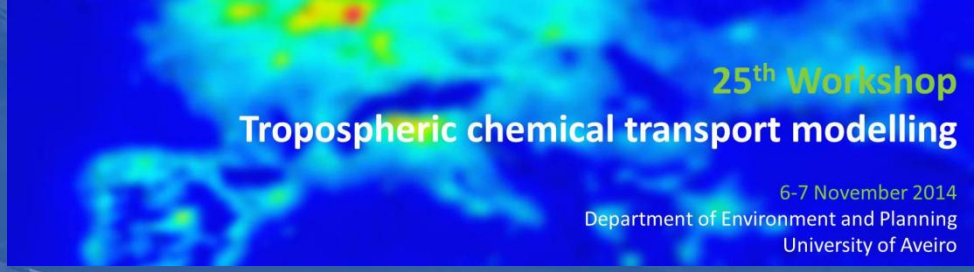
www.bsc.es



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Spatial representativeness of air quality monitoring stations using a 1-km resolution modeling system: application to Andalusia

M.T. Pay[1,*], M. Guevara[1], D. Trillo-Montero[2], A. Lozano-García[2], J.M. Baldasano[1,3]

[1]Earth Sciences Department, Barcelona Supercomputing Center, Barcelona, Spain
[2]Environmental and Water Agency of Andalusia, Seville, Spain
[3]Environmental Modeling Laboratory, Technical University of Catalonia, Barcelona, Spain
*now at: Laboratoire de Météorologie Dynamique, École Polytechnique, Palaiseau Cedex, France

25th GLOREAM-EURASAP Workshop on Tropospheric Chemical Transport Modelling
Aveiro (Portugal), November 6-7, 2014

Spatial representativeness of air quality monitoring stations

- Spatial representativeness is a critical information when
  - choosing the location of monitoring sites
  - assessing the effects on population of short-long term exposure to air pollution (Righini et al., 2013)
- Spatial representativeness is complex to quantify because it depends on:
  - ambient and local factors affecting the station (topography, emission sources, meteorological patterns, urban/rural environment)
  - the measured pollutant
  - the temporal basis used to estimate the SR.
- Spatial representativeness is harmonized defined:

AQUILA and FAIRMODE are promoting the discussion concerning spatial representativeness Kracht (2014)

> ***"Representativeness is the extent to which a set of measurements taken in a space-time domain reflects the actual conditions in the same or different spacetime domain taken on a scale appropriate for a specific application."***
>
> **(Nappo et al. 1982)**

> ***[The area of representativeness is ...] "... the area in which the concentration does not differ from the concentration measured at the station by more than a specified amount."***
>
> **(Larssen et al. 1999)**

> ***"A monitoring station is representative of a location if the characteristic of the differences between concentrations over a specified time period at the station and at the location is less than a certain threshold value."***
>
> **(Spangl et al. 2007)**

Spatial Representativeness in the AQ EU Legislation

- Monitoring criteria in the Air Quality Directive (2008/50/EC)

Type of station	Representativeness
Urban	A few km^2
Suburban	Some tens of km^2
Rural	Some hundreds of km^2
Rural background	1000 to 10 000 km^2

- There is no definitions about the "spatial representativeness" of monitoring stations in the Air Quality Directive.
- There is a need to develop tools for its quantitative assessment.

Annex VIII – Directive 2008/50/EC

Type of station	Objectives of measurement	Representativeness (1)	Macroscale siting criteria
Urban	Protection of human health: to assess the exposure of the urban population to ozone, i.e. where population density and ozone concentration are relatively high and representative of the exposure of the general population	A few km^2	Away from the influence of local emissions such as traffic, petrol stations, etc.; vented locations where well mixed levels can be measured; locations such as residential and commercial areas of cities, parks (away from the trees), big streets or squares with very little or no traffic, open areas characteristic of educational, sports or recreation facilities
Suburban	Protection of human health and vegetation: to assess the exposure of the population and vegetation located in the outskirts of the agglomeration, where the highest ozone levels, to which the population and vegetation are likely to be directly or indirectly exposed occur	Some tens of km^2	At a certain distance from the area of maximum emissions, downwind following the main wind direction/directions during conditions favourable to ozone formation; where population, sensitive crops or natural ecosystems located in the outer fringe of an agglomeration are exposed to high ozone levels; where appropriate, some suburban stations also upwind of the area of maximum emissions, in order to determine the regional background levels of ozone
Rural	Protection of human health and vegetation: to assess the exposure of population, crops and natural ecosystems to sub-regional scale ozone concentrations	Sub-regional levels (some hundreds of km^2)	Stations can be located in small settlements and/or areas with natural ecosystems, forests or crops; representative for ozone away from the influence of immediate local emissions such as industrial installations and roads; at open area sites, but not on summits of higher mountains
Rural background	Protection of vegetation and human health: to assess the exposure of crops and natural ecosystems to regional-scale ozone concentrations as well as exposure of the population	Regional/ national/ continental levels (1 000 to 10 000 km^2)	Station located in areas with lower population density, e.g. with natural ecosystems, forests, at a distance of at least 20 km from urban and industrial areas and away from local emissions; avoid locations which are subject to locally enhanced formation of ground-near inversion conditions, also summits of higher mountains; coastal sites with pronounced diurnal wind cycles of local character are not recommended.

(1) Sampling points should, where possible, be representative of similar locations not in their immediate vicinity.

Spatial Representativeness
Review of methods

Method	Reference	Advantage	Disadvantage
Passive sampling surveys	• Thornburg et al., 2009	• Accuracy of continuous measurement	• Limited spatial resolution • Limited temporal resolution • Limited availability of data and pollutant
Objective factors (land cover)	• Janssen et al., 2012; Piersanti et al., 2013	• High spatial coverage (2D) • Easily available • Urban areas	• Limited accuracy based on statistical analyses an expert judgement • Only long term (annual averaged) representativeness
Emission variability	• Cremola et al., 2013	• High spatial coverage (2D) • Primary pollutants	• Limited accuracy based on statistical analyses an expert judgement
Modelled concentration similarity	• Vitali et al., 2013	• High spatial (3D) and temporal coverage (year) and resolution (1km, 1hr) • Flexibility • Easy to cover several pollutants • Secondary pollutants • Rural areas	• Resource intensive • Depending on input parameters (emissions, meteorology) • Shortcomings of model accuracy
Semi-variogram of modelled concentration	• Solazzo et al., 2014	• The same as modelled concentration similarity • Quantify the spatial representativeness uncertainty.	• The same as modelled concentration similarity • Mathematically complex (geoestatiscical model)

Objective and methods

- CTM are potential tools to quantitatively assess about spatial representativeness: they have a high spatial and temporal coverage, deal with primary and secondary pollutants, consider the distribution of pollution sources, and take into account meteorological and topographical features affecting the surrounding of the monitoring sites

  → **EVALUATION**

- The objective is **implement the "*concentration similarity method*" to analyze and quantify the spatial representativeness (SR) of an air quality monitoring network using a 1-km resolution modelling system**:
  - To determinate how to estimate the maximum discrepancy (DM).
  - To study the influence of SR as a function of the pollutant, temporal basis and station type.
  - To analyze the effectiveness of the air quality network considering it capability to cover most emission sources and populated area.

- **Method**:
  - To quantify the SR of the air quality monitoring stations of Andalusia (southern Spain) for O_3, NO_2, SO_2, PM10 and PM2.5 and the year 2013.
  - Air quality model: CALIOPE Air Quality Forecasting System running at 1-km horizontal.

CALIOPE: Air Quality Forecasting System

CALIOPE modules

Meteorology
- WRF-ARWv3.5
- 38 sigma levels (top 50 hPa)
- IBC: GFS (NCEP)
- 33 layers/50 hPa

Emissions
- HERMESv2
- EU: HERMES-DIS (EMEP data)
- Spain: HERMES-BOUP

Chemistry
- CMAQv5.0.1
- CB05/AERO5
- BC: NCAR MOZART4
- 15 layers/ 50 hPa

Mineral dust
- BSC-DREAM8bv2
- Mineral PM10 and PM2.5

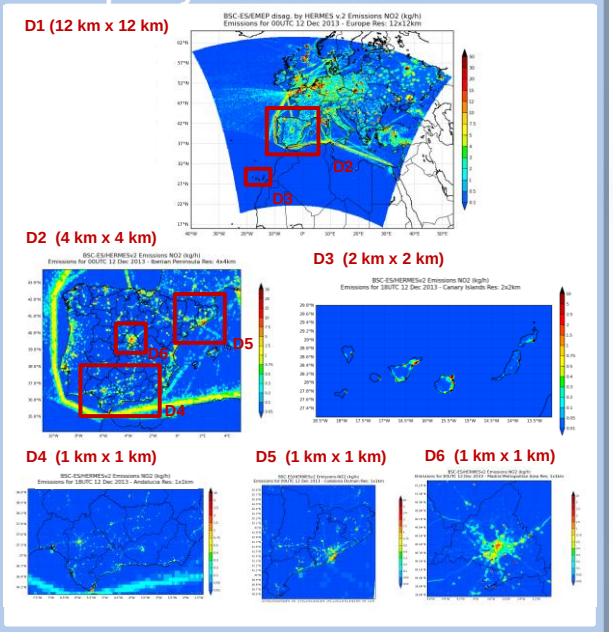
Post-process
- Kalman Filter (puntual y 2D)

