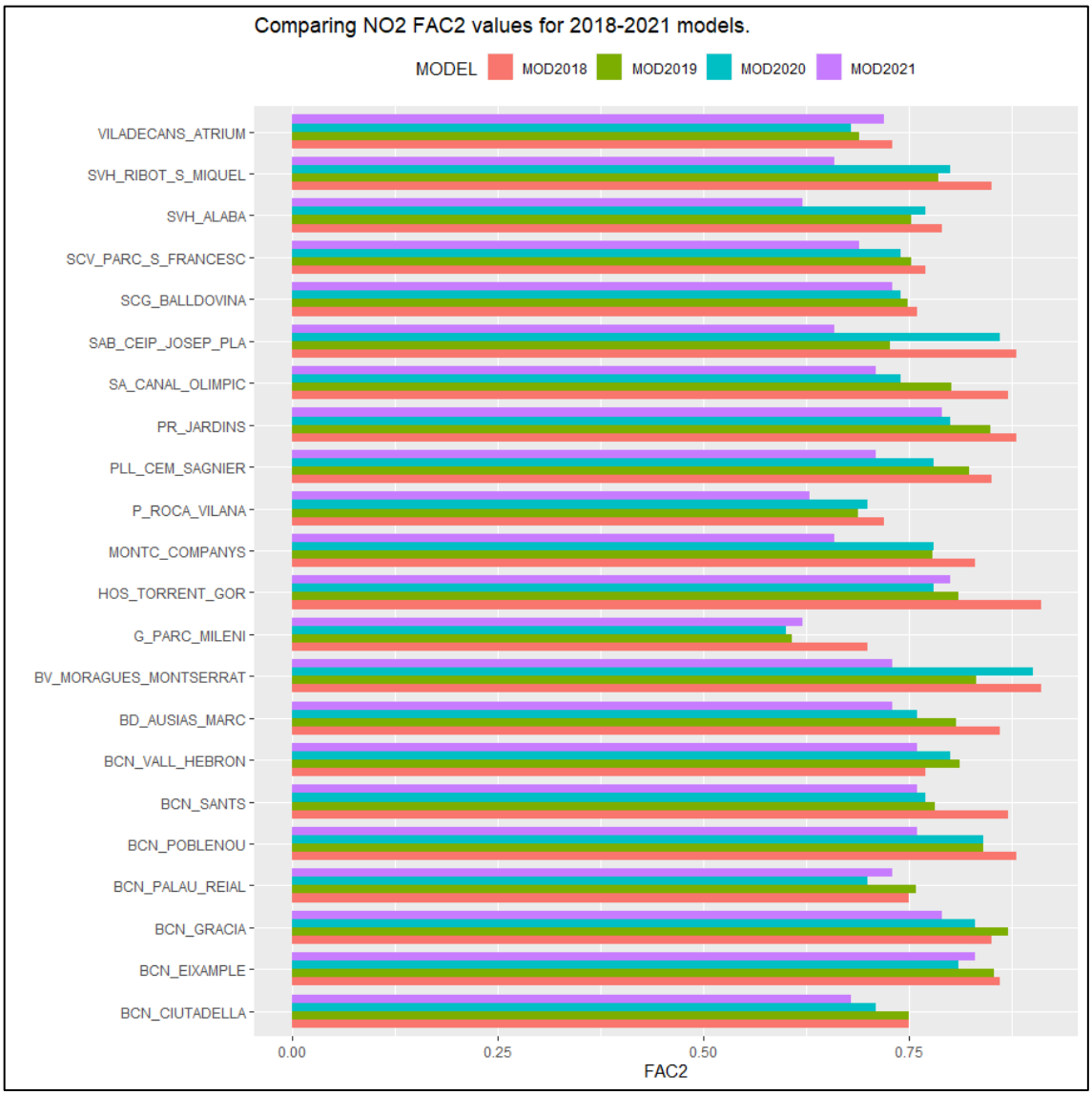
The results of the model indicate, in general, a similar correlation for each station for all the years of analysis, showing its consistency. However, in 2021, the values are probably the lowest in the period (from 2018 to 2021). This may be due to lack of PBLH data (meteorological variable) for 2021. In addition, some emission data for 2021 are still provisional, since statistics on fuel consumption (natural gas) and emissions from large industries have not yet been published. It is very likely that such situation has affected the correlation.

On the other hand, the best correlation results have been obtained in 2020, perhaps due to the lower contribution of traffic due to COVID-19 lockdowns. The variability of urban air pollution has been smoothed out during 2020. In addition, the better emission characterization profiles of the road transport due to the data exploitation from the LEZ cameras has contributed to a more realistic behaviour of the model in that sense.

In conclusion, the hourly correlation shows a good performance and means for the model to represent well the locations characteristics and the variability of the input data, especially the meteorology and the data from the considered emissions inventory.

The following figure shows, for NO_2, the FAC2 factor between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 32. NO_2 FAC2 factor between the results of the model and the real observations of the XVPCA stations, for each year of analysis

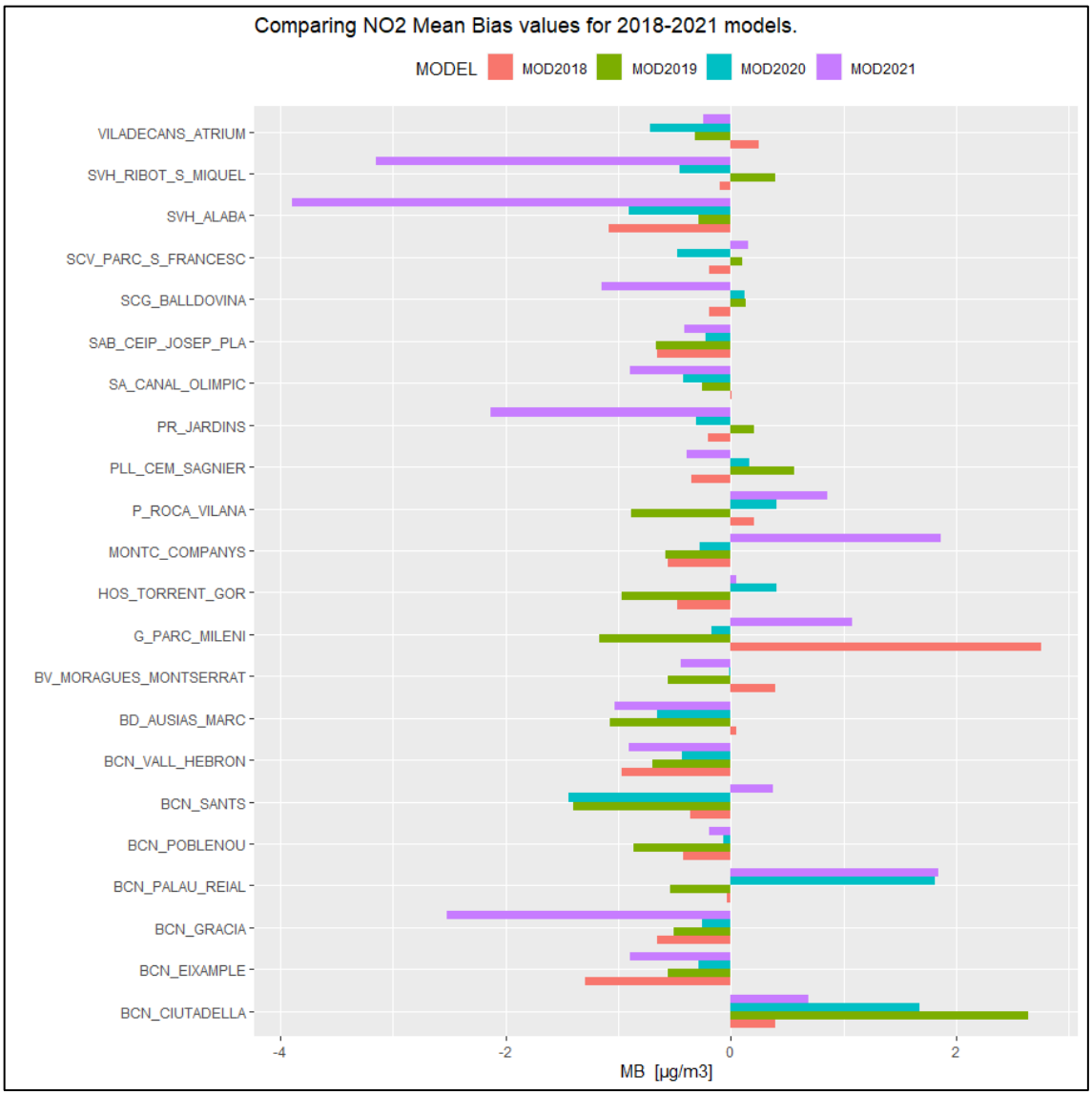


Source: Barcelona Regional

The FAC2 (fraction of predictions within a factor of two of observations) is the most robust measure, because it is not overly influenced by high and low outlier. The acceptable values for this parameter are between 0,5 and 2,0, with 1,0 being the desired value. The graphic below shows that the model has good performance, with values ranging between 0,6 and 0,91.

The following figure shows, for NO_2, the Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 33. NO_2 Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

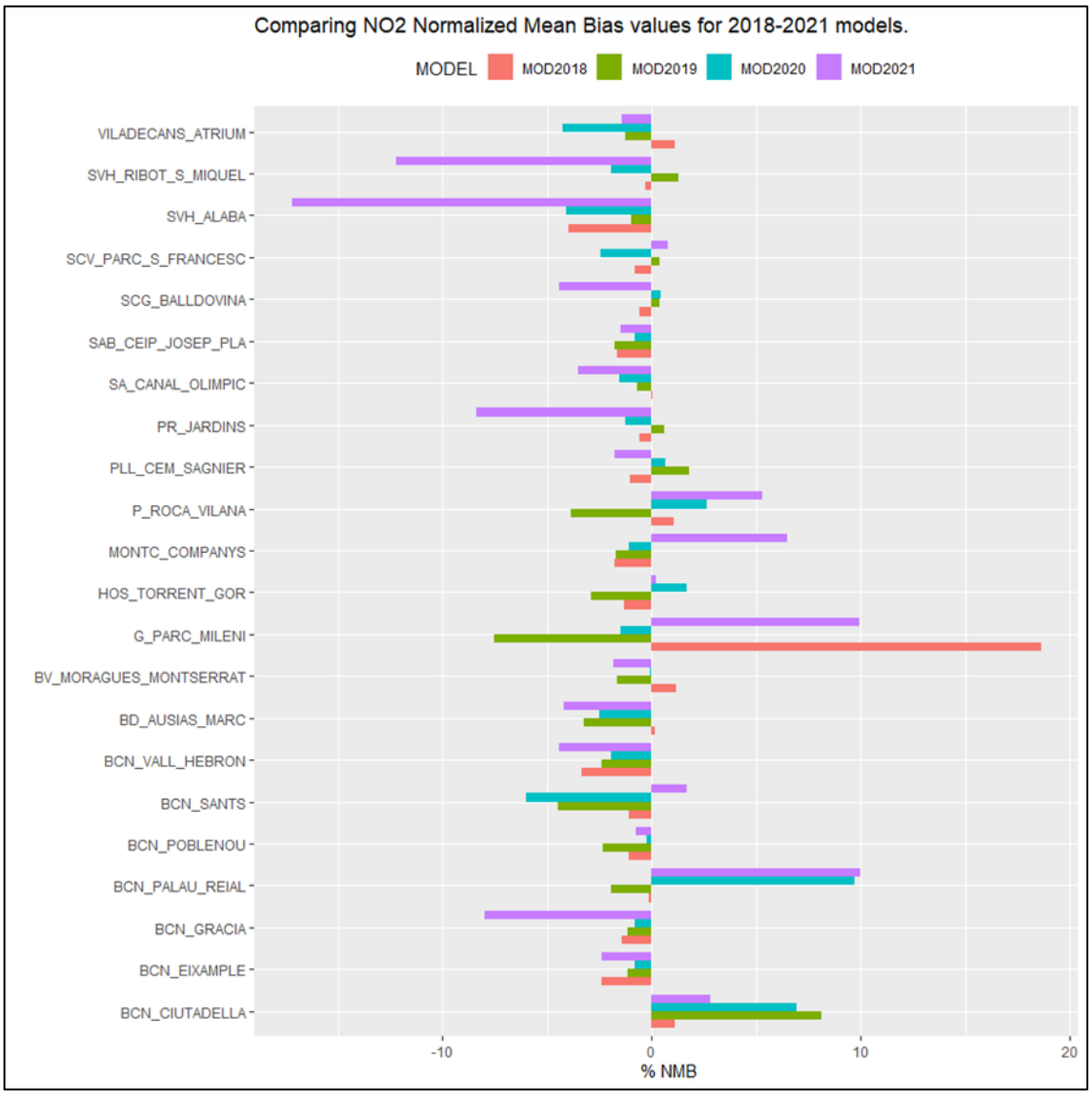


Source: Barcelona Regional

As can be seen in the previous figure, the model tends to underestimate the NO_2 concentrations. The total average is -0,28 µg/m^3 (all years and all stations). However, the station with the highest concentration overestimation is Ciutadella station (located in Barcelona).

