UNIÓN EUROPEA
Fondos estructurales
Invertimos en su futuro

Comunidad de Madrid

Programas de Actividades de I + D entre grupos de investigación de la Comunidad de Madrid en Tecnologías 2018

PAIDIR TEC 2018

AIRTEC-CM

EVALUACIÓN INTEGRAL DE LA CALIDAD DEL AIRE URBANO Y CAMBIO CLIMÁTICO

P2018/EMT-4329

Estimation of $PM_{2.5}$ ambient concentration in Madrid using XGBoost models

J.M. Cordero, J. Li, D. de la Paz, P. Koutrakis, R. Borge

3ª Reunión del Comité Cientifico-Técnico

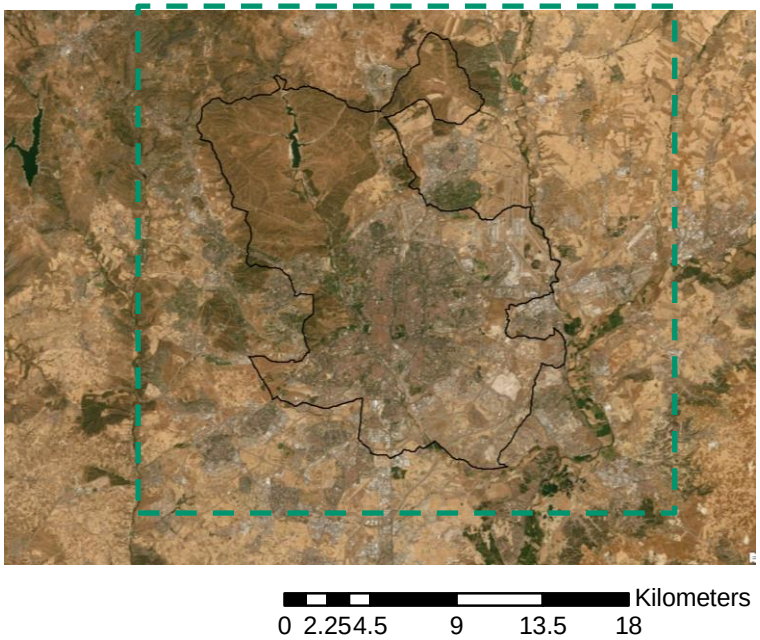
23 Noviembre 2021, CIEMAT, Madrid / zoom

OUTLINE

OUTLINE

- Domain and Variables } Starting information
- The 2-step-ML model } Modelling
- Results and Validation } Results and Evaluation
- Exposure } First studies on exposure
- Conclusions

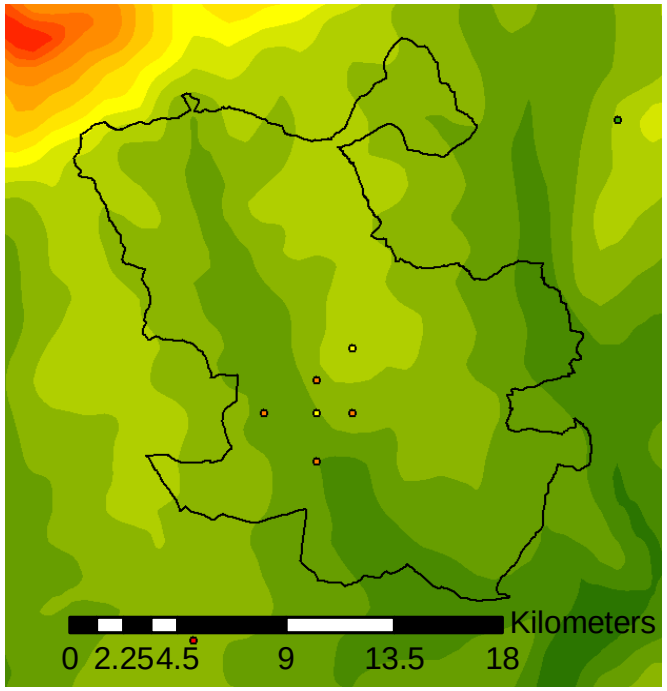
Domain: Municipality of Madrid and its surroundings, 1x1 km, 2015



elevation

- 500-600
- 600-650
- 650-700
- 700-750
- 750-800
- 800-850
- 850-900
- 900-950
- 950-1000
- +1000

m



Station	Lon	Lat	PM2.5_avg	PM2.5_max	Type
E. Aguirre	-3.6887207	40.4220314	12.32	39	Traffic
Casa de Campo	-3.7516174	40.4216385	9.5	36	Suburban
Cuatro Caminos	-3.7140808	40.4410362	10.66	37	Traffic
Pza. Castilla	-3.6891174	40.4603386	9.96	34	Traffic
Méndez Álvaro	-3.713562	40.3931465	11.1	39	Background
Castellana	-3.7138672	40.4218788	10.54	33	Traffic
Leganés	-3.8003235	40.2872162	13.38	43.47	Traffic
Algete	-3.5012817	40.5954056	6.66	29.12	Background