Air quality forecast

O_3, NO_2, SO_2, CO, PM10, PM2.5, Benzene

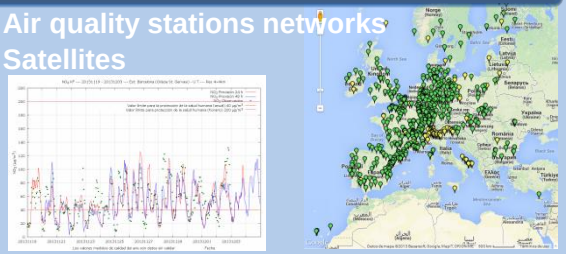
48h forecast

- Concentration maps, emis, meteo
- Air quality indexes



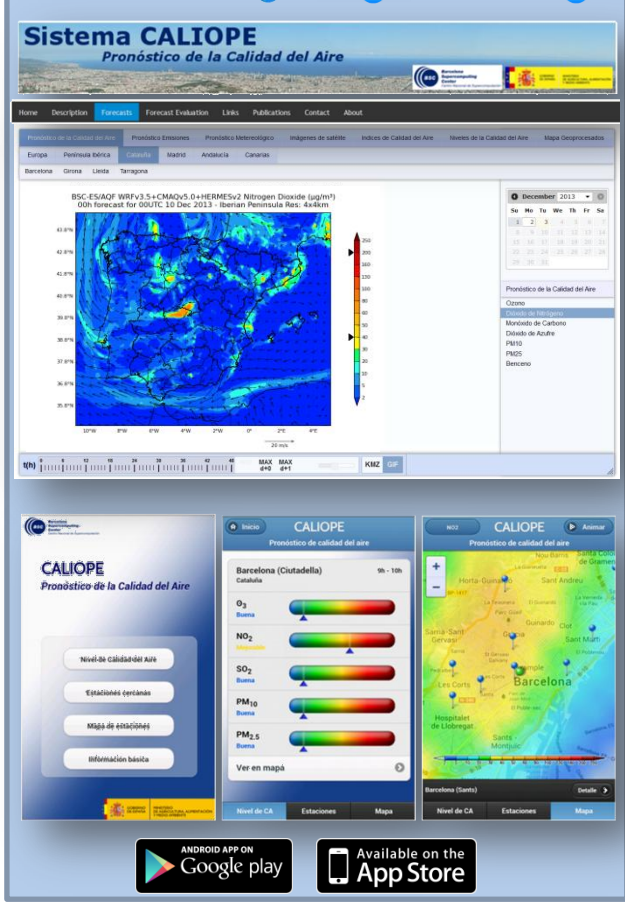
Near real-time evaluation

- Air quality stations networks
- Satellites

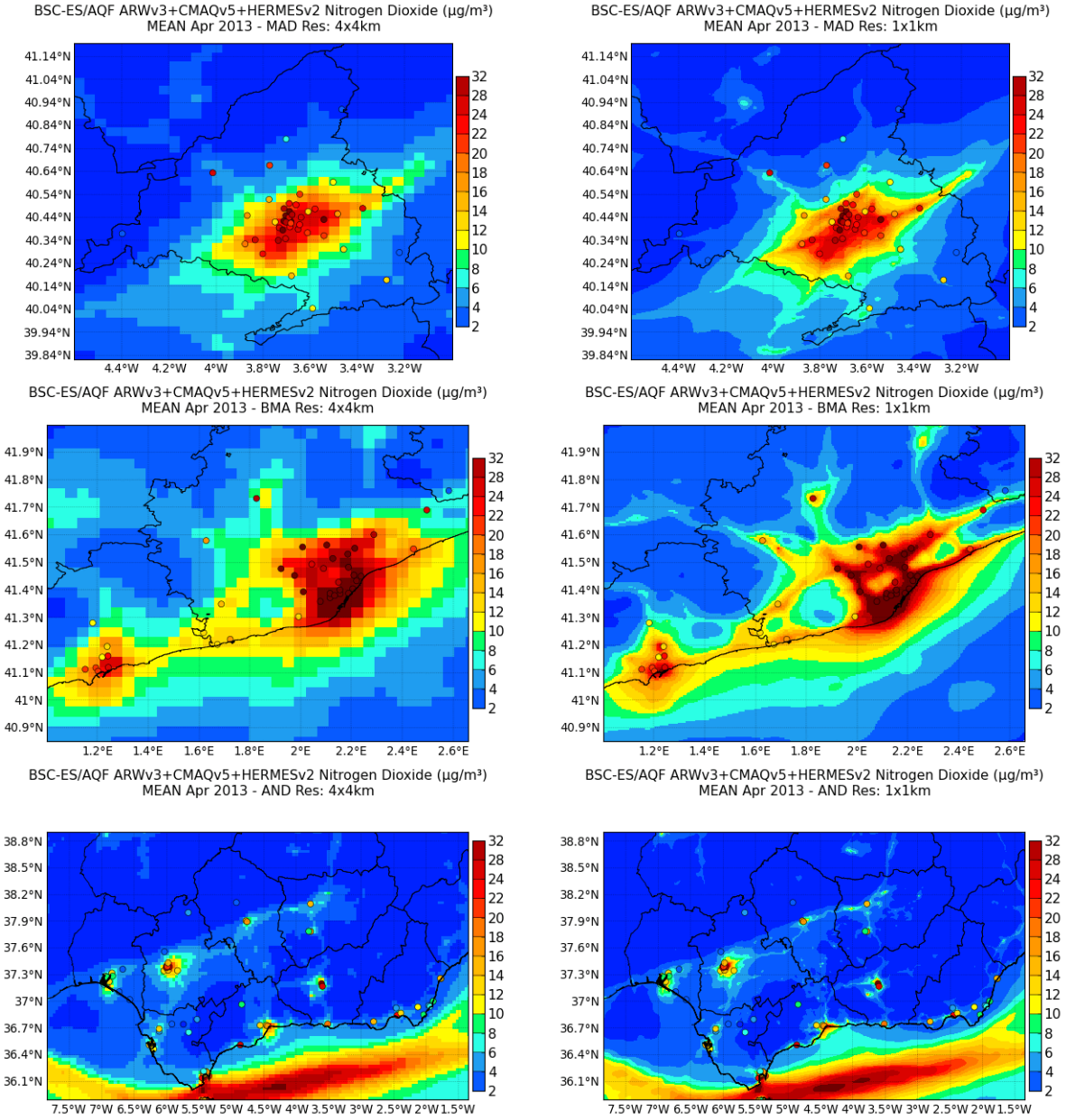


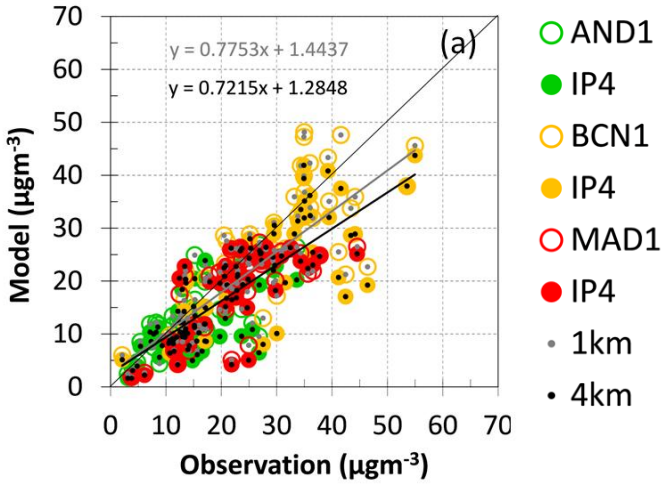
Diffusion

- Web (www.bsc.es/caliope)
- Smartphone



Forecasted NO_2: mean concentration April 2013 → 1 km





(e)	r_{NO2}	
	4km	1km
All	0.79	0.81
AND	0.62	0.71
BCN	0.72	0.73
MAD	0.70	0.74

Pay et al., GMD 2014

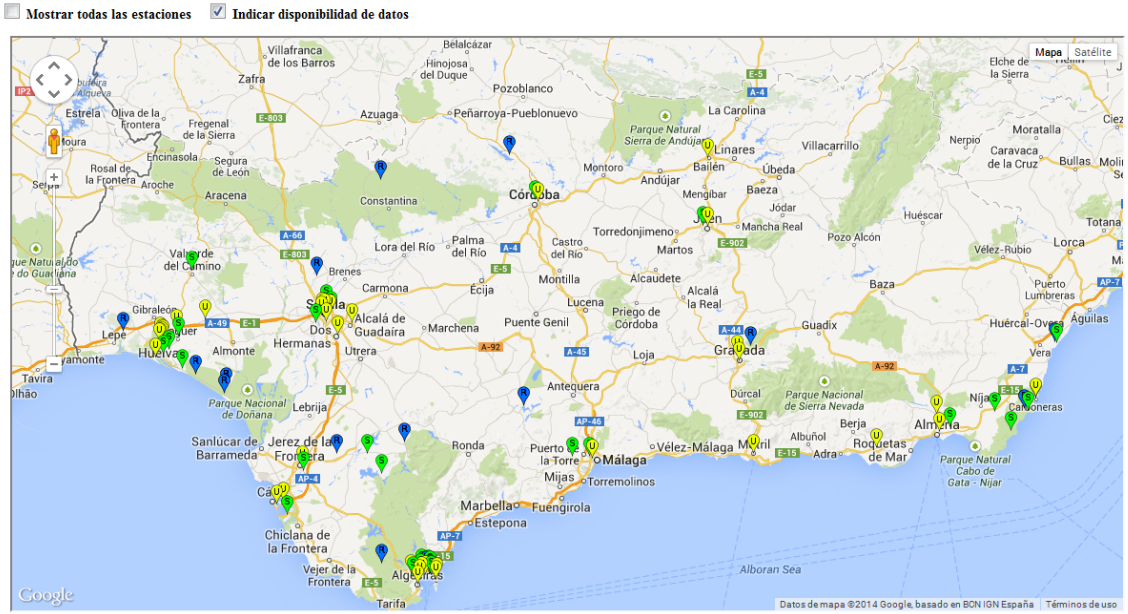
Air quality monitoring network in ANDALUCÍA

Dominio	Total	Rural	Suburbanas	Urbanas
UE	573	291 (RF)	275 (SF)	
IP	445	117	127	201
CAN	33	5	10	18
AND	90	18	32	40
CAT	84	33	26	25

Número de estaciones por dominio y contaminante

Dominio	O_3	NO_2	SO_2	PM_{10}	$PM_{2.5}$
EU	465	308	119	210	47
IP	290	345	250	223	43
CAN	25	32	33	32	28
AND	49	76	71	56	13
CAT	49	61	45	21	2

	O3			NO2			SO2			PM10			PM2.5		
	U	S	R	U	S	R	U	S	R	U	S	R	U	S	R
Andalucía	19	19	11	35	26	15	35	23	13	29	18	8	3	7	3



Concentration similarity method

- The monitoring station is representative of a wider area if all concentrations in this area differ by less than a threshold from the concentrations in the site of interest
- Concentrations are modelled running a high resolution air quality model
- The condition of “concentration similarity” is assessed with a threshold (maximum discrepancy, DM) (Jansen et al., 2008; Vitali et al., 2013).
- A point measurement is representative of the average in a larger area if the probability that the squared difference between point and area measurement is smaller than a certain threshold more than 90% of the time (Nappo et al., 1982)

$$f_{sito}(x,y) = \frac{\sum_{i=1}^{N_t} flag}{N_t}, \quad conflag = \begin{cases} 1, & \frac{|C(X_{sito},Y_{sito},Z_0,t_i) - C(x,y,Z_0,t_i)|}{C(X_{sito},Y_{sito},Z_0,t_i)} < 0.2 \\ 0, & \frac{|C(X_{sito},Y_{sito},Z_0,t_i) - C(x,y,Z_0,t_i)|}{C(X_{sito},Y_{sito},Z_0,t_i)} > 0.2 \end{cases}$$