The following figure shows, for NO_2, the Normalized Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 34. NO_2 Normalized Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.



Source: Barcelona Regional

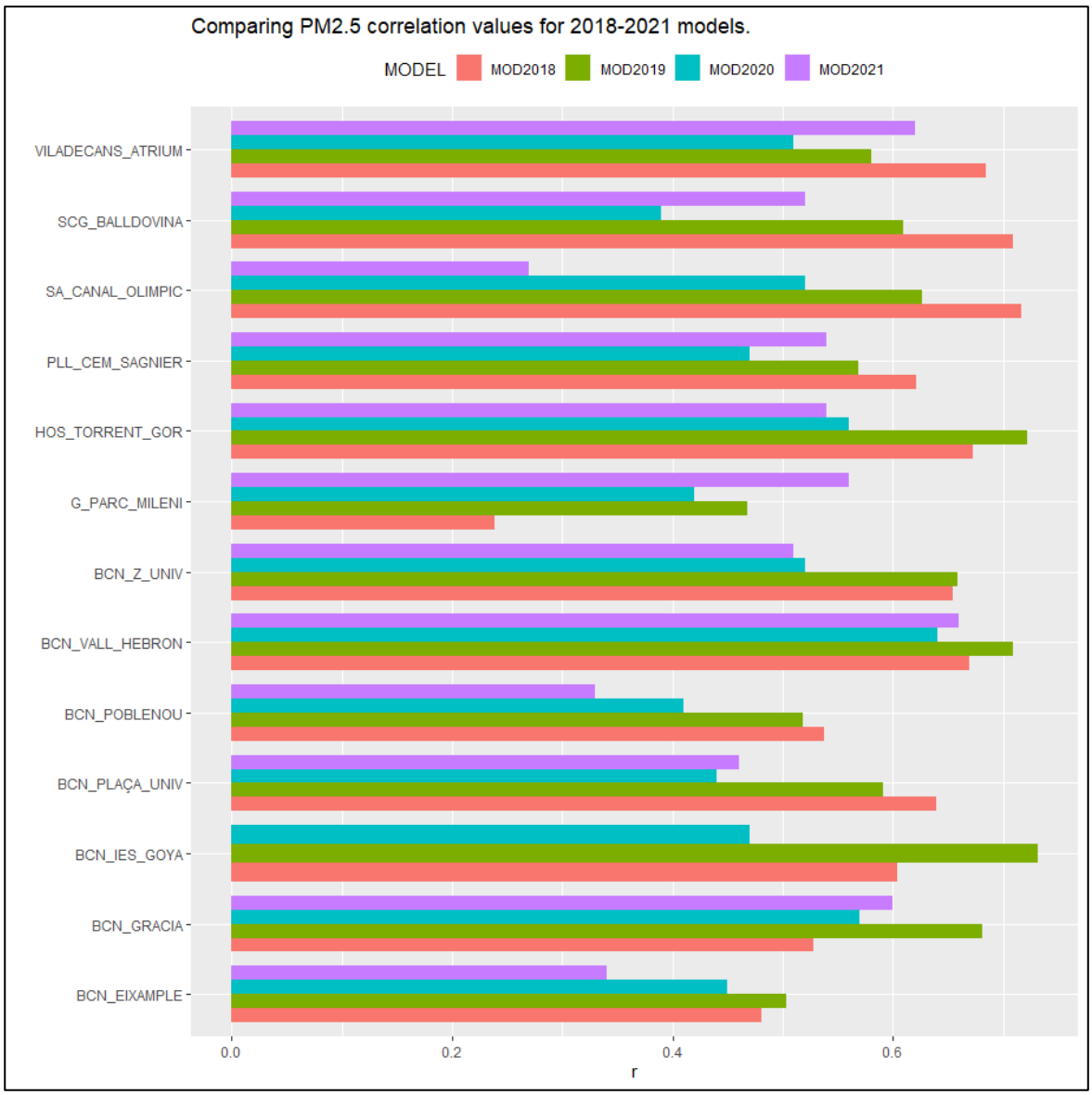
As can be seen in the previous two graphics, the MB is located between -4 µg/m^3 and 3 µg/m^3. As a fraction of observed value, this represents between -20% and 20% of the observed values. This means a good model fit to observed value. The fit is even better if the Air Quality Directive Objectives (AQD, 2008) are considered. Such Directive sets the data quality objective value at 15% for O_3, 15% for NO_2 and 25% for particulate matter (PM_{10}). Note these values are meant as the maximum limit allowed for each pollutant (Thunis et .al, 2012).

$PM_{2,5}$ performance evaluation

In relation to $PM_{2,5}$, the evaluation is based in daily measurement data, since XVPCA stations only measure this pollutant in daily basis.

The following figure shows, for $PM_{2,5}$, the correlation level (r parameter) between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 35. $PM_{2,5}$ correlation level (r parameter) between the results of the model and the real observations of the XVPCA stations, for each year of analysis (2018 to 2021).



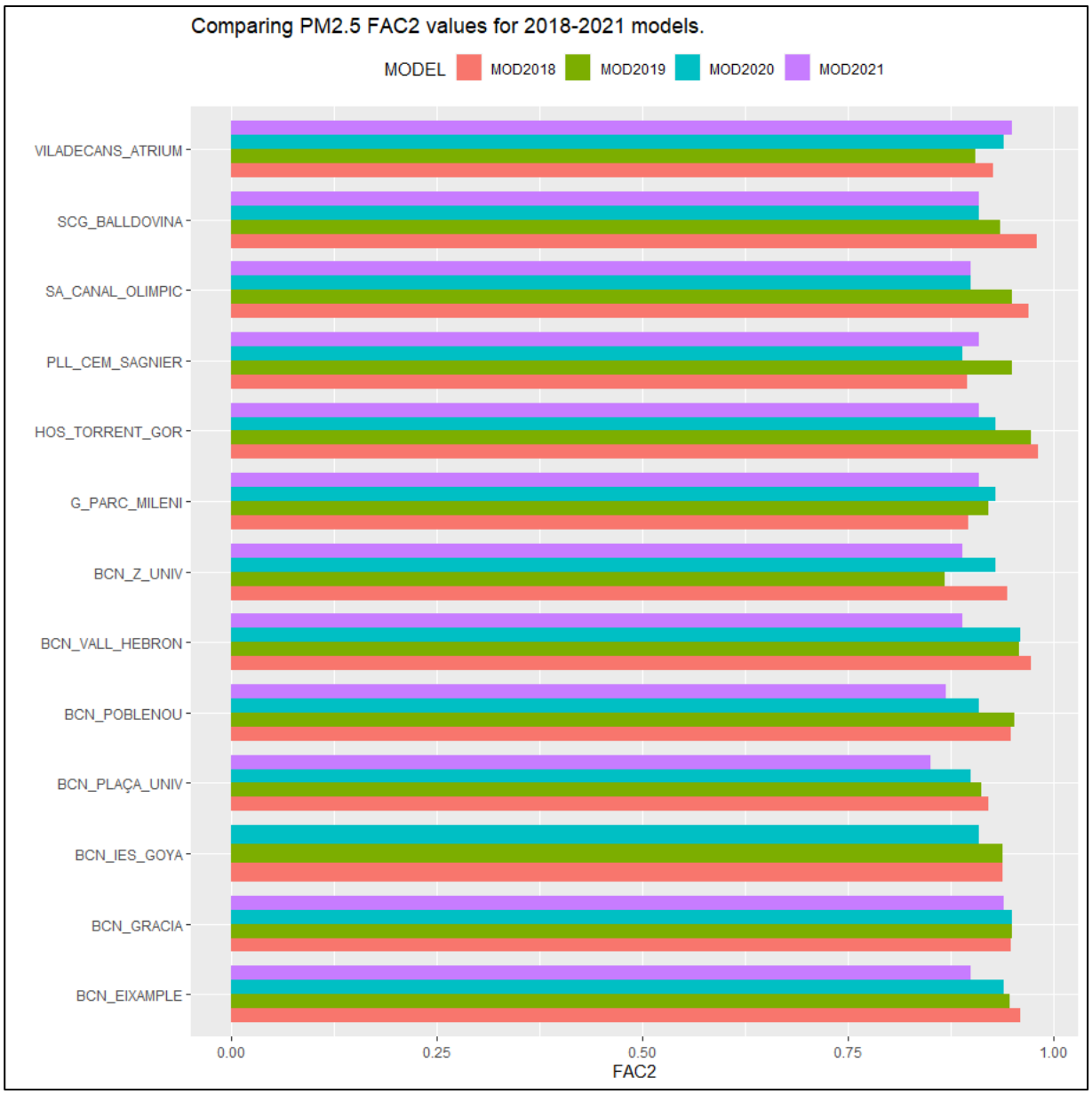
Source: Barcelona Regional

In general, the correlation is between 0,25 and 0,65. The average of the r parameter for all years is 0,55. Parc del Mil·leni is, again, the station with the poorest performance (2018). Vall d'Hebron is a background station that has, on average, the best performance for all years for this. This may be

due to the influence of the whole city of Barcelona and not only the direct influence of the road traffic, as happens in other traffic stations.