Maximums

- > 35 μg/m^3
- > 30 μg/m^3
- < 30 μg/m^3

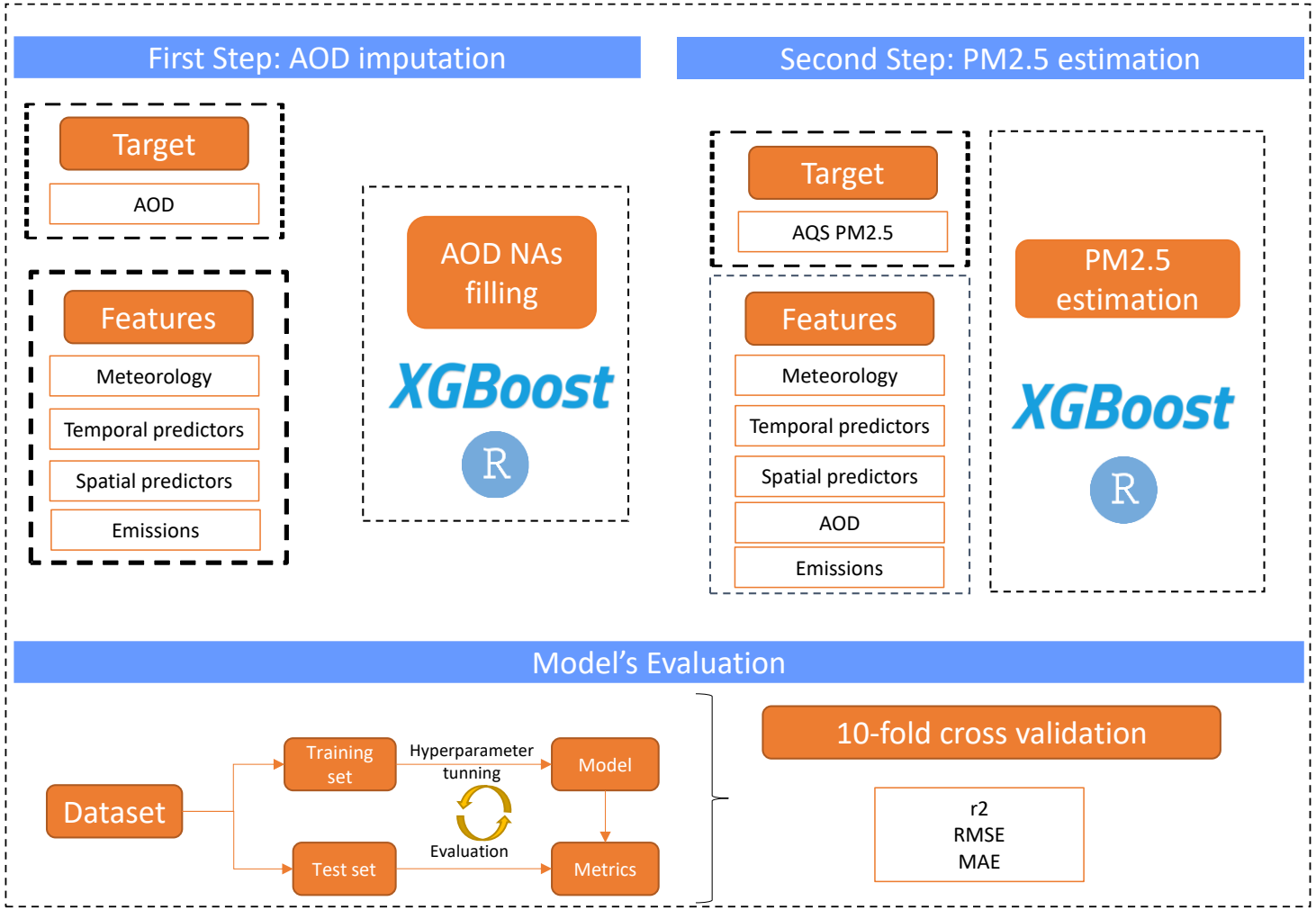
Independent Variables

Variable	Units	Type
Elevation	m	
Traffic	Vh_km	
Land Use	WRF 33 cat	
Population	Hab/km2	Spatial
Wind speed	m/s	Temporal
SLP	hPa	Temporal
q (specific humidity)	g/kg	
Temperature	K	
NDVI MOD	MOD13A3	
PBL Heigth	m	
Layer 1	g/s	
Layer 2	g/s	Emission
Resuspension	g/s	Emission