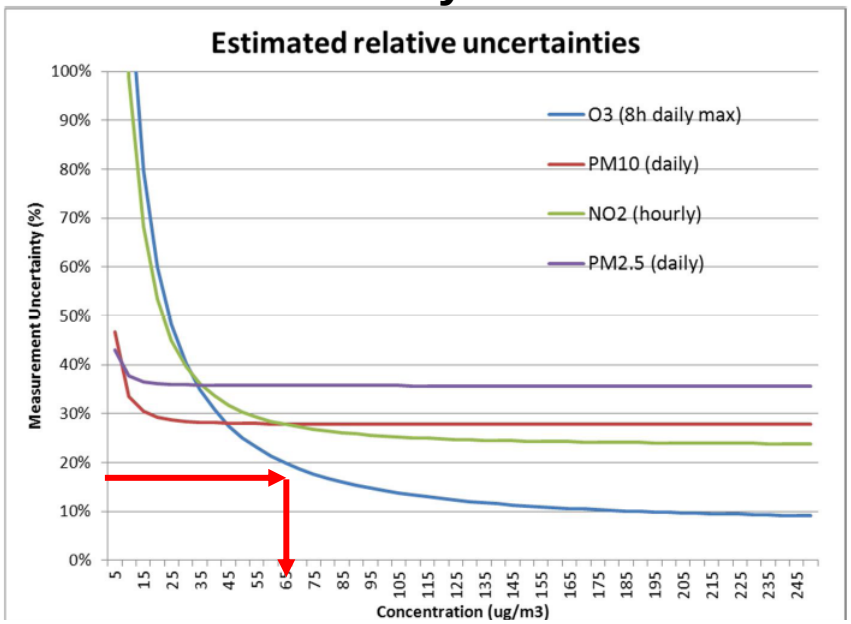
Review of valued for maximum discrepancy (DM)

DM	Pollutant	Reference
15% 25%	O_3, NO_2, SO_2 PM10, PM2.5	2008/50/EC
20%	All pollutants	Jansen et al. (2008) Vitali et al. (2013)
20% 100%	O_3, NO_2, SO_2, PM10 SO_2 (<4 μgm^{-3} annual mean, < 25 μgm^{-3} daily average, < 70 μgm^{-3} hourly average) NO_2 (< 13 μgm^{-3} annual mean, < 50 μgm^{-3} hourly average)	Martin et al. (2013)

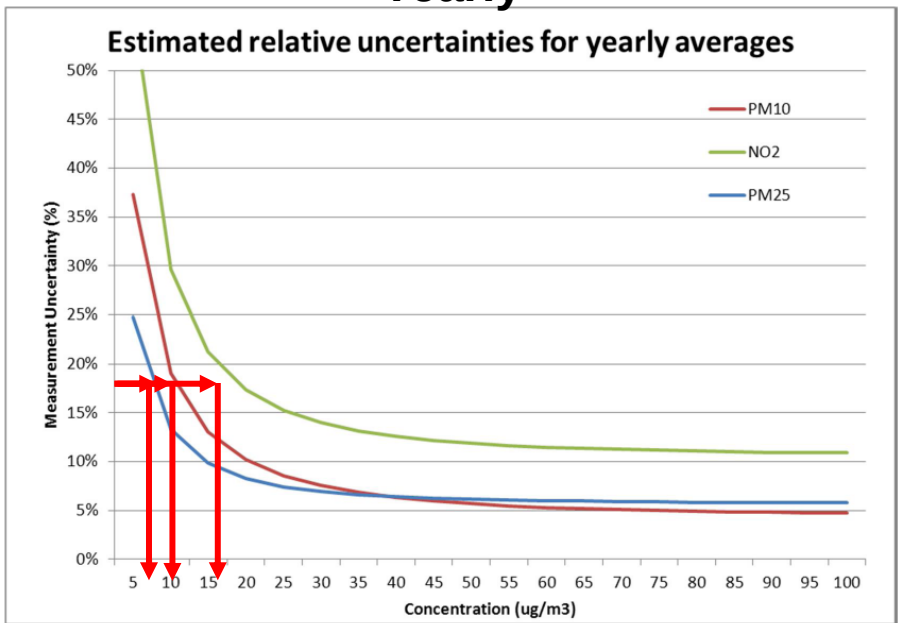
DM base on the observation uncertainty
DM = 20 % is the most common value

Formulation of the observation uncertainty
Thunis et al. (2013), Pernigotti et al. (2013), FAIRMODE

Daily



Yearly



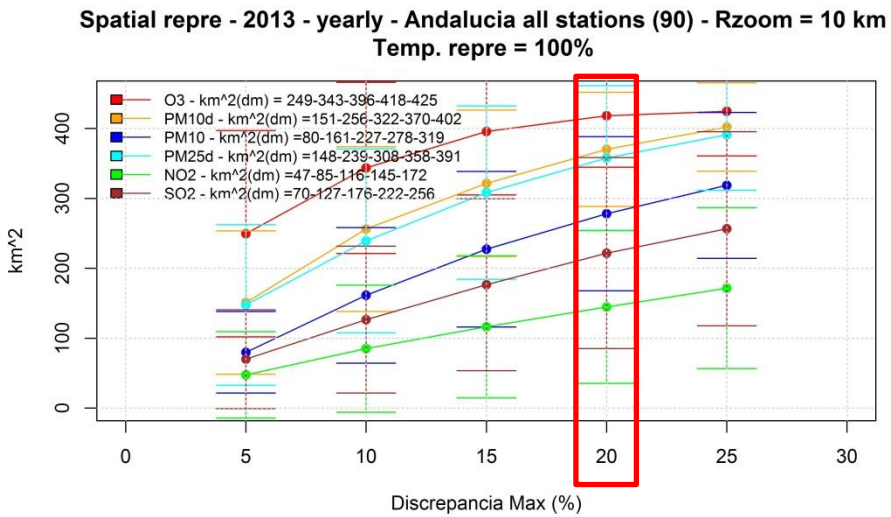
According to the FAIRMODE estimated uncertainty of the observation as a function of the concentration :

- For **hourly/daily basis**: DM = 20% is strict enough for NO_2, PM10 and PM2.5 because the uncertainty for these pollutants on an daily basis is always > 25% (especially at low concentration). In the case of O_3, the DM = 20% is the same order of magnitude as the uncertainty corresponding to the mean ozone concentration (65.5 $\mu g/m^{-3}$) over the Andalucía stations.
- For **annual basis**: DM = 20% is strict enough for low concentration (< 15 $\mu g/m^{-3}$ for NO_2, < 10 $\mu g/m^{-3}$ for PM10 and < 5 $\mu g/m^{-3}$ for PM2.5) because of the estimated observation uncertainty increase fast with the decrease of concentration. However, DM = 20% is relax enough for higher concentrations (> 15 $\mu g/m^{-3}$ for NO_2, > 10 $\mu g/m^{-3}$ for PM10 and > 5 $\mu g/m^{-3}$ for PM2.5) because the estimated observation uncertainty is usually < 15 %.

Spatial representativeness as a function of pollutant: *Andalucía case on in a yearly basis*

Mean spatial representativeness of each stations over a Radio = 10 km from the station

Yearly



Mean spatial representativeness for each Andalucía AQN as a function of pollutant. For any dm, the spatial representativeness:

$$O_3 > PM10 > SO_2 > NO_2$$

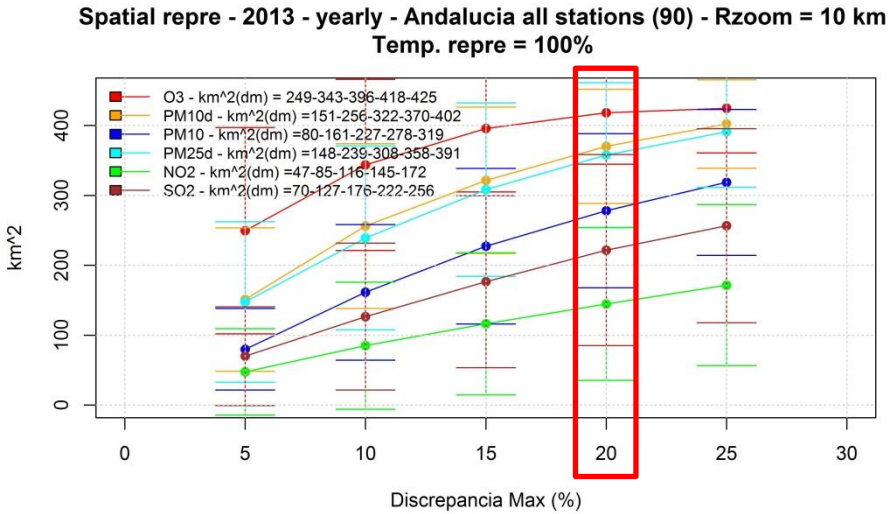
At DM = 20%:

- O_3 spatial representativeness : 418 km^2
- PM10 spatial representativeness: 278 km^2
- SO_2 spatial representativeness: 222 km^2
- NO_2 spatial representativeness: 145 km^2

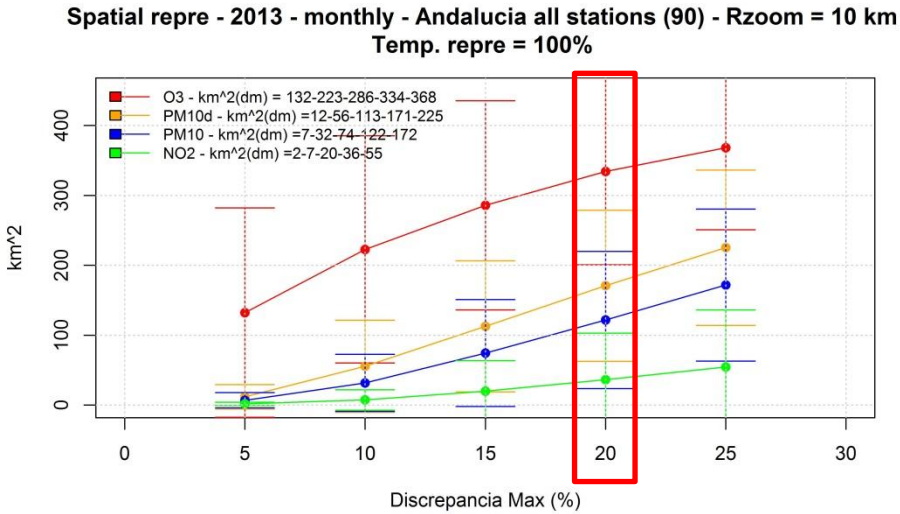
Spatial representativeness as a function of temporal resolution: *Andalucía case*

Mean spatial representativeness of each stations over a Radio = 10 km from the station