The following figure shows, for $PM_{2,5}$, the correlation level (r parameter) between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 36. $PM_{2,5}$ FAC2 factor between the results of the model and the real observations of the XVPCA stations, for each year of analysis

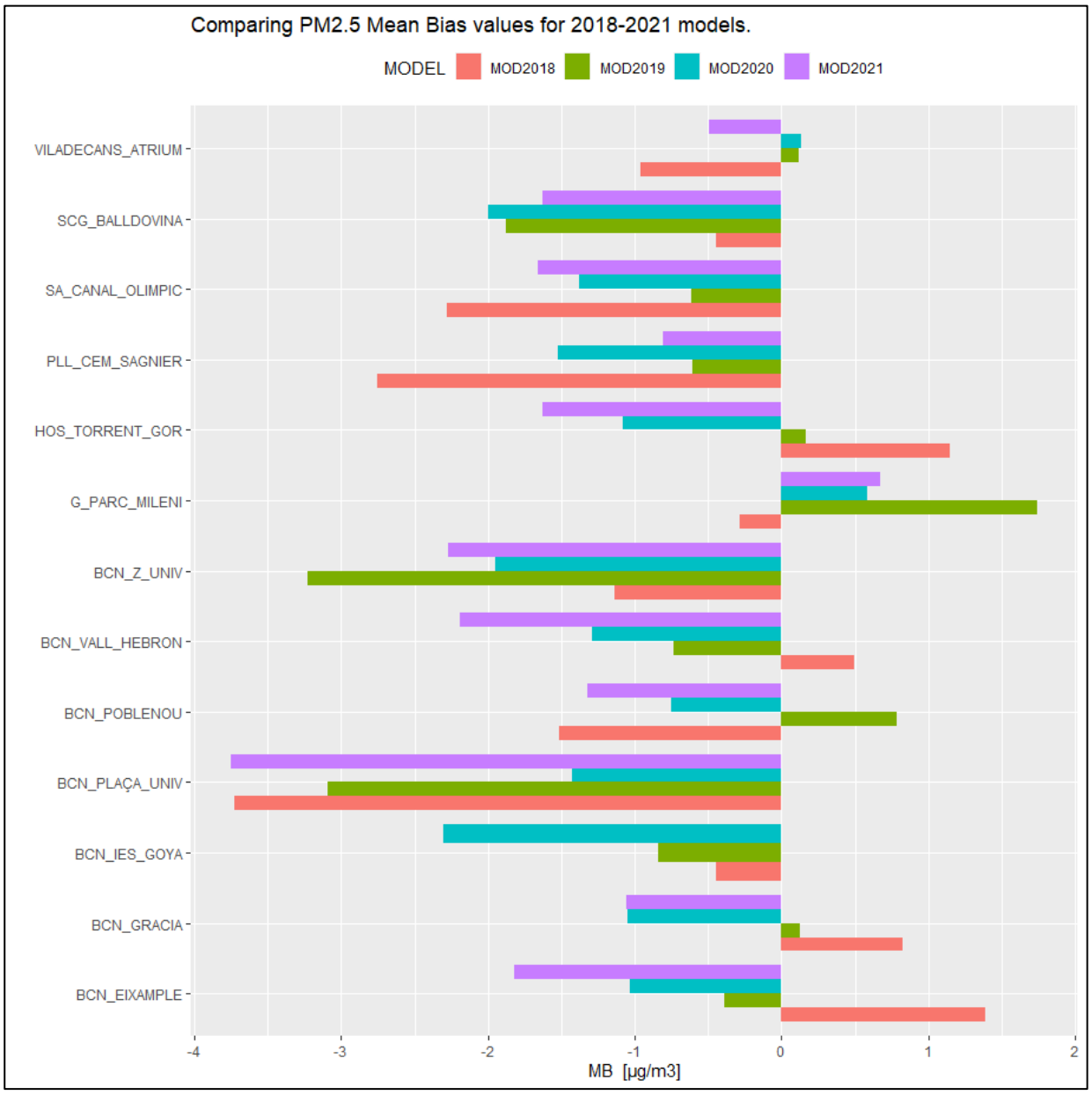


Source: Barcelona Regional

The FAC2 shows good performance for $PM_{2,5}$ for all stations, near to 1, maybe because the measurements are in daily basis.

The following figure shows, for $PM_{2,5}$, the Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 37. $PM_{2,5}$ Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.



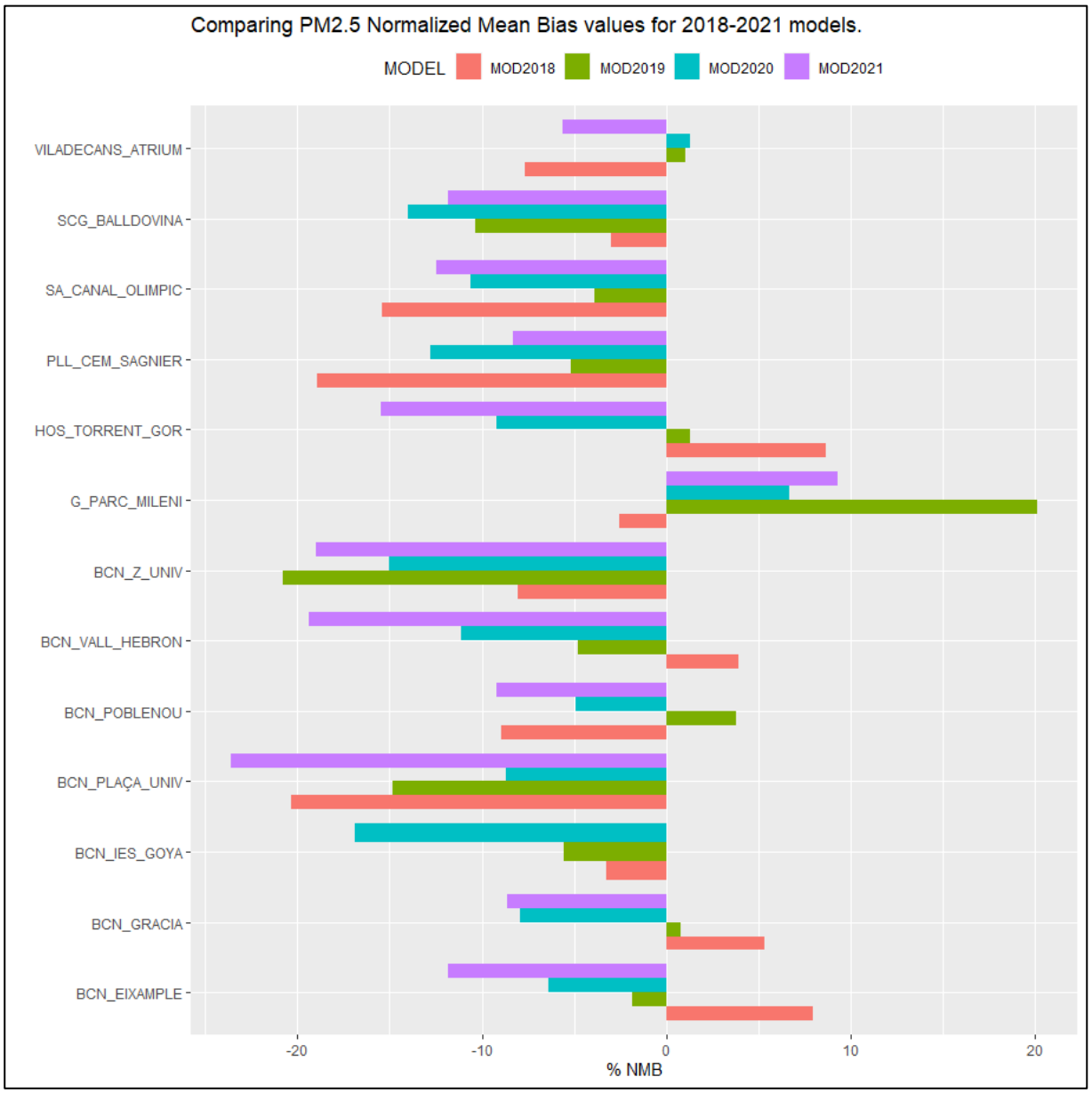
Source: Barcelona Regional

The model tends to underestimate the concentration of $PM_{2,5}$, with values between -3,6 µg/m^3 and 1,8 µg/m^3 in average. The MB is -1 µg/m^3 for all years and stations

As can be seen in the last figure, Plaça Universitat (located in Barcelona) has the poorest results for the MB (-3 µg/m^3 on average). Maybe the emissions are not well characterized for this location and an emission checking is needed, since there are several bus stops that can be affecting the model results.

The following figure shows, for $PM_{2,5}$, Normalized Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.

Figure 38. $PM_{2,5}$ Normalized Mean Bias between the results of the model and the real observations of the XVPCA stations, for each year of analysis.



Source: Barcelona Regional

Referring to the NMB, its value ranges between -20 and 20% for practically all years and all stations (except Plaça Universitat in 2021 and Zona Universitària in 2019). In this sense, the model shows good behaviour for this pollutant too.

2. PREGNANT LOCATION DATA PROCESSING

The general methodology is based on modelling the air pollution exposure for each of the participants of the project, according to her location at each time period.

Geolocation data by participants at different time periods are needed to feed the air quality model. Two different approaches have been considered. On one hand, GPS data have been used for some time periods in order to know the participant's location at each time of the day (ExpoApp[16] data and commuting). On the other hand, in some cases, commuting routes and home locations have been located with GIS tools according to the explanation for each participant (surveys).

This chapter is focused on the first approach (GPS data), since the second data and reference points are directly introduced to the model by means of GIS tools.

2.1. Commute and ExpoApp data processing

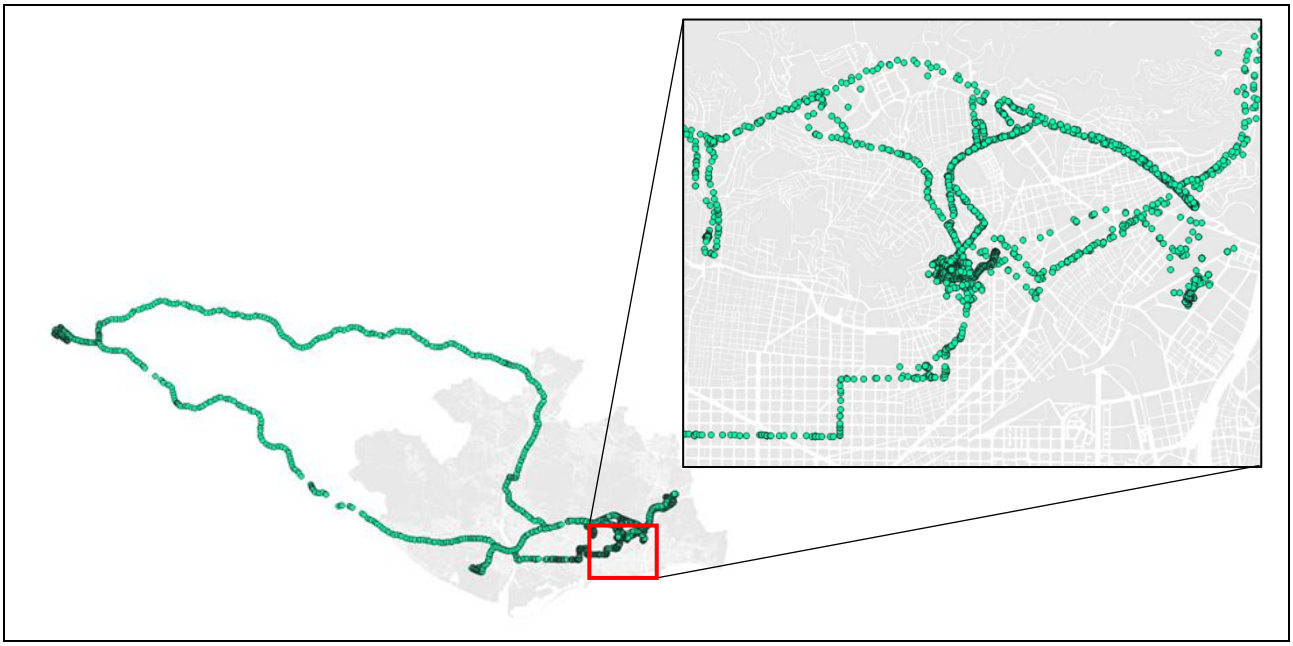
Mainly, GPS data has been used. Nevertheless, it generates a large volume of data that is needed to clean and process in order to reduce the number of points feed to the air quality model properly. In this sense, this chapter exposes the followed methodology.