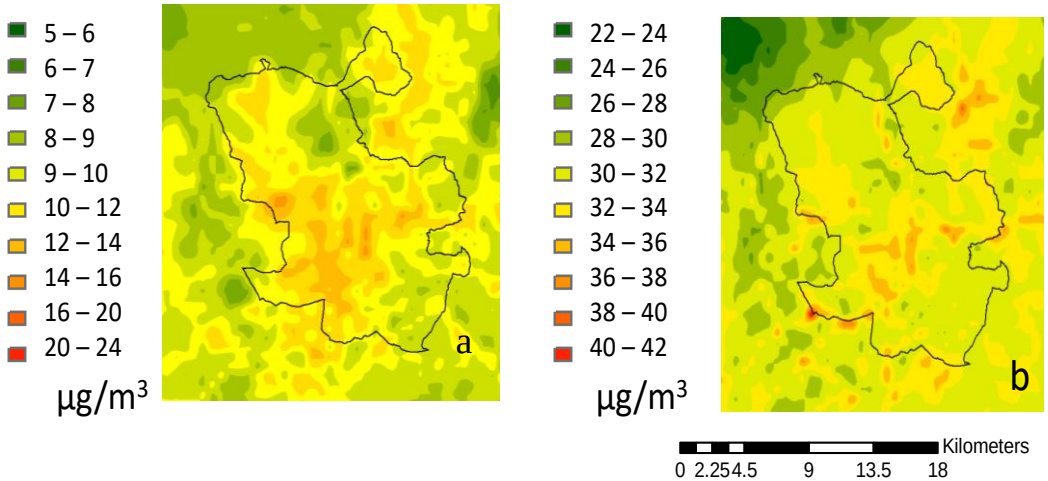
✓ Dependent variable: AOD (MAIAC MCD19A2 collection 06)

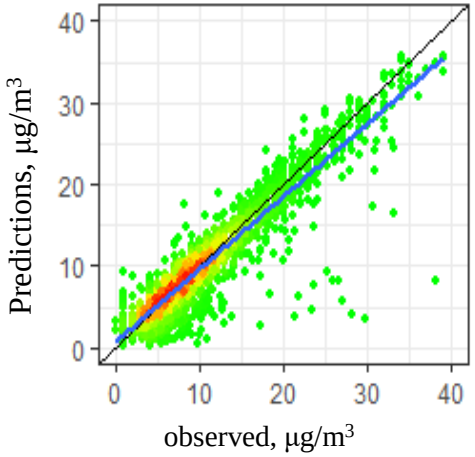
✓ Dependent variable: PM2.5 (Ground AQS Observations)

2-step-ML-algorithm



Results 1/2



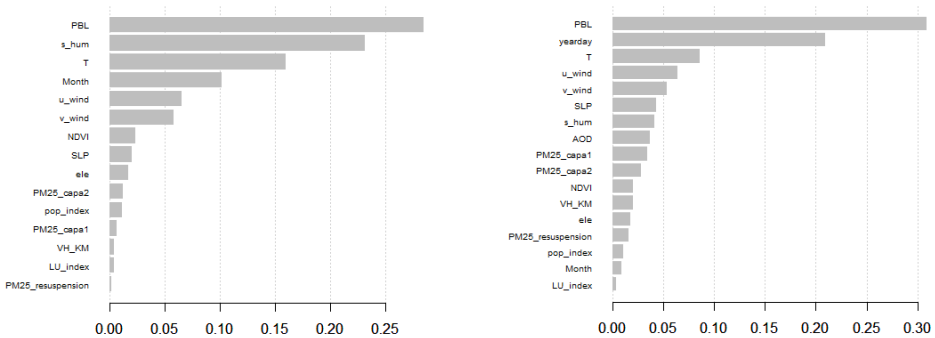


a Map of yearly PM2.5 averaged concentrations, the contribution of main highways can be seen
b Day of máximum PM2.5 concentration due to saharian dust episode advection,
The prediction of AEMET was 40 µg/m3