Yearly



Monthly

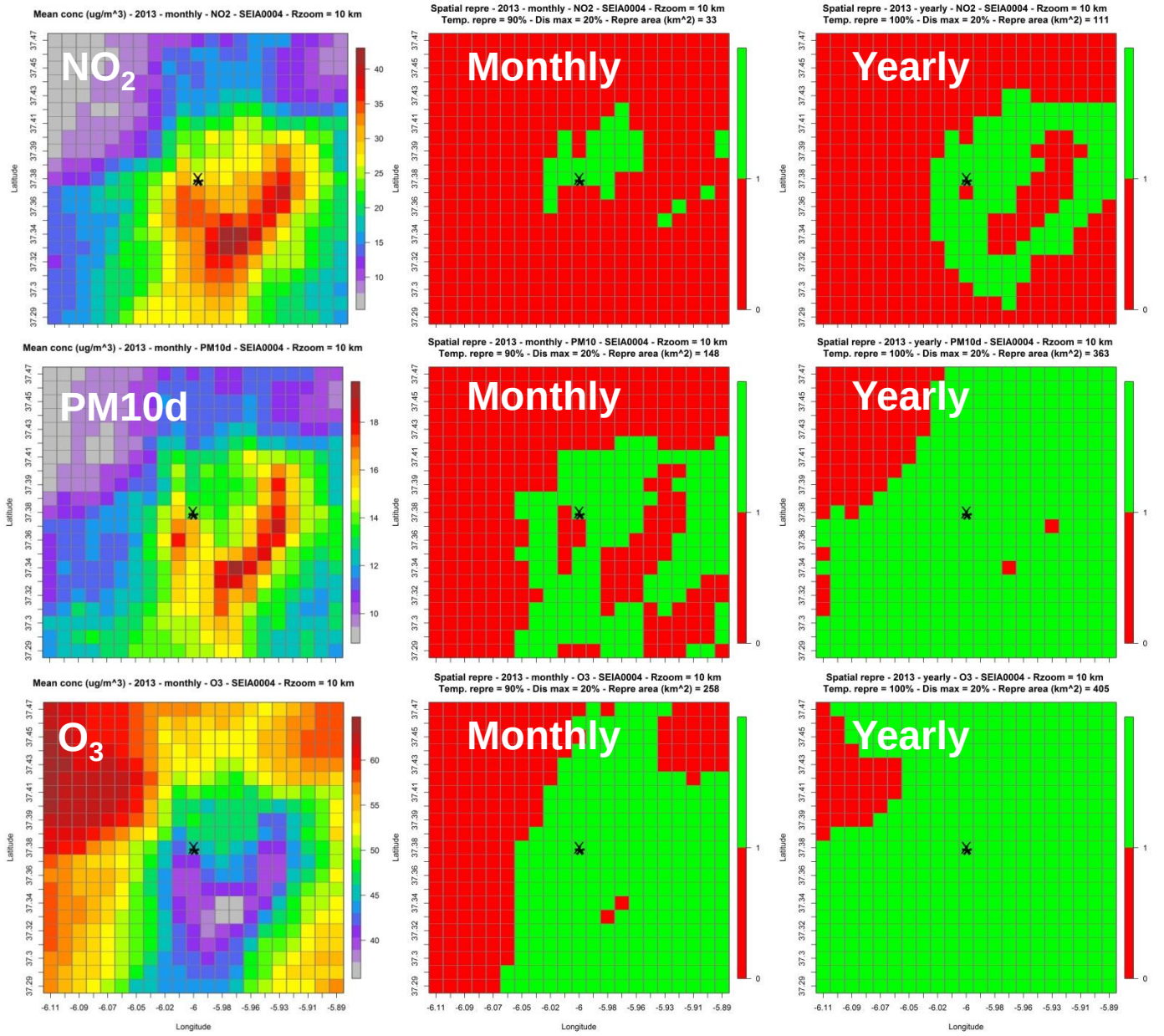


Mean spatial representativeness for each Andalucía AQN as a function of the temporal basis (annual vs monthly). At DM = 20%:

- O_3 spatial representativeness : 334/418 km^2 (month/annual) → 25% increase
- NO_2 spatial representativeness: 36/145 km^2 (month/annual) → 303% increase
- PM10d spatial representativeness: 171/370 km^2 (month/annual) → 116% increase

annual representativeness > monthly representativeness

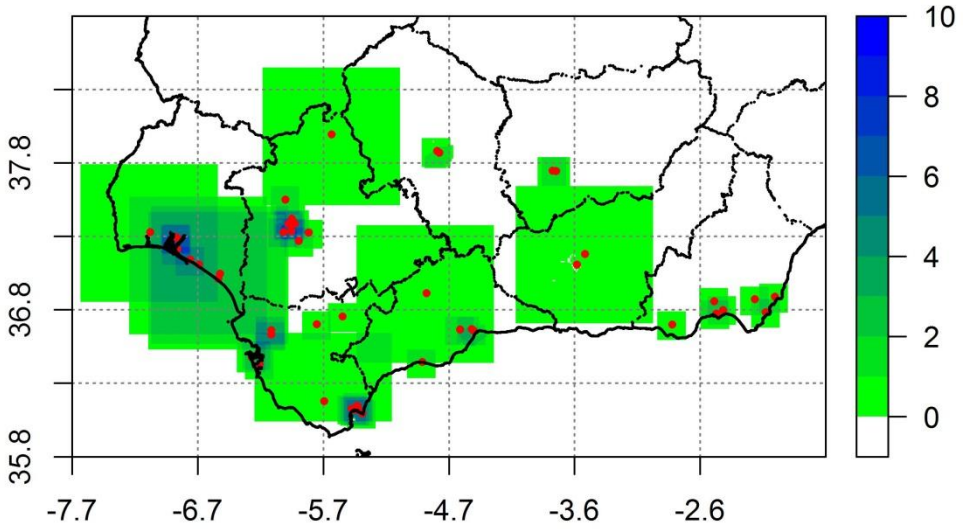
Spatial representativeness as a function of temporal resolution | PRINCIPE station (dis.max = 20%)



Frequency of spatial representativeness (FreqSR) Explanation

Frequency of spatial representativeness (FreqSR) = number of times a grid cell is representative for one or more stations

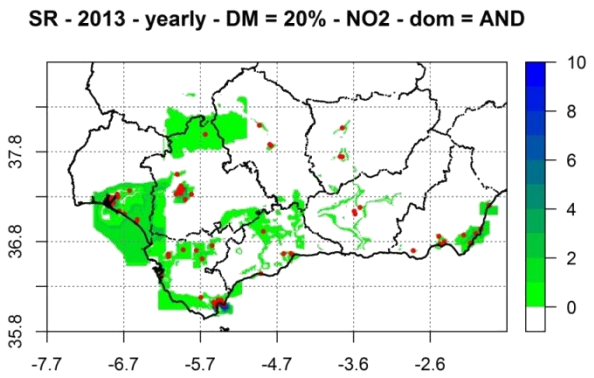
FreqSR – 2013- yearly – DM = 20% – O_3

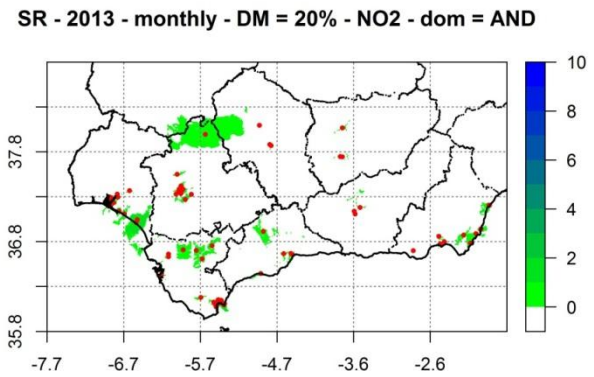


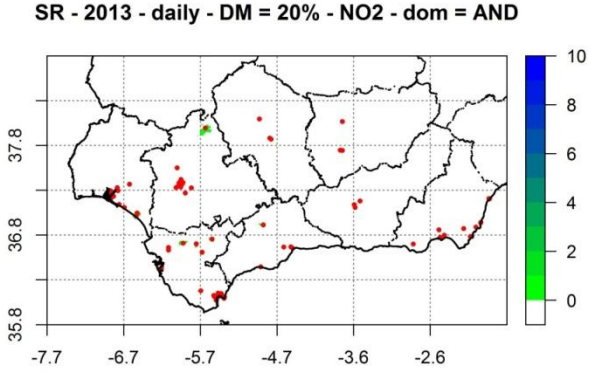
According to the 2008/50/EC, the maxima spatial representativeness (SRmax) is constrained to:

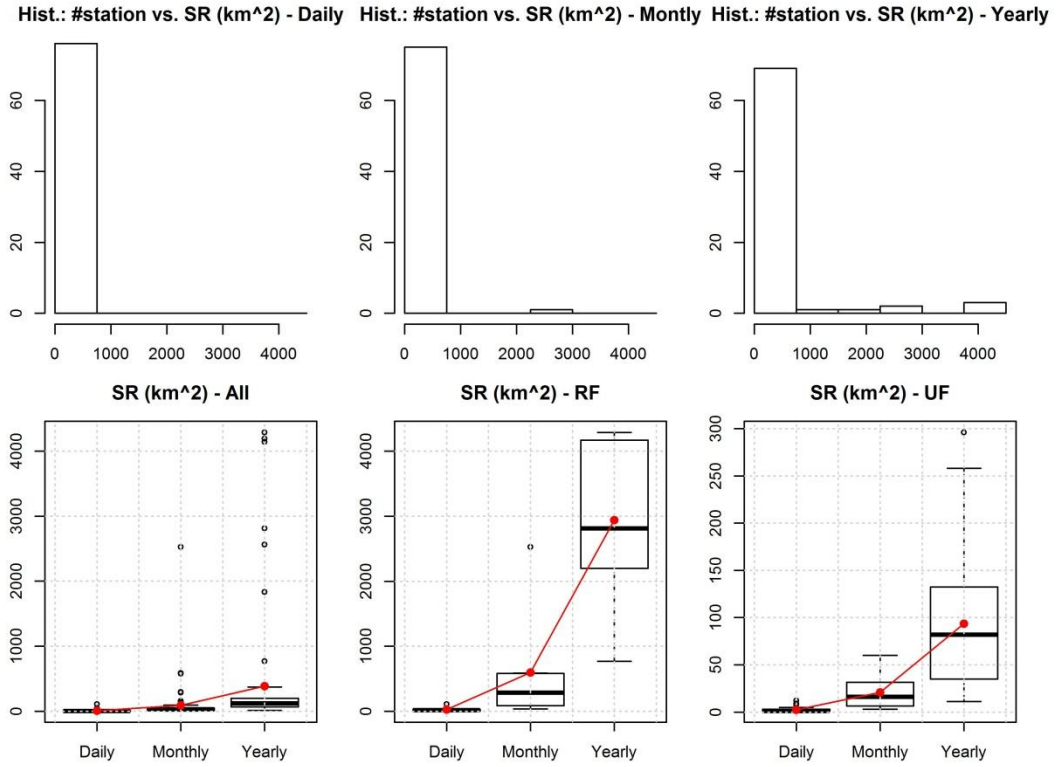
- SR max (Rural Background) = 10.000 km^2 (r = 50 km)
- SR max (Urban, Suburban, Rural Industrial, Rural Traffic) = 400 km^2 (r = 10 km)