2.1.1. Input data

Input data is mainly a set of GPS traces provided by GPS devices and mobile phones from the participant subjects in the study. These GPS traces were captured for one week, at different time periods. The figure below shows, for a specific participant, an overview of the whole extent (and zoomed view) of the input GPS data of her route.

[16] App to track geolocation by means of GPS.

Figure 39. Example of raw GPS data for a participant (general overview and zoom over Barcelona and surrounding areas)



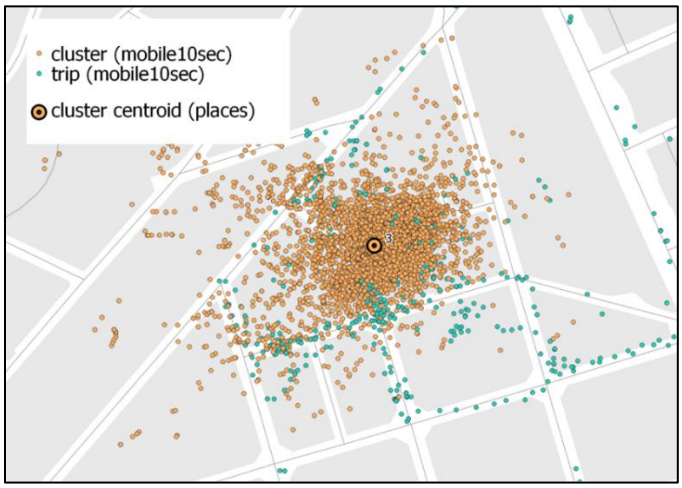
Source: Barcelona Regional

The GPS data is structured in two main tables or layers:

- ✓ **Mobile10sec**: original data resampled every 10 seconds. There are about 38.000 records.
- ✓ **Places**: the centroids of the identified clusters (places where the subject stayed for a long time)

Next figure shows, for a certain analysis zone, such structure of the data.

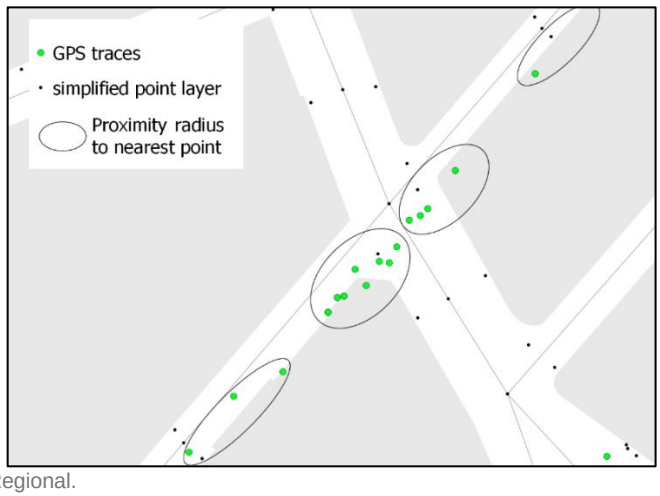
Figure 40. Types of GPS data.



Source: Barcelona Regional

To reduce the number of points for modelling, a technique has been implemented to assign the GPS traces (green dots) to the nearest point of the Simplified Point Layer (black dots), using a set of rules that are described in the following lines. Next figure shows both GPS traces and SPL's (Simplified Point Layers).

Figure 41. Aggregation of data: GPS traces and SPL's (Simplified Point Layers).



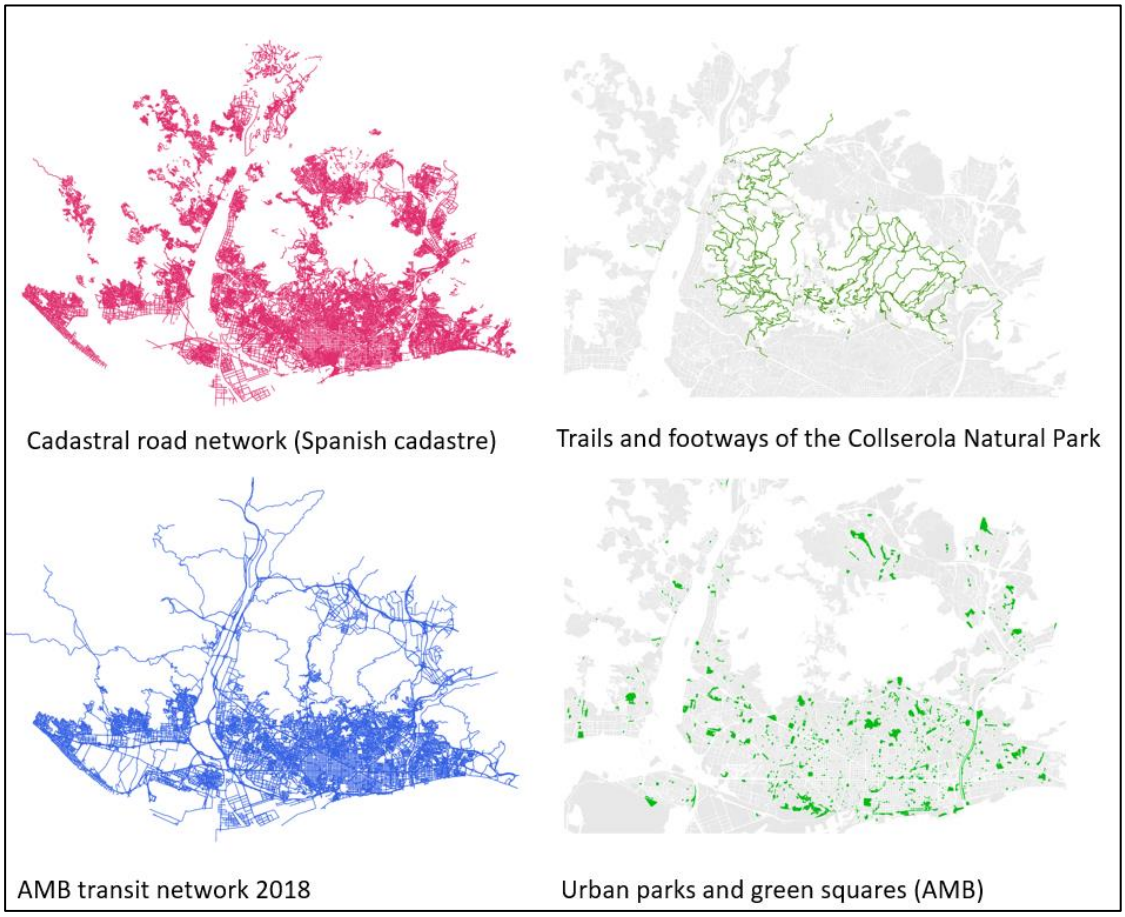
Source: Barcelona Regional.

2.1.2. Simplified Point Layer

To generalize input data, a Simplified Point Layer (SPL), which represents the road network and other walkable areas, has been created.

This SPL has been key to reduce the amount of points. In order to create them, a set of layers have been utilised (shown in the following picture).

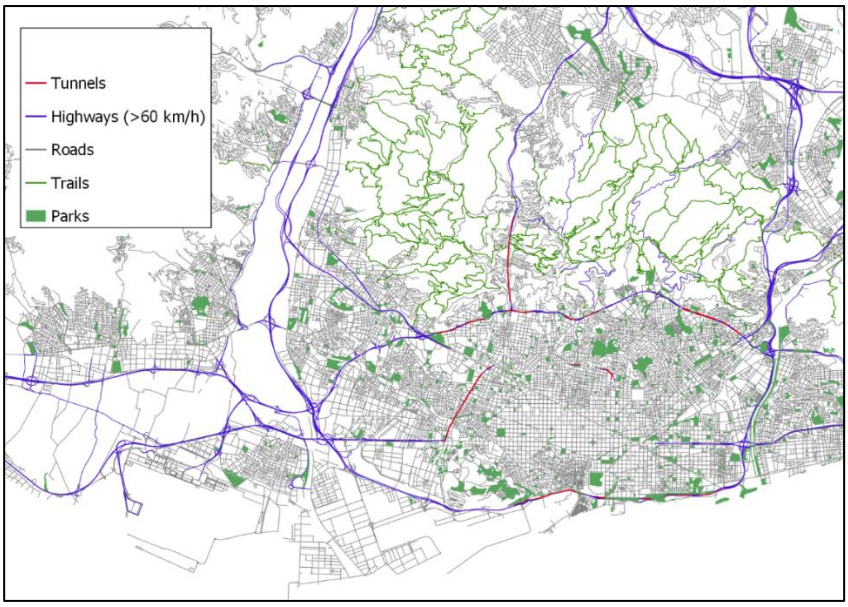
Figure 42. Data used to generate the Simplified Point Layer (SPL).



Source: Barcelona Regional.

Once the layer was merged, a SPL has been used to assign the GPS traces, shown in next figure.

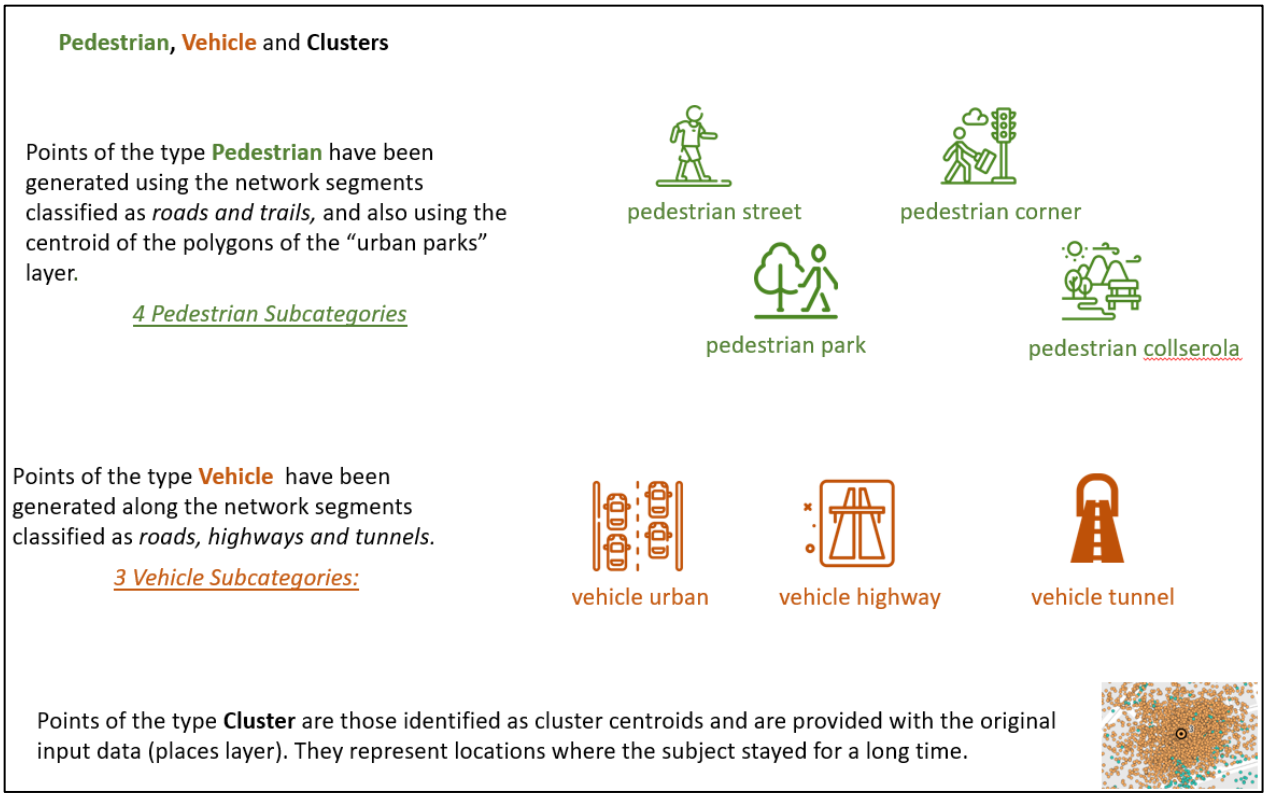
Figure 43. Simplified Point Layer (SPL), made up of several sublayers.



Source: Barcelona Regional.

The SPL points have been classified in three main categories: pedestrian, vehicle and clusters. Such categories are shown in the following figure.

Figure 44. Point classification categories according SPL.



Source: Barcelona Regional.

The considered criteria for different locations is described below.

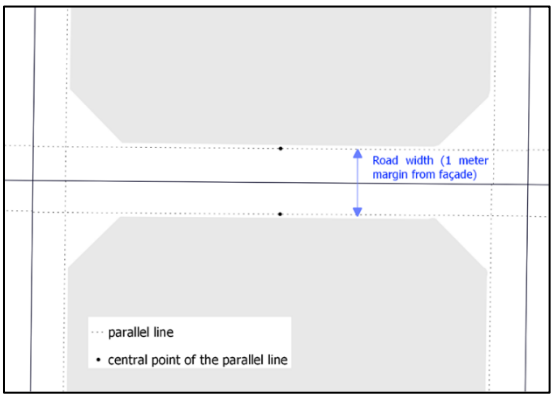
Pedestrian street

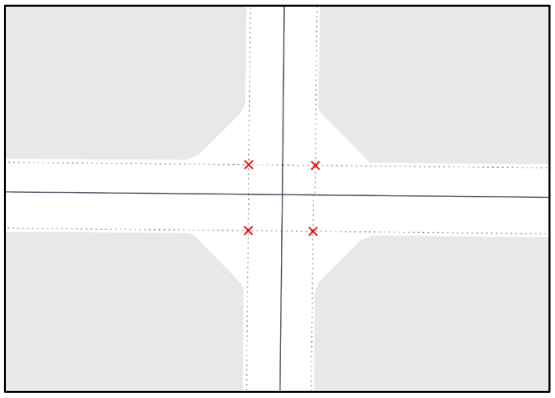
A parallel line has been generated on both sides of the road, 1 meter away from the building façade.

The point is generated at the centre of this line, where the sidewalk would be.

Road segments longer than 150 metres have been divided each 100 meters.

No line has been generated if the space between the facade and the road segment is < 2 meters.



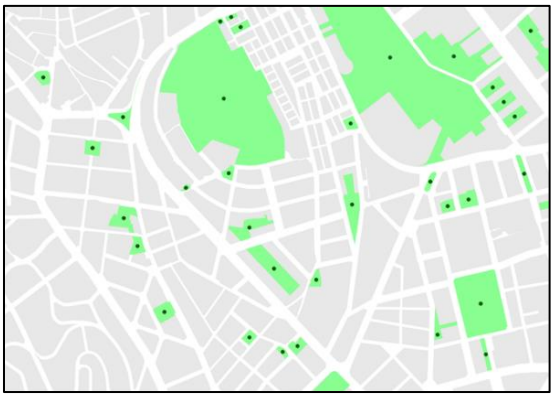


Pedestrian corner

The points classified as pedestrian corner are generated at the intersection of the parallel lines created in the previous step.

Pedestrian park

Pedestrian parks: central point of the park geometry (inside).



Pedestrian Collserola

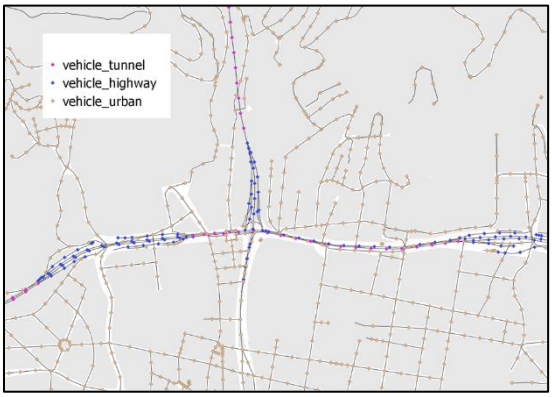
Pedestrian Collserola: points generated every 50 meters along the trails.

Vehicle urban, vehicle highway and vehicle tunnel

All points of type *vehicle* have been generated using the central point of each road segment and its ends.

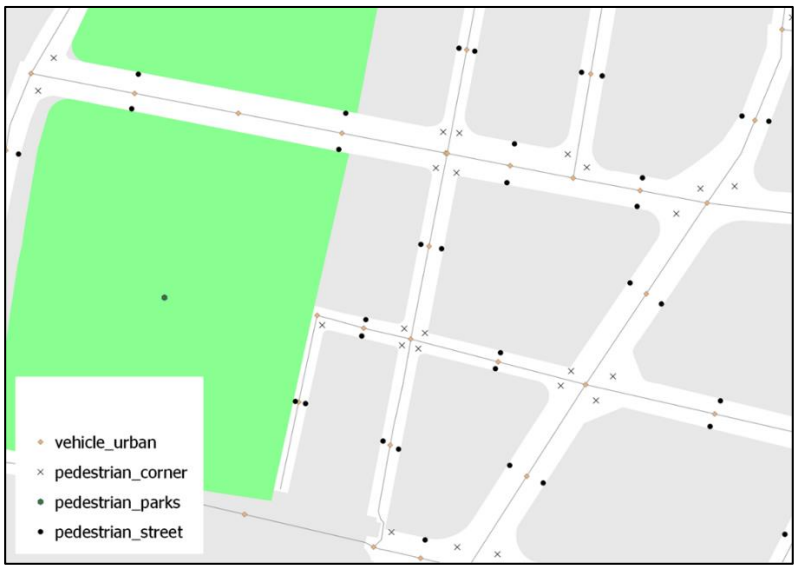
Road segments > 150 metres have been divided every 100 meters.

Subcategories are assigned using the road type of the section (tunnel, highway and urban).



As an example, a detailed view of the final SPL is shown in the following picture.

Figure 45. Detailed view of a portion of the final SPL.



*The other point types that are missing in this image are *pedestrian_collserola*, *vehicle_highway, vehicle_tunnel and cluster_long_stay*

Source: Barcelona Regional.

2.1.3. Processing the GPS traces input data

The process has been divided in four phases:

1. Cleaning and processing the GPS traces to compute travel time and travel speed.

2. Classification of the GPS traces into these following categories:

   - ✓ **Cluster:** trace identified as cluster.
   - ✓ **Out of domain:** trace location is out of bounds of the study area.
   - ✓ **Pedestrian:** walking trip according to travel speed.
   - ✓ **Vehicle:** vehicle trip according to travel speed.

3. Assignment by proximity of the GPS traces classified as *Pedestrian* or *Vehicle* to the Simplified Point Layer

4. Grouping of the GPS traces and aggregation.

2.1.3.1. Cleaning and processing the GPS traces to compute travel time and travel speed

Geolocation precision of each GPS trace is recorded in the “accuracy” field, which expresses the margin of error as radius (in meters). Physical obstacles, such as buildings or trees, may cause some specific low accuracy GPS readings. In order to perform more precise travel speed calculations (and improve overall precision), all GPS traces with an “accuracy” < 30 meters have been removed from the analysis.

Another event that has an impact in accuracy is when the subject is travelling by Metro or other underground means of transport. To identify those situations, a threshold of 5 minutes of consecutive low accuracy readings has been used. In case this threshold is exceeded, all the time spent in this situation is aggregated in a single record and stored under the field “is_undergroud”.

With these two approaches, nearly 20% of the original GPS traces were removed or marked as “is_undergroud” in the sample subject example.

- ✓ 38.097 original records.
- ✓ 29.938 after filtering

Once the records are filtered, travel time and travel speed are calculated.

2.1.3.2. Classification of the GPS traces by categories:

✓ **Clusters**

Cluster GPS traces are already identified in the original data, but there are many records classified as *trip* that could be reclassified as *cluster*, as seen in the following example.

Table 10. Example of cluster data reclassification.

gps_trace_ID	type	timestamp	seconds
231	cluster	17:24:13	20
250	cluster	17:24:33	20
266	cluster	17:24:53	20
285	trip	17:25:13	16
287	cluster	17:25:29	4
301	trip	17:25:33	11
308	cluster	17:25:44	9
323	trip	17:25:53	12
329	cluster	17:26:05	8
341	trip	17:26:13	9
351	trip	17:26:22	9
367	trip	17:26:31	16
375	trip	17:26:47	7
385	trip	17:26:54	10
397	trip	17:27:04	11
410	trip	17:27:15	13
413	cluster	17:27:28	6

These records clearly could be reclassified as *cluster* because the trips only last a few seconds

This trip only lasts 75 seconds before "re-entering" the cluster again

gps_trace_ID	type	timestamp	seconds	new_type
231	cluster	17:24:13	20	cluster
250	cluster	17:24:33	20	cluster
266	cluster	17:24:53	20	cluster
285	trip	17:25:13	16	cluster
287	cluster	17:25:29	4	cluster
301	trip	17:25:33	11	cluster
308	cluster	17:25:44	9	cluster
323	trip	17:25:53	12	cluster
329	cluster	17:26:05	8	cluster
341	trip	17:26:13	9	cluster
351	trip	17:26:22	9	cluster
367	trip	17:26:31	16	cluster
375	trip	17:26:47	7	cluster
385	trip	17:26:54	10	cluster
397	trip	17:27:04	11	cluster
410	trip	17:27:15	13	cluster
413	cluster	17:27:28	6	cluster

To reclassify these records, a threshold of 300 seconds (5 minutes) has been established.

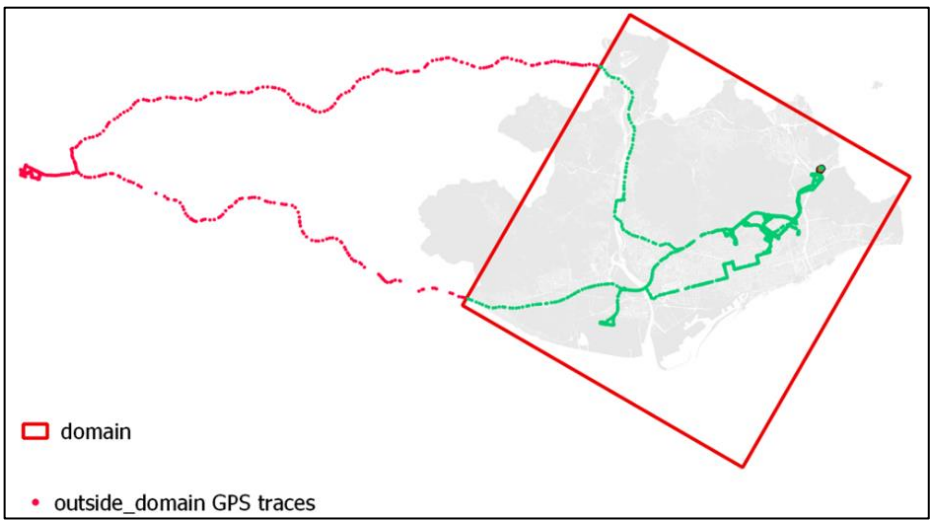
All the trips shorter than this amount of time are reclassified as cluster

Source: Barcelona Regional.