NO_2 spatial representativeness



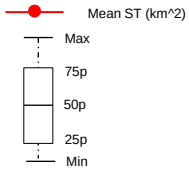




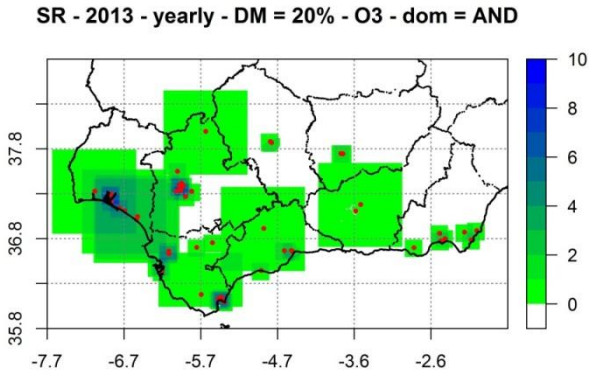


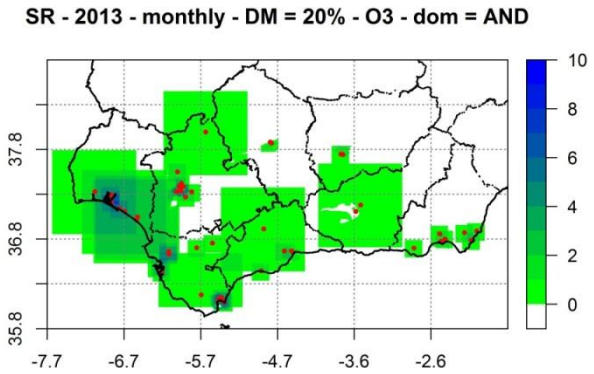
type	n	SRT	SRimd	SRimm	SRimy
All	76	428(d) / 5821(m) / 20365(y)	5.8±13.7	87.2±301.4	383±908.8
RF	7	-	26.3±37.1	596.7±881.9	2942.6±1350
R	15	-	17.3±27.8	327.3±637.4	1483.9±1668.2
U	35	-	2.4±2.5	20.7±16.3	93.4±71.1
S	26	-	3.7±4.5	38.3±36	137.8±71.6

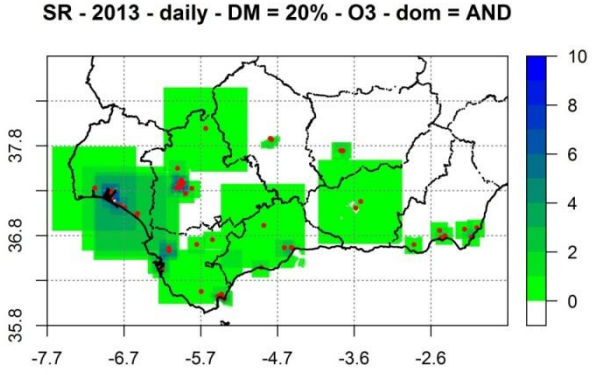
SRT = Total Spatial Representativeness (in km^2) of the monitoring network in daily (d), monthly (m) and yearly (y)
SRimd = Mean Spatial representativeness of a stations (in km^2) in a daily basis.
SRimm = Mean Spatial representativeness of a stations (in km^2) in a monthly basis.
SRimy = Mean Spatial representativeness of a stations (in km^2) in a yearly basis.

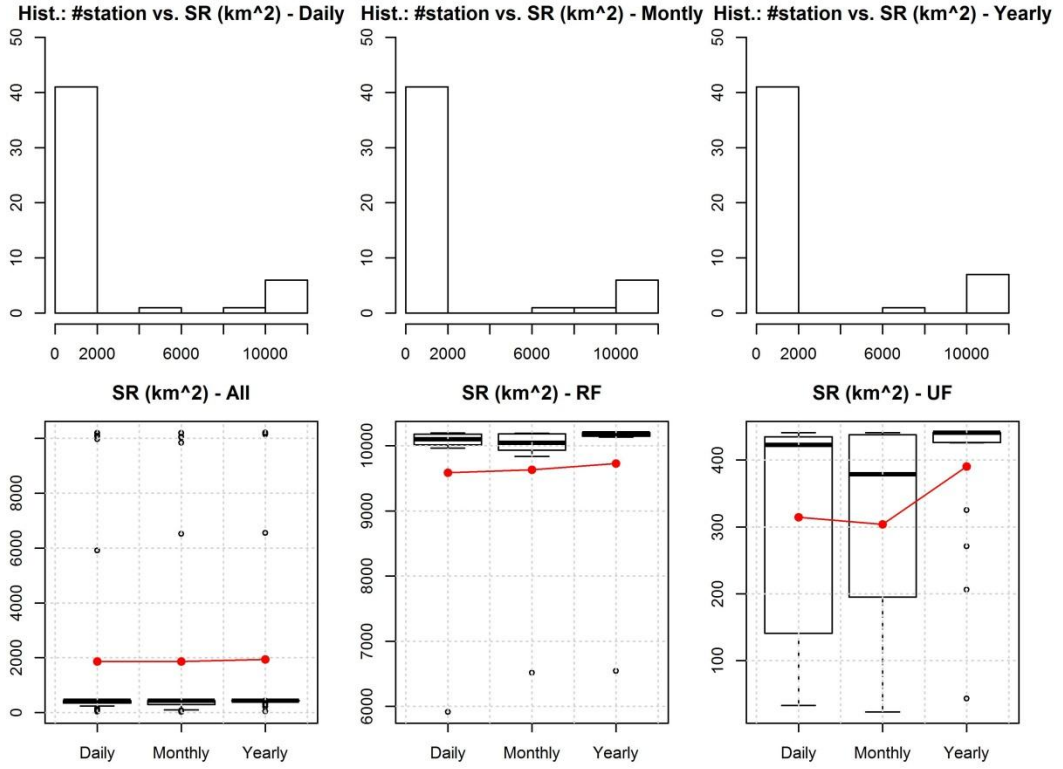


O_3 spatial representativeness



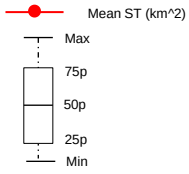




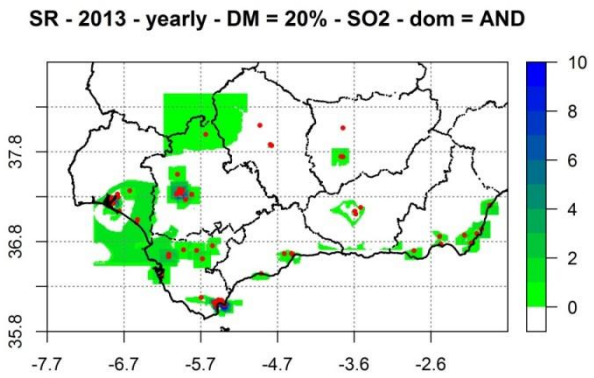


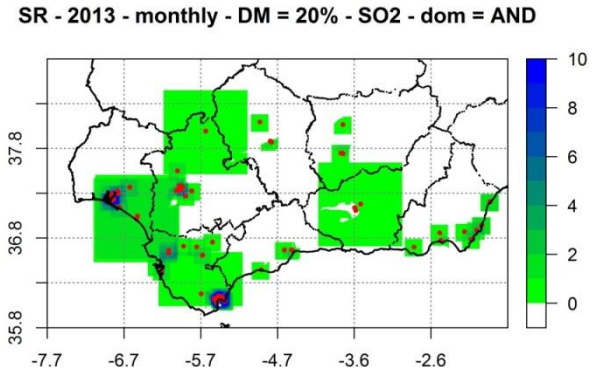
type	n	SRT	SRimd	SRimm	SRimy
All	49	58487(d) / 58755(m) / 59563(y)	1859.3±3497.3	1865.8±3500.4	1935.1±3512.9
RF	8	-	9586.9±1486.3	9629.8±1264.1	9727.4±1286.9
R	11	-	7092.5±4449.5	7122.8±4422	7194.7±4469.3
U	19	-	314.6±157.1	303.6±148.7	390.5±107.5
S	19	-	374.3±106	384.3±82.3	434.7±12.5

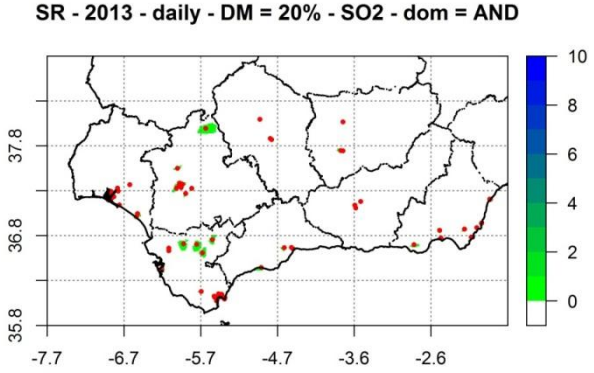
SRT = Total Spatial Representativeness (in km^2) of the monitoring network in daily (d), monthly (m) and yearly (y)
SRimd = Mean Spatial representativeness of a stations (in km^2) in a daily basis.
SRimm = Mean Spatial representativeness of a stations (in km^2) in a monthly basis.
SRimy = Mean Spatial representativeness of a stations (in km^2) in a yearly basis.