✓ **Outside domain**

This classification is given to the GPS traces that fall outside of the bounds of the study area. An example of this case is presented in the figure below.

Figure 46. Example of GPS data measurement domain.



Source: Barcelona Regional.

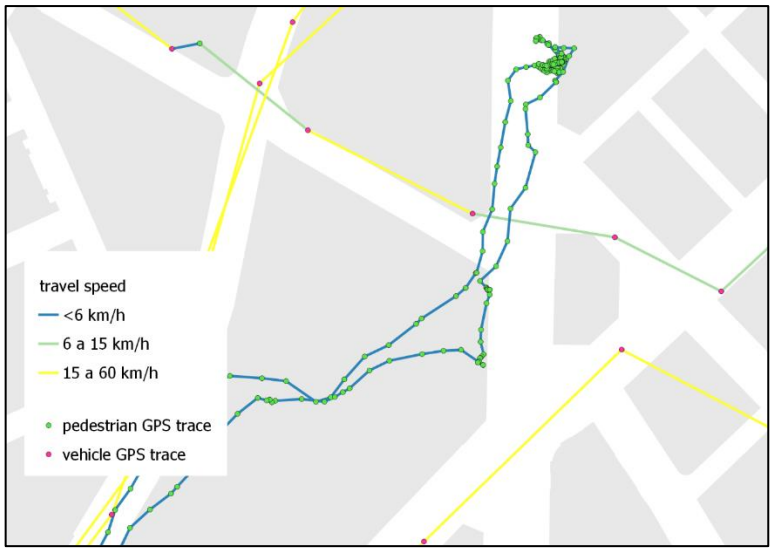
✓ **Seed travel classification for pedestrian and vehicle**

GPS traces are identified as “pedestrian” or “vehicle” according to travel speed.

- If travel speed < 6 km/h, GPS trace is considered as pedestrian.
- If travel speed > 6 km/h, GPS trace is considered as vehicle.

The following figure shows, for a certain region, an example of the speed calculation results and GPS trace assignation along urban ways.

Figure 47. GPS traces identification according to travel speed (“pedestrian” or “vehicle”).

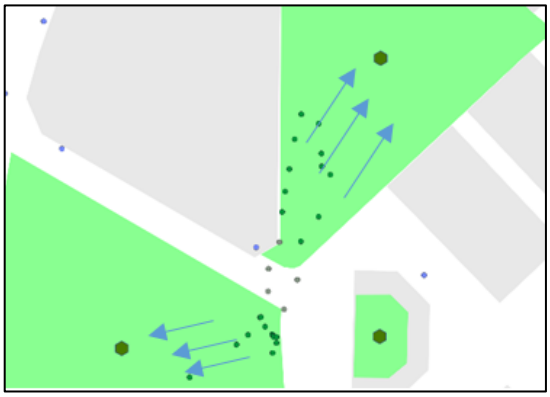


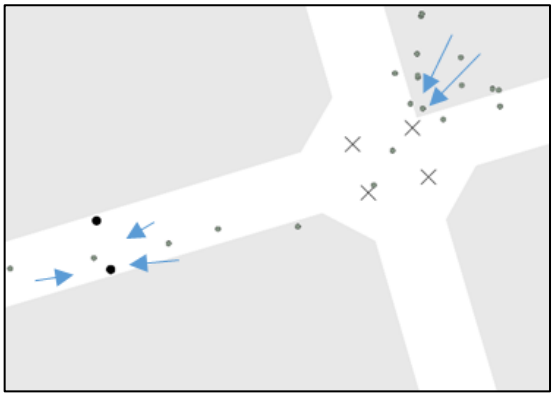
Source: Barcelona Regional.

2.1.3.3. Assignment by proximity to the Simplified Point Layer of the GPS traces classified as Pedestrian or Vehicle

Pedestrian and *Vehicle* GPS traces are assigned to the points of the Simplified Point Layer by using the following rules, applied in the presented order:

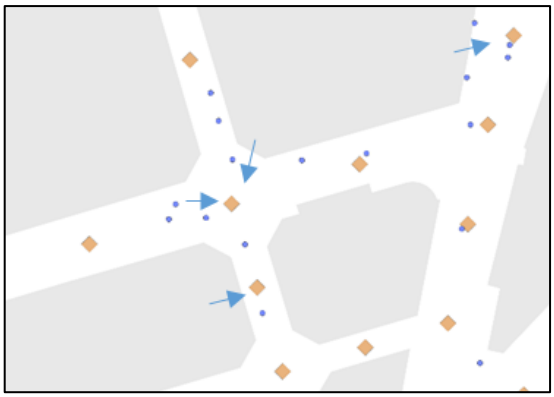
Pedestrian GPS traces that fall within the boundaries of a park are assigned to the centroid of that park.





Remaining pedestrian GPS traces are assigned to the closest point of the type *pedestrian* within a distance of 50 meters. (If no point of that type is found at that distance, then any point is used)

Vehicle GPS traces are assigned to the closest point of the type *vehicle*



2.1.3.4. Grouping of the GPS traces and aggregation

GPS traces are grouped according to the point id of SPL they have been assigned to, and also taking into account the time sequence of the GPS traces itself.

Time sequence has to be considered because a same point of the SPL can have several GPS data assigned to different dates (e.g. points generated near the home of the participant subject, back and forth trips… etc). Once the groups are identified, data is aggregated.

2.1.4. GPS trace processing results

Through this methodology 38.079 records in the original input data have been processed and 2.298 records after aggregation have been obtained (which means 94% of the records reduced).

The following table shows the output format of the simplified file.

Table 11. Output format for the simplified file.

Group ID	id of the group identified used for aggregating the GPS traces
weekday	number (1-7) representing the day of the week
Point ID	id of the point the GPS traces were assigned to
Cluster ID	id of the cluster, if GPS traces belonged to a cluster
GPS count	number of aggregated GPS traces
point type	point type (pedestrian_street, pedestrian_corner, vehicle_urban… etc)
FIRST_localdate	earliest date of the group
LAST_localdate	latest date of the group
seconds_sum	sum of the total time of the group, in seconds

Source: Barcelona Regional.

Finally, the output file with results by pregnant are shown as follows.

Table 12. Output file header example.

Group ID	weekday	Point ID	Cluster ID	GPS count	point type	x_coord	y_coord	FIRST_localdate	LAST_localdate	seconds_sum
1	1	416724	3	1588	cluster_long_stay	430921.662	4586217.025	19/11/2018 17:14:29	19/11/2018 23:59:59	24330
1	2	416724	3	1561	cluster_long_stay	430921.662	4586217.025	19/11/2018 23:59:59	20/11/2018 8:47:26	31647
2	2	95777	-	2	pedestrian_street	430957.8197	4586185.38	20/11/2018 8:47:26	20/11/2018 8:47:42	16

Source: Barcelona Regional.

2.2. Other location data processing

The other type of location data used to model the air quality exposure of the pregnant control group comes mainly from home (of participants) data location, some commuting and reference locations. This type of data has been introduced to the model utilising GIS tools.

3. PREGNANT CONTROL GROUP AIR POLLUTION EXPOSURE MODELLING RESULTS

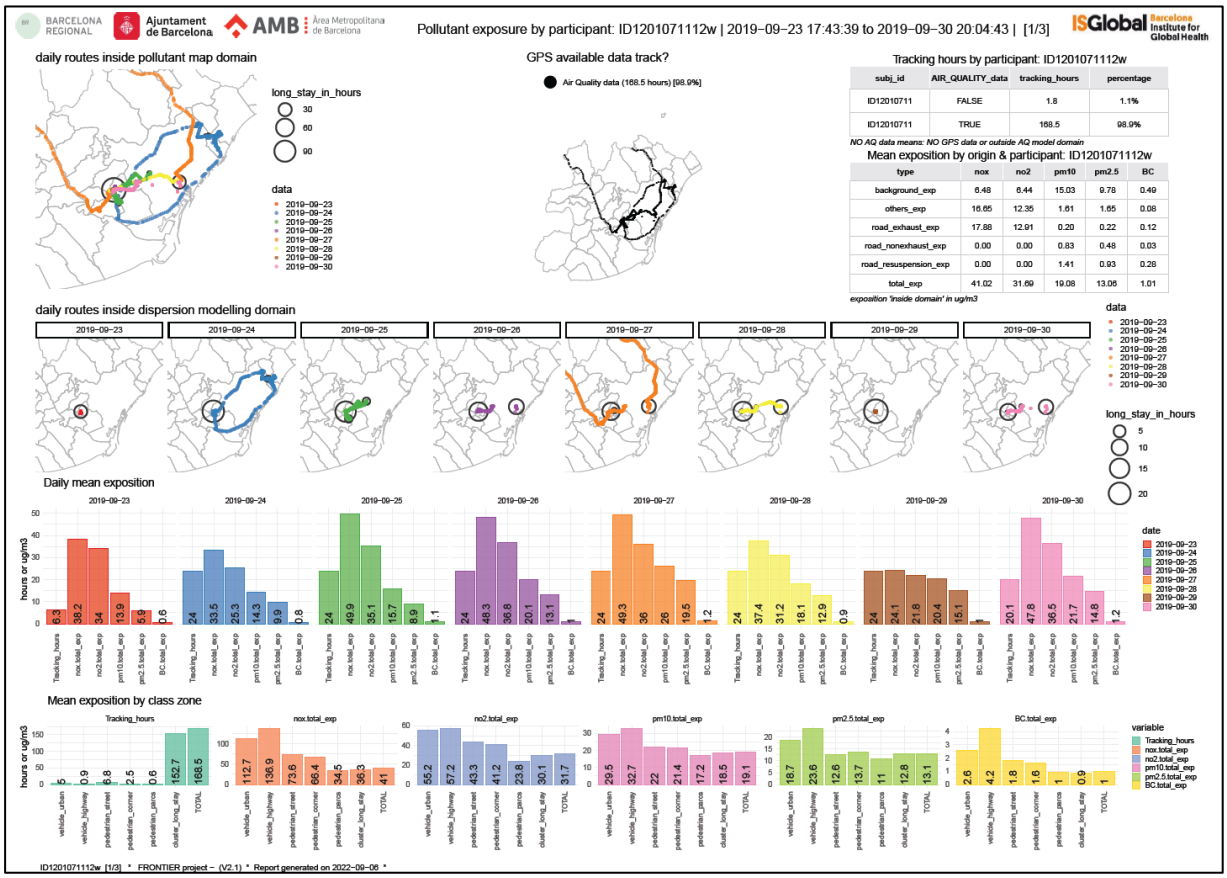
This chapter shows a sample of the main results once the air quality modelling is performed, using the control sample of pregnant at the period 2018 -2021 at the Metropolitan Area of Barcelona.

3.1. Results by participant

The following figure presents an example of the commute results for a specific participant. The dashboard shows a sample of the daily route based on GPS track data for a week in 2019. The hourly pollutant exposition was estimated for different emissions sources and considering pedestrian and vehicle receptor types. The mean pollutant exposition for the period is summarized in the table.

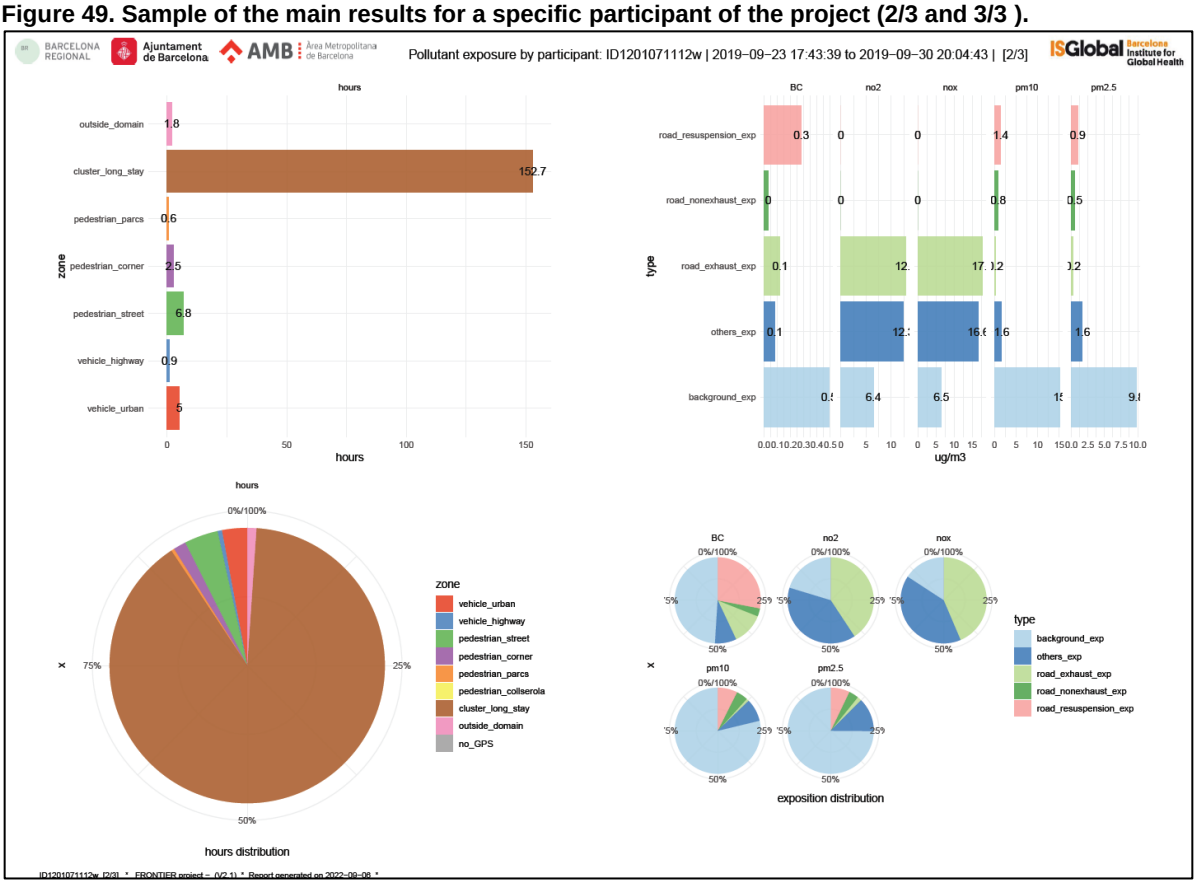
This kind of result has been process for each participant by type of input data (1084 for commuting, 978 for ExpoApp, 1083 homes and 36 reference points).

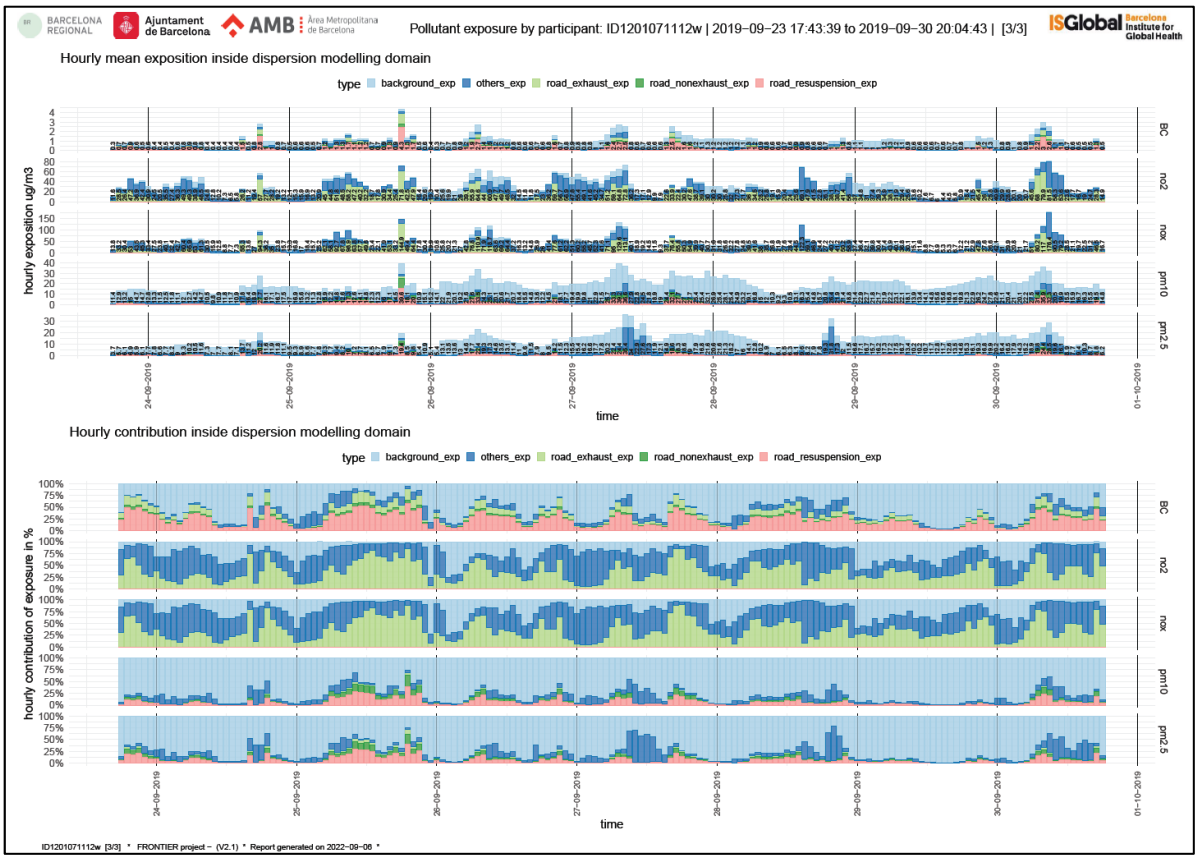
Figure 48. Sample of the main results for a specific participant of the project (1/3).



Source: Barcelona Regional.

Figure 49. Sample of the main results for a specific participant of the project (2/3 and 3/3).





Source: Barcelona Regional

3.2. Average results by input data type

The table shows a summary of the average values for the mean and the standard deviation values for the total of pregnant control group modelling. The results are shows according to the input data type.

Table 13. Summary of the average values for the mean and the standard deviation for the total of pregnant modelling (µg/m^3).

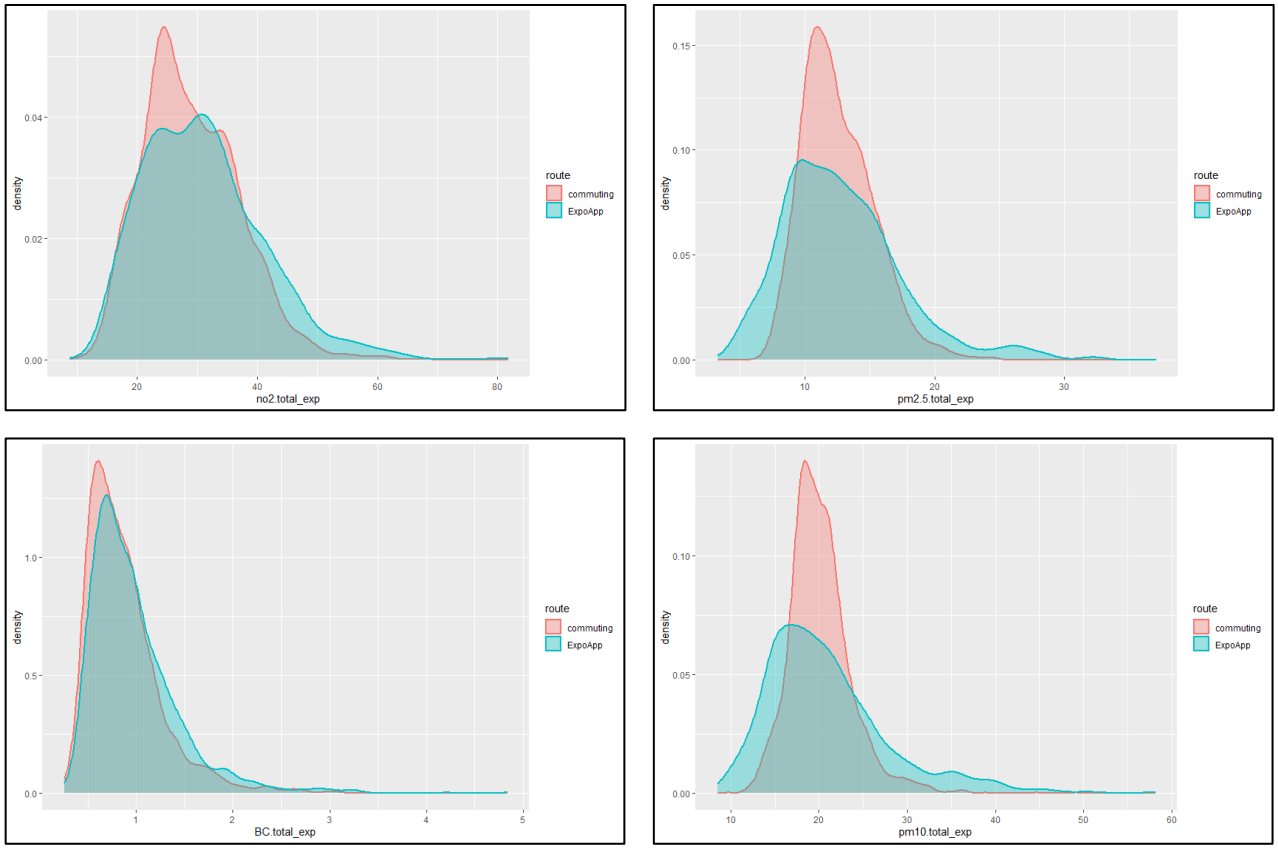
Route/Place	N_SUBJ(#)	NO_2_MEAN	NO_2_SD	PM25_MEAN	PM25_SD	BC_MEAN	BC_SD	PM_{10}_MEAN	PM_{10}_SD
Commuting	1084	29,1	8,10	12,6	2,80	0,90	0,40	20,2	3,50
ExpoApp	978	31,0	10,1	12,9	4,70	1,00	0,50	20,9	7,00
Homes	1083	29,0	19,1	12,5	7,78	0,88	0,74	19,9	12,1
Reference	36	27,5	18,3	12,3	7,80	1,10	1,00	22,6	12,0

Source: Barcelona Regional. N_SUBJ(#): amount of participants.