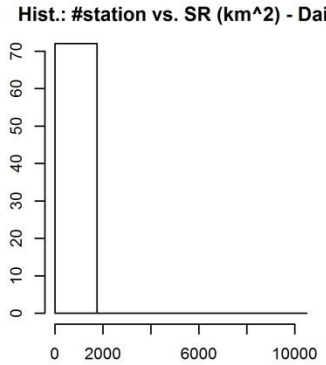


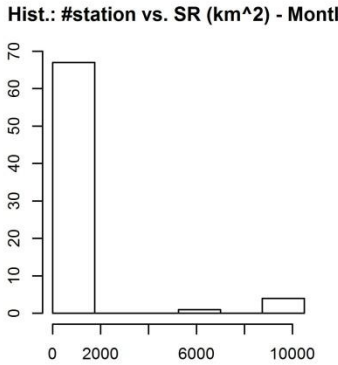
SO_2 spatial representativeness

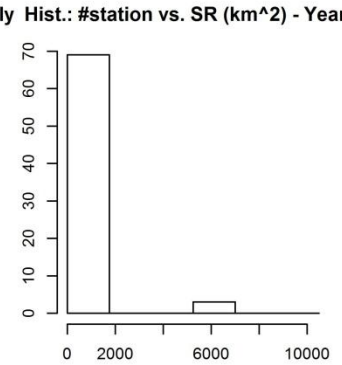


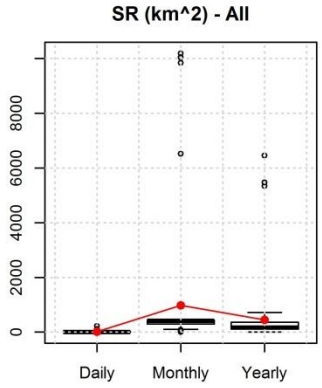


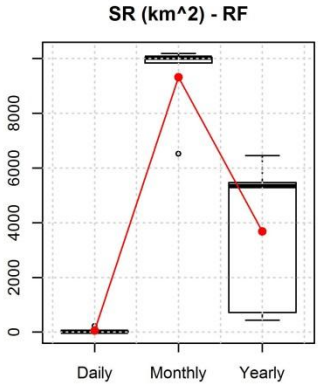


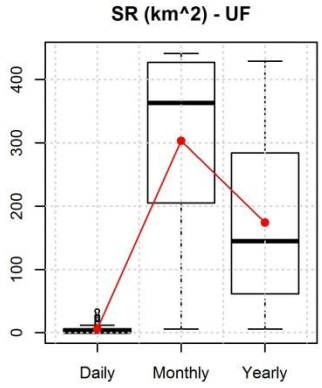






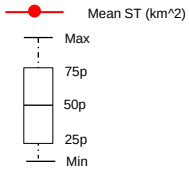




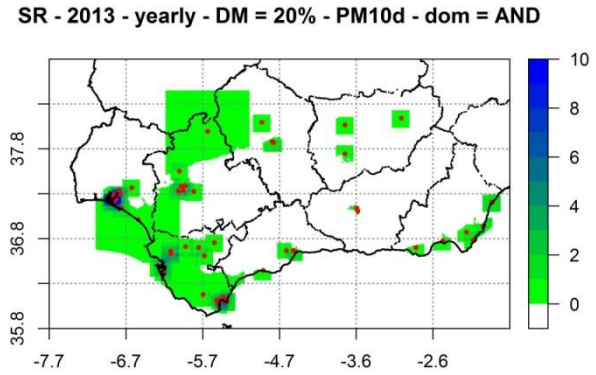


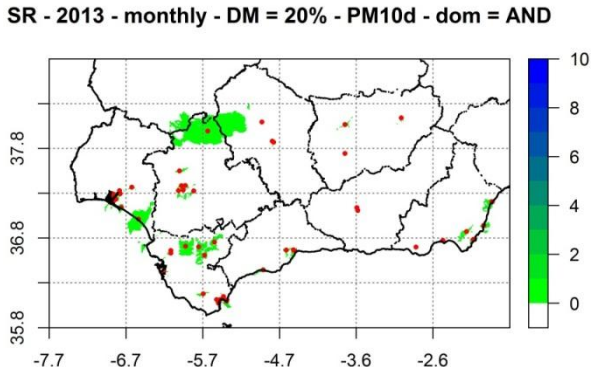
type	n	SRT	SRimd	SRimm	SRimy
All	72	1045(d) / 44291(m) / 20748(y)	15±33.2	974.4±2329.4	442.5±1129.4
RF	5	-	56±94.6	9321.8±1573.9	3684.2±2867
R	13	-	38±64.6	3842.2±4599.7	1564.8±2407.6
U	36	-	6.2±9	303.3±139.9	173.9±133.6
S	23	-	15.9±26.7	403.9±58.9	228.6±127

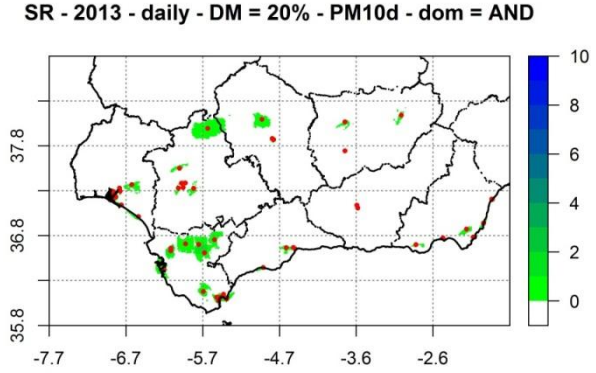
SRT = Total Spatial Representativeness (in km^2) of the monitoring network in daily (d), monthly (m) and yearly (y)
SRimd = Mean Spatial representativeness of a stations (in km^2) in a daily basis.
SRimm = Mean Spatial representativeness of a stations (in km^2) in a monthly basis.
SRimy = Mean Spatial representativeness of a stations (in km^2) in a yearly basis.

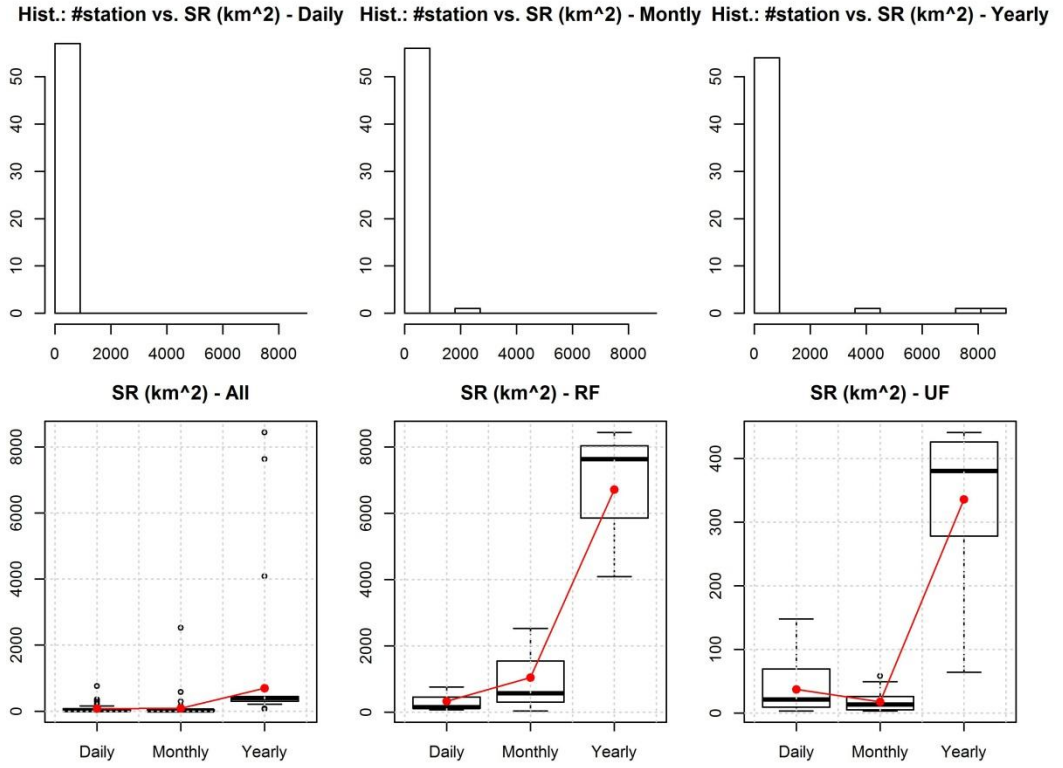


PM10 spatial representativeness



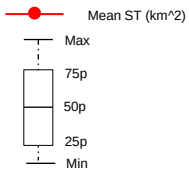




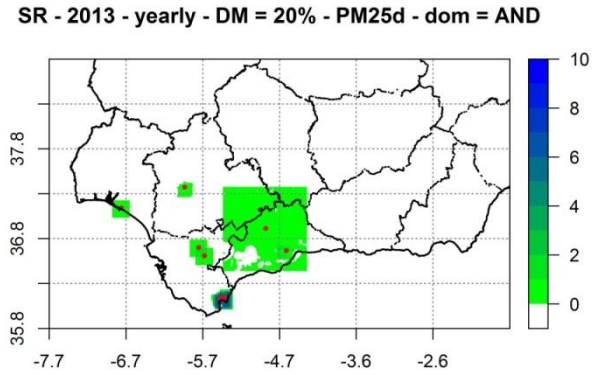


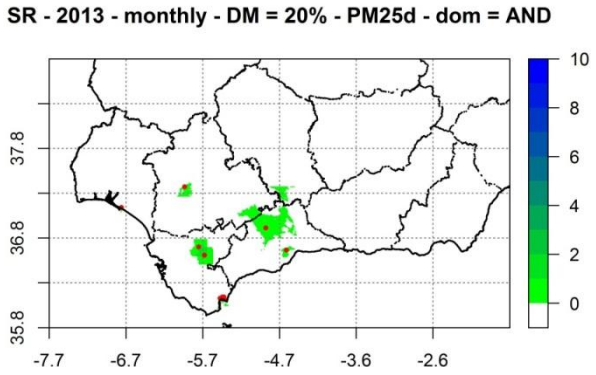
type	n	SRT	SRimd	SRimm	SRimy
All	57	3943(d) / 4906(m) / 27997(y)	77.3±122.6	89.4±339.7	697.4±1500.1
RF	3	-	329.7±372.2	1045.7±1311	6719.7±2314.5
R	8	-	240.5±236	464±854.8	2792.9±3479.1
U	30	-	37.1±37.4	17.6±14.7	335.9±106.4
S	19	-	72.1±91.1	45.1±39.9	385.7±63.8

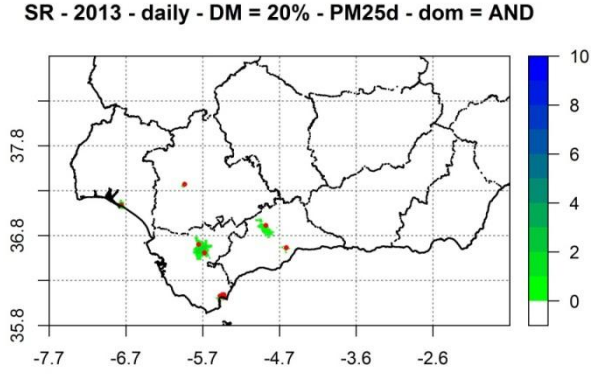
SRT = Total Spatial Representativeness (in km^2) of the monitoring network in daily (d), monthly (m) and yearly (y)
SRimd = Mean Spatial representativeness of a stations (in km^2) in a daily basis.
SRimm = Mean Spatial representativeness of a stations (in km^2) in a monthly basis.
SRimy = Mean Spatial representativeness of a stations (in km^2) in a yearly basis.