As can be seen, NO_2 mean concentration is between 27,5 µg/m^3 and 31 µg/m^3. For $PM_{2,5}$, concentration ranges between 12,3 µg/m^3 and 12,6 µg/m^3. In relation to BC, the concentration lies between 0,88 µg/m^3 and 1,0 µg/m^3. As for PM_{10}, concentration is between 19,9 µg/m^3 and 20,9 µg/m^3, depending of the route/place data type. Still, the range of results is similar for all data types.

The density distribution for the different pollutants (NO_2, $PM_{2,5}$, PM_{10} and BC) for the commuting and ExpoApp data is shown in the following graphic.

Figure 50. Density distribution for NO_2, $PM_{2,5}$, PM_{10} and BC for the commuting and ExpoApp data.

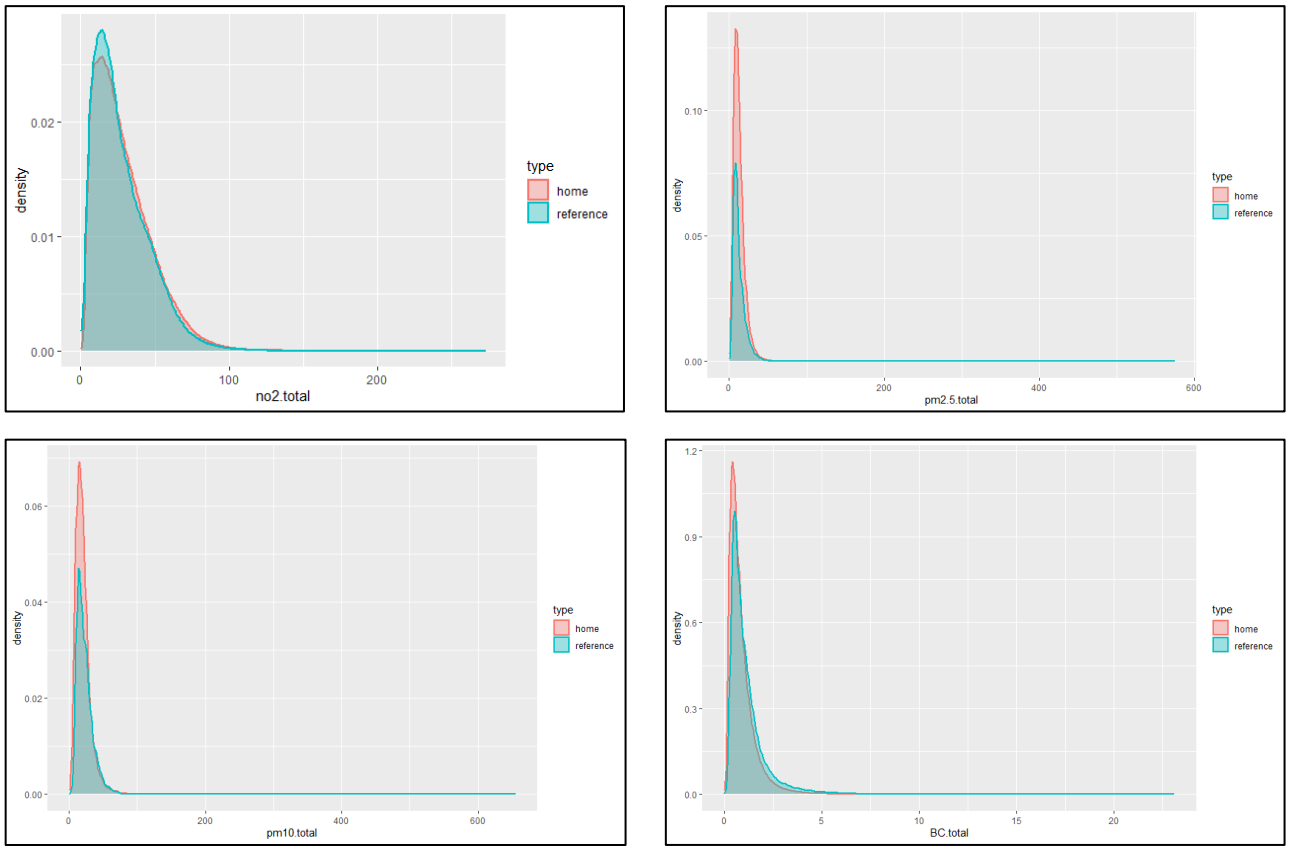


Source: Barcelona Regional

The graphics below present the density distribution for the different pollutants (NO_2, $PM_{2,5}$, PM_{10} and BC) for participant homes and other reference points.

The density distribution shows, for commuting routes, that the less dispersion of the results is referred to the ExpoApp routes. This is maybe due to the amount of modelling hours for commuting, being 20 times greater than ExpoApp data.

Figure 51. Density distribution for NO_2, $PM_{2,5}$, PM_{10} and BC for homes and reference data.



Source: Barcelona Regional

3.3. Computational cost

Computational cost is referred to the execution time that simulation have taken. In this sense, only the ADMS Urban execution time has been considered.

45600 hours CPU have been spent at 56 CPUs in parallel (34 days), mainly for commuting routes (72% of total execution time) and ExpoApp (25% of total execution time). Remaining time has been for home and reference points simulation

The simulation have been carried out in a cluster:

- Window 10 Operating System
- Machines with CPU AMD EPYC 7302 at 2,99 GHz.

4. CONCLUSIONS

In general conclusion, **the model exhibits good performance** for NO_2 and $PM_{2,5}$ for all years (from 2018 to 2021) in most air quality station locations, for BC the available measurement is limited to two stations and two years, however, the results have a good performance.

In average the main statistics in hourly basis for **NO_2** reveal:

- r parameter value of 0,61
- the value of MB is -0,28 µg/m^3
- and NMB value of 0,78%
- an FAC2 factor of 0,77

Regarding **$PM_{2,5}$**, the model yields these statistics for daily basis data.

- r parameter value of 0,55
- the value of MB is -1,02 µg/m^3
- and NMB value of 6,9%
- an FAC2 factor of 0,93

Referring to **BC**, the model yields these statistics for daily basis data.

- r parameter value of 0,81
- the value of MB is -0.26 µg/m^3
- and NMB value of 17,8%
- an FAC2 factor of 0,96

Spatially, the model offers good performance, especially in the Barcelona conurbation, where most of the AMB population is concentrated.

Since, geolocation data by participants at different time periods are needed to feed the air quality model, two different approaches have been considered. On one hand, GPS data have been used for some time periods in order to know the participant's location at each time of the day (ExpoApp data and commuting). On the other hand, in some cases, commuting routes and home locations have been located with GIS tools according to the explanation for each participant (surveys).

To generalize input GPS data (GPS traces), a Simplified Point Layer (SPL) has been created to represent the road network, as well as other walkable areas. In this sense, the used layers have been trails and footways of the Collserola Natural Park, AMB transit network (2018), Urban parks and green squares (AMB) and *Cadastral road network.*

This SPL has been key to reduce the amount of points to model. By means of this methodology, once the initial 38.079 records of the original input data were processed, 2.298 records have been obtained (which means 94% reduction of the records).

In summary, the total amount of pregnant women modelled according to the input data is 1.084 for commuting, 978 for ExpoApp, 1.083 homes and 36 reference points.

To the whole control group, NO_2 mean concentration lies between 27,5 µg/m^3 and 31 µg/m^3. For $PM_{2,5}$, concentration is between 12,3 µg/m^3 and 12,6 µg/m^3. In relation to BC, the concentration ranges between 0,88 µg/m^3 and 1,0 µg/m^3. As for PM_{10}, concentration is between 19,9 µg/m^3 and 20,9 µg/m^3, depending of the route/place data type.

5. REFERENCES

1. Amato F., Karanasiou A., Morenob, T., Alastuey A., Orza J.A., Lumbreras J., Borge R., Boldo, Linares C., Querol, X. Emission factors from road dust resuspension in a Mediterranean freeway Atmospheric Environment Volume 61, December 2012, Pages 580-587, https://doi.org/10.1016/j.atmosenv.2012.07.065
2. Fulvio Amato; Peter Zandveld; Menno Keuken; Sander Jonkers; Xavier Querol; Cristina Reche; Hugo A.C. Denier van der Gon; Martijn Schaap. 2016. Improving the modeling of road dust levels for Barcelona at urban scale and street level. Atmospheric Environment, ISSN: 1352-2310, Vol: 125, Page: 231-242 (https://doi.org/10.1016/j.atmosenv.2015.10.078)
3. Aeropuertos Españoles y Aeronavegación Aérea (AENA, statistics data, https://wwwssl.aena.es/csee/Satellite?pagename=Estadisticas/Home)
4. AQD, 2008. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe (No. 152). Official Journal.
5. Barcelona Regional and Ajuntament de Barcelona, Model de contaminació local de Barcelona – 2017,
6. Berkowicz, R. OSPM - A Parameterised Street Pollution Model. Environ Monit Assess 65, 323–331 (2000). https://doi.org/10.1023/A:1006448321977
7. Corporación de Reservas Estratégicas de Productos Petrolíferos CORES. Boletín Estadísticos de Hidrocarburos mensuales (https://www.cores.es/es/estadisticas)
8. Departament d'Acció Climàtica, Alimentació i Agenda Rural http://mediambient.gencat.cat/ca/05_ambits_dactuacio/atmosfera/
9. Departament de Territori i Sostenibilitat (DTS) Mapa de Cobertes del Sòl de Catalunya 2013, (http://territori.gencat.cat/es/01_departament/12_cartografia_i_toponimia/bases_cartografiques/medi_ambient_i_sostenibilitat/bases_miramon/territori/mapa-dusos-i-cobertes-del-sol/, gener del 2017).
10. European Environment Agency. EMEP/EEA air pollutant emission guidebook 2019 (https://www.eea.europa.eu/publications/emep-eea-guidebook-2019).
11. Institut Català d'Energia (ICAEN). Sectorial energy consumption statistics (personal request)
12. Institut Cartogràfic i Geològic de Catalunya – ICGC. municipal cadastre
13. Meteosim https://www.meteosim.com/es/
14. M.J. Barnes, T.K. Brade, A.R. MacKenzie, J.D. Whyatt, D.J. Carruthers, J. Stocker, X. Cai, C.N. Hewitt, Spatially-varying surface roughness and ground-level air quality in an operational dispersion model, Environmental Pollution, Volume 185, 2014, Pages 44-51, ISSN 0269-7491, https://doi.org/10.1016/j.envpol.2013.09.039.
15. Port de Barcelona Open data (https://opendata.portdebarcelona.cat/en/)
16. Registro Estatal de Emisiones y Fuentes Contaminantes. EPER-PRTR (https://prtr-es.es/Informes/InventarioInstalacionesIPPC.aspx)
17. Salizzoni, P., Marro, M., Soulhac, L. et al. Turbulent Transfer Between Street Canyons and the Overlying Atmospheric Boundary Layer. Boundary-Layer Meteorol 141, 393–414 (2011). https://doi.org/10.1007/s10546-011-9641-1
18. Servei Meteorològic de Catalunya (SMC) Meteorological data from XEMA (https://www.meteo.cat/observacions/xema)
19. Thunis, Pederzoli and Pernigottiet, (2012) Performance criteria to evaluate air quality modelling applications Atmospheric Environment 59:476–482 (10.1016/j.atmosenv.2012.05.043)
20. Tsyro, S.G. 2001: Description of the Lagrangian Acid Deposition Model. (http://www.emep.int/mscw/eudm_acid_model.pdf, gener del 2018).

21. Venkatram, A., Karamchandani, P., Pai, P. And Goldstein, R., 1994. The Development and Application of a Simplified Ozone Modelling System (SOMS). Atmos. Enriron, 28 pp. 335-3678.