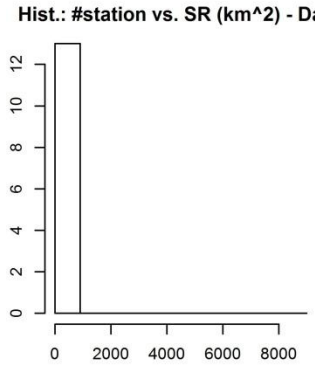


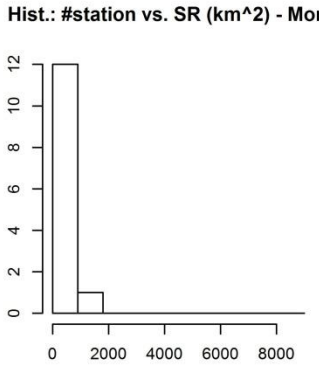
PM2.5 spatial representativeness

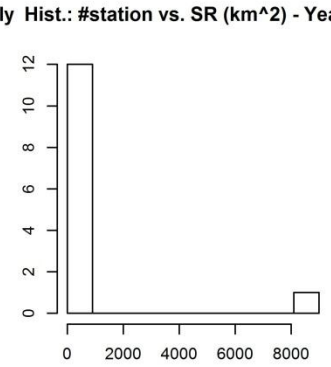


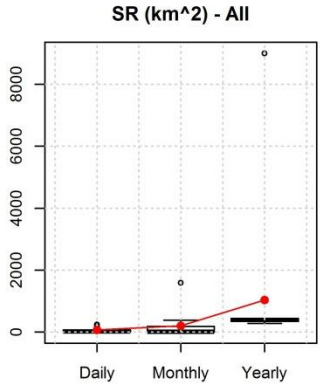


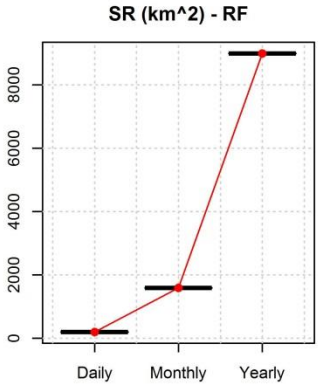


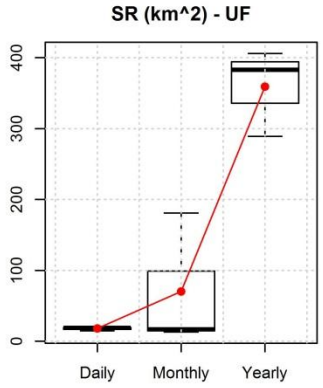






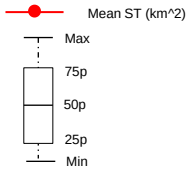


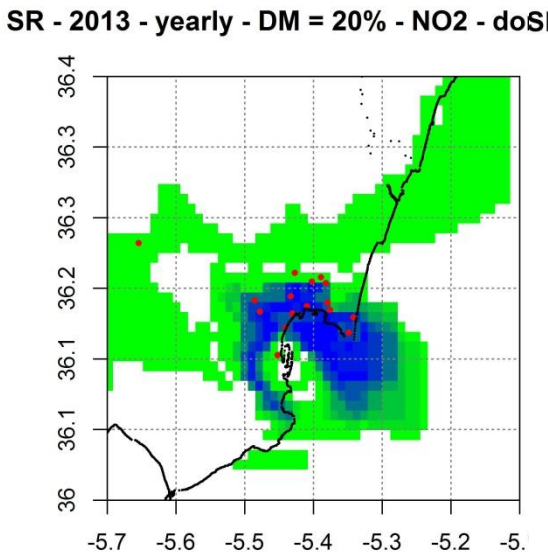


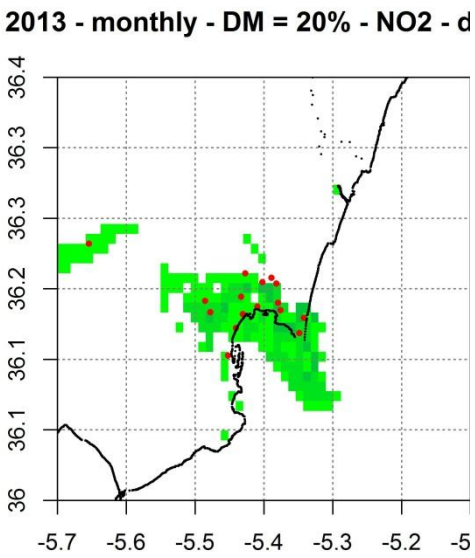


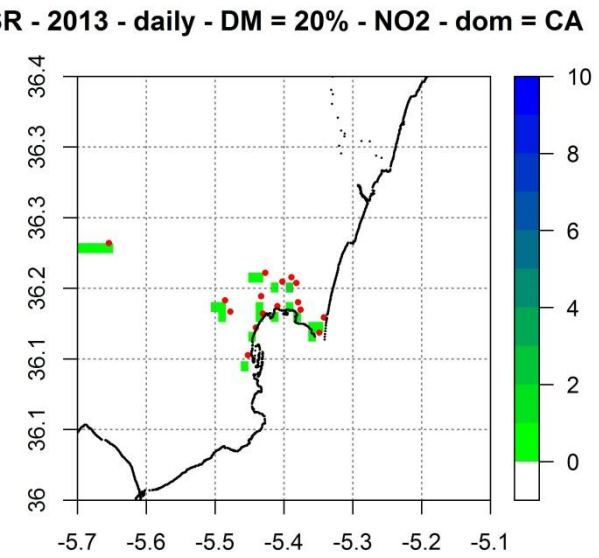
type	n	SRT	SRimd	SRimm	SRimy
All	13	735(d) / 2567(m) / 11087(y)	68.4±91.1	207.6±437.6	1039.8±2389.4
RF	1	-	195±NA	1590±NA	8990±NA
R	3	-	71±107.4	531.7±916.5	3254.3±4967.2
U	3	-	18±2.6	70.3±95.9	359.3±62
S	7	-	88.9±104.9	127.6±174.5	382.3±67.3

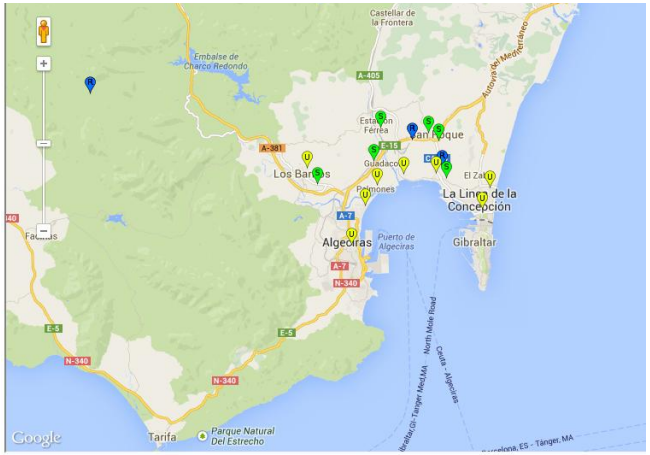
SRT = Total Spatial Representativeness (in km^2) of the monitoring network in daily (d), monthly (m) and yearly (y)
SRimd = Mean Spatial representativeness of a stations (in km^2) in a daily basis.
SRimm = Mean Spatial representativeness of a stations (in km^2) in a monthly basis.
SRimy = Mean Spatial representativeness of a stations (in km^2) in a yearly basis.

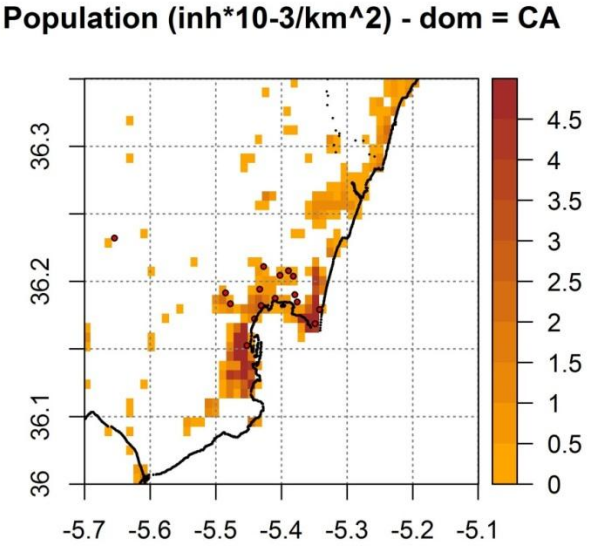


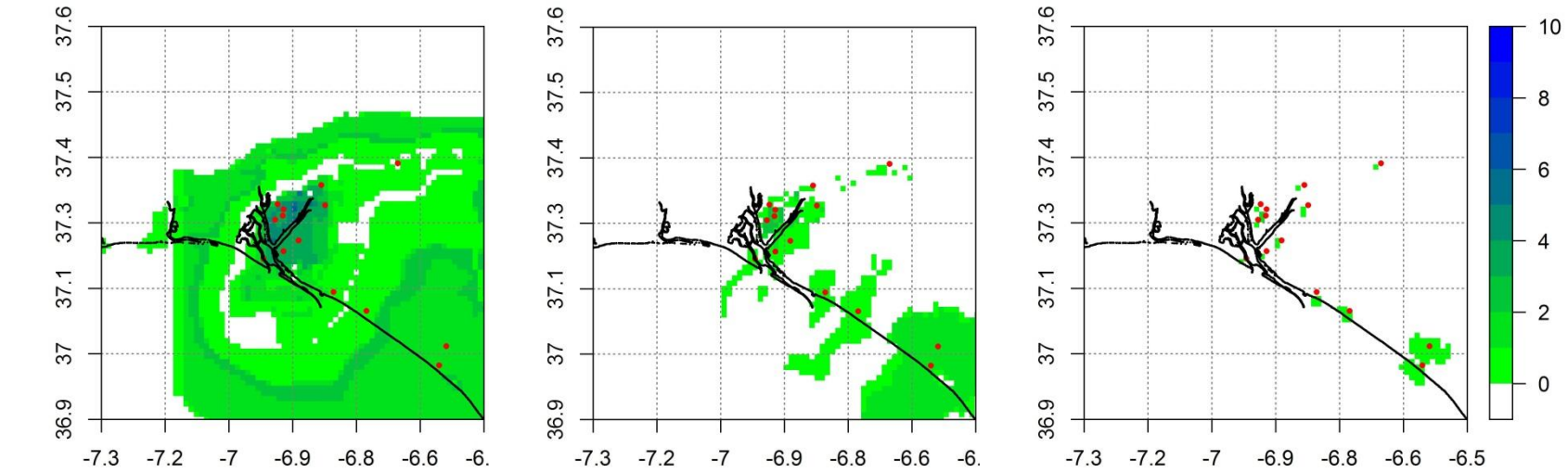

SR - 2013 - yearly - DM = 20% - NO2 - do
36.4
36.3
36.3
36.2
36.1
36
-5.7
-5.6
-5.5
-5.4
-5.3
-5.2
-5.

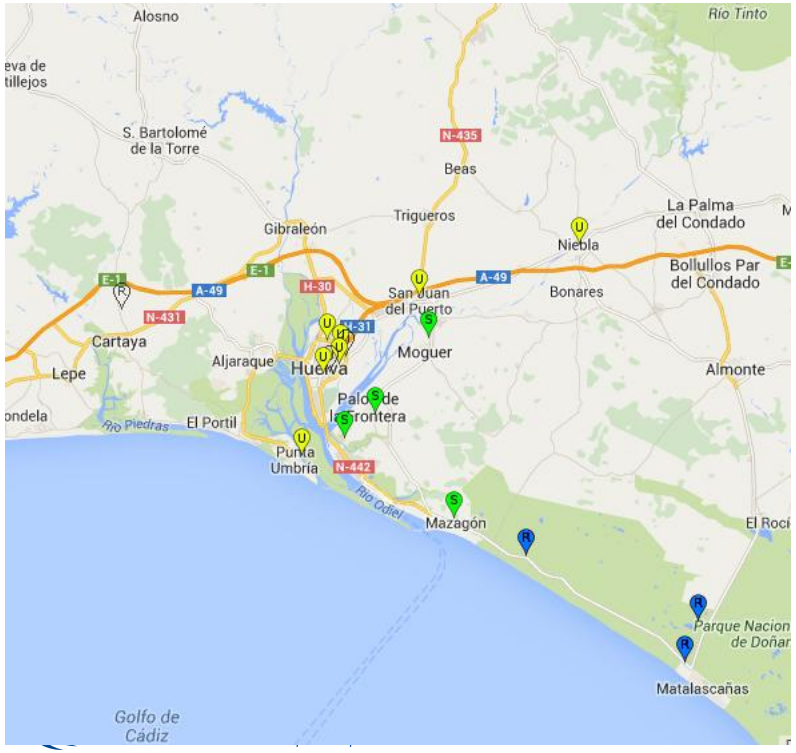

SR - 2013 - monthly - DM = 20% - NO2 - d
36.4
36.3
36.3
36.2
36.1
36
-5.7
-5.6
-5.5
-5.4
-5.3
-5.2
-5

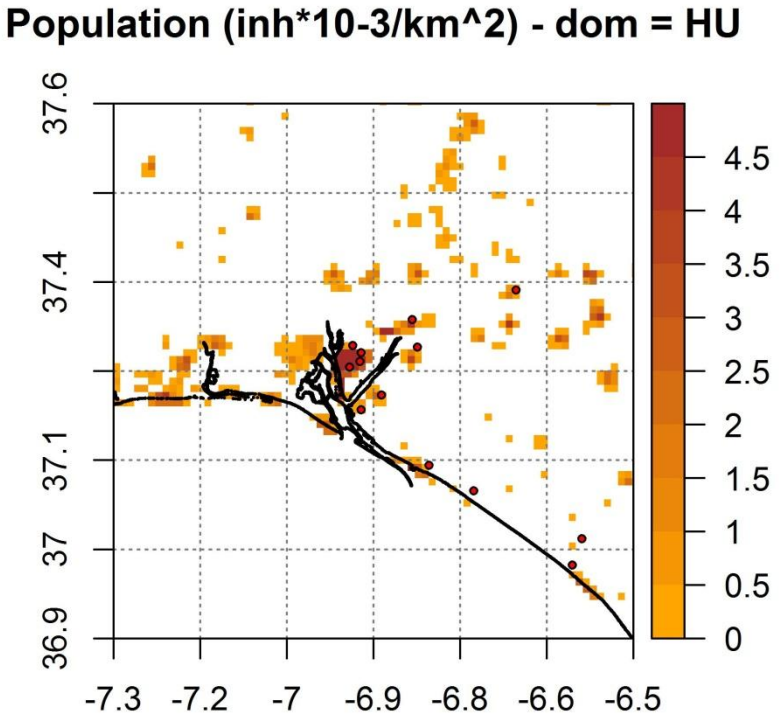

SR - 2013 - daily - DM = 20% - NO2 - dom = CA
36.4
36.3
36.3
36.2
36.1
36
-5.7
-5.6
-5.5
-5.4
-5.3
-5.2
-5.1
10
8
6
4
2
0


Castellar de la Frontera
Embalse de Charco Redondo
A-405
Estación Férrea
Los Barrios
Guadacorte
Palmones
Algeciras
Puerto de Algeciras
La Línea de la Concepción
Gibraltar
El Zabal
A-381
A-7
N-340
E-5
Tarifa
Parque Natural Del Estrecho
Google

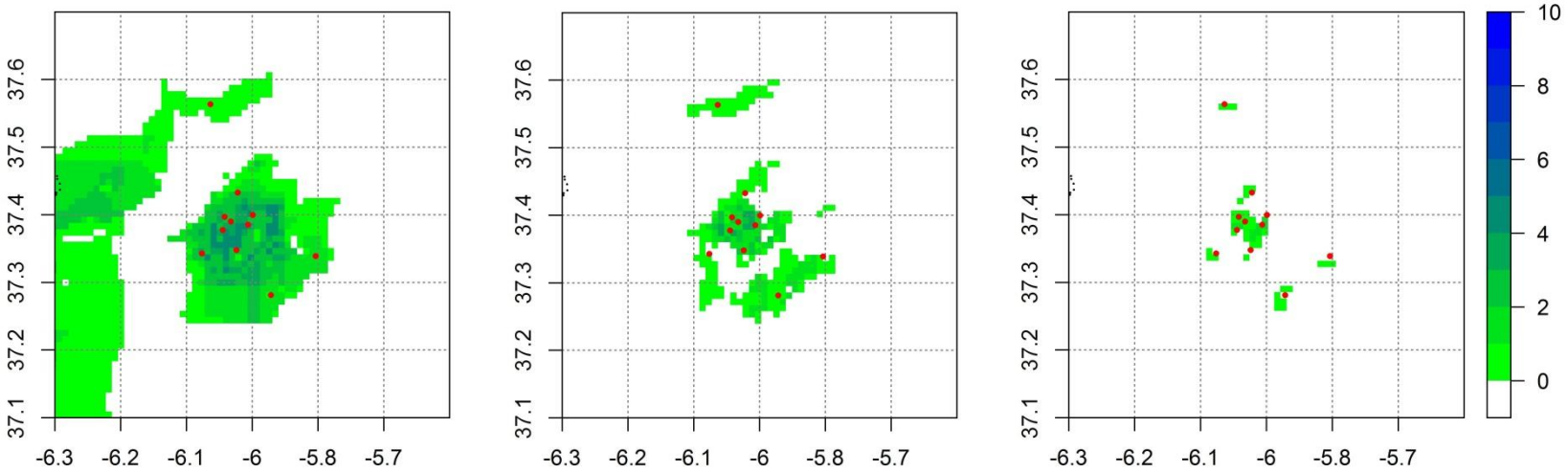

Population (inh*10-3/km^2) - dom = CA
36.3
36.2
36.1
36
-5.7
-5.6
-5.5
-5.4
-5.3
-5.2
-5.1
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0

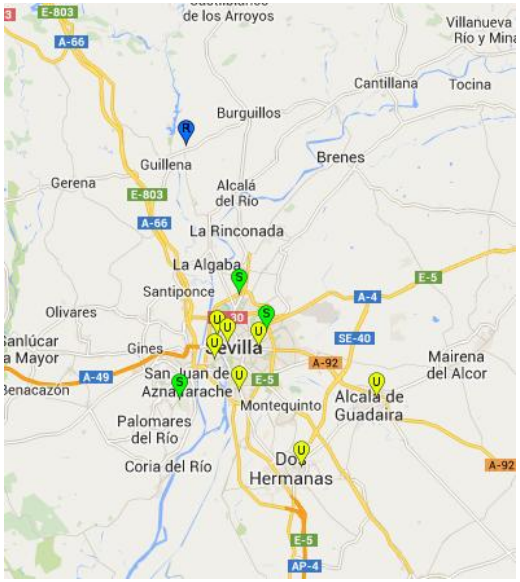

SR - 2013 - yearly - DM = 20% - NO2 - doSR - 2013 - monthly - DM = 20% - NO2 - d
SR - 2013 - daily - DM = 20% - NO2 - dom = HU
37.6
37.5
37.4
37.3
37.1
37
36.9
-7.3
-7.2
-7
-6.9
-6.8
-6.6
-6.5
10
8
6
4
2
0


Alosno
Río Tinto
S. Bartolomé de la Torre
Beas
Trigueros
Gibraleón
La Palma del Condado
Niebla
Bollullos Par del Condado
Bonares
San Juan del Puerto
Moguer
Cartaya
Aljaraque
Huelva
Lepe
Almonte
Palos de la Frontera
El Portil
Río Piedras
Punta Umbría
Río Odiel
Mazagón
Matalascañas
Golfo de Cádiz
N-435
N-431
N-442
A-49
H-30
H-31
E-1

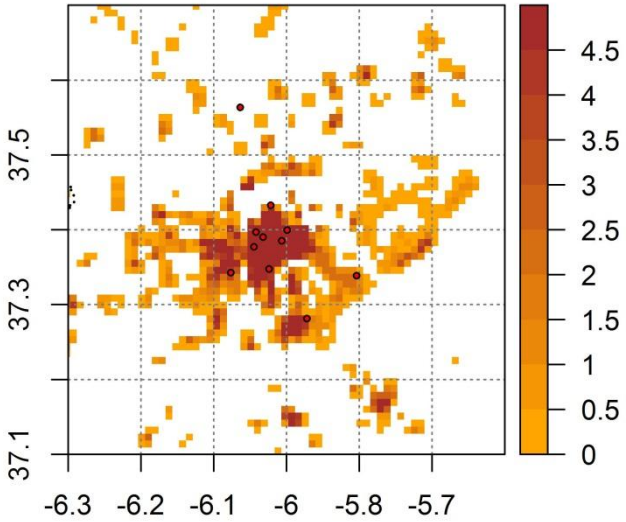

Population (inh*10-3/km^2) - dom = HU
37.6
37.4
37.1
37
36.9
-7.3
-7.2
-7
-6.9
-6.8
-6.6
-6.5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0

SR - 2013 - yearly - DM = 20% - NO2 - doSR - 2013 - monthly - DM = 20% - NO2 - d SR - 2013 - daily - DM = 20% - NO2 - dom = SE





Population (inh*10-3/km^2) - dom = SE



Summary and conclusions

- The “Modelled concentration similarity” methodology was used to estimate the spatial representativeness of the Andalucía air quality monitoring network using the AQFS CALIOPE 1-km simulations.
- Spatial representativeness is high sensitive to the 1) pollutant, 2) temporal basis, 3) station type, and 4) location.
- Several sensitivity tests were applied over the Andalucía monitoring network:
  - Test to temporal resolution (yearly, monthly and daily) → SR increases when the temporal resolution decrease, SR is higher for yearly basis than for monthly and daily basis.
  - Test to maxima discrepancy (DM = 5, 10, 15, 20 %) → According to the bibliographical review and the sensitivity test performed, 20% of DM for all the pollutant could be an conservative selection.
- Current O_3 stations the spatial coverage is high enough to monitoring areas with high concentrations. However, PM2.5 stations are not representative of air quality in highly dense populated areas such (e.g., Seville and Malaga).
- Consequently, the **spatial representativeness** (SR) is specific, and not universal