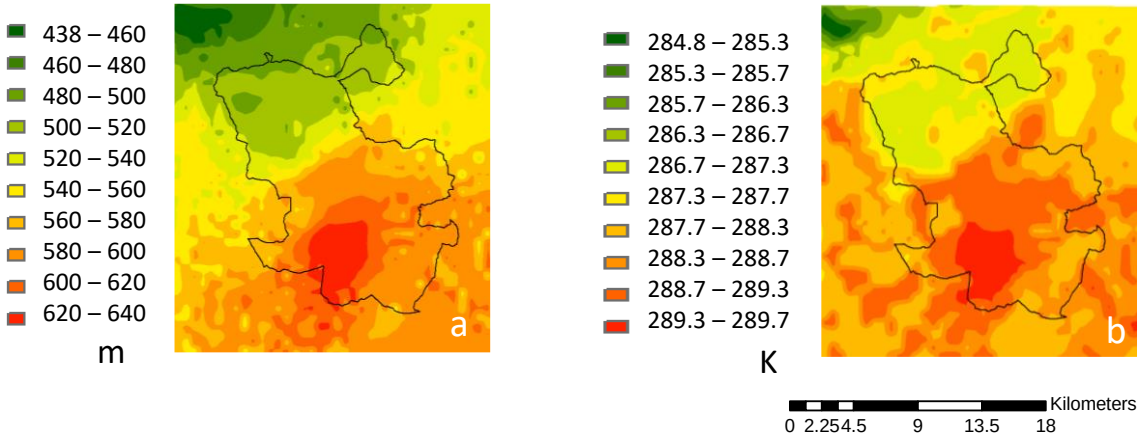
Results of the 10-fold cross-validation

r2	RMSE (µg/m3)	MAE (µg/m3)
0.95±0.007	1.51±0.14	1.06±0.07

Results 2/2



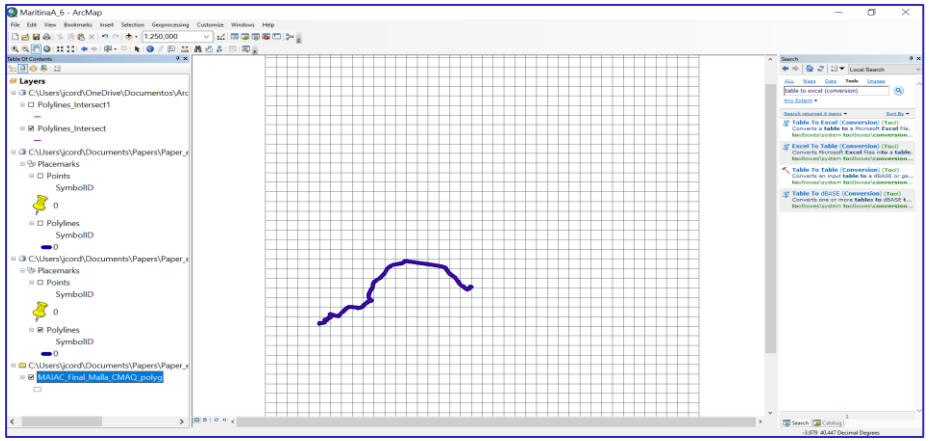
% Relative Importance for the models



2015 averaged PBL Heigth (a) and temperature (b) at the domain studied.

Deduction→ With a large city occupying a considerable space of the Domain, PBLH (T, friction…) affects hugely the AODs distribution.

Exposure 1/4



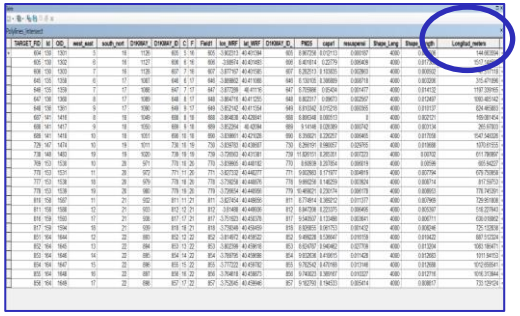
Person A trajectory

Grid with PM2.5 info

Combination

To .xlsx

Length_m by cell



Exposure 2/4

1) Assuming no daily variations in the PM2.5 concentrations

$$\overline{PM2.5} = \frac{\left(C_{rel_{global_{casa}}} * 13.5\,h + C_{rel_{global_{trabajo}}} * 9\,h + 2x45/60 \sum C_{rel_{global_{celda}}} \frac{time_s}{t_mov}\right)}{24\,h} C_global_mean$$

home work travel

µg/m³ d

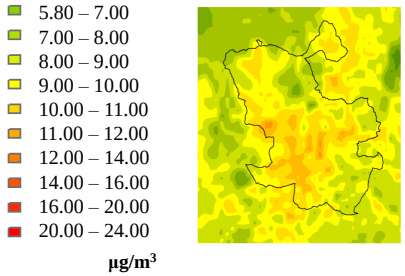
2) Introducing daily variations in the PM2.5 concentrations with ft

$$Exposure = \frac{\left[\sum_i^{1:9,21:24}(C_{PM2.5})\,(f_{t,i}) * 1h + \sum_i^{11:19}(C_{PM2.5})\,(f_{t,i}) * 1h + \left(\sum_{i=2}^{N-2}(C_{PM2.5}, i)\,(f_{t,10})\frac{time_{s,i}}{t_{mov}}\right)1h + \left(\sum_{i=2}^{N-2}(C_{PM2.5}, i)\,(f_{t,20})\frac{time_{s,i}}{t_{mov}}\right)1h\right]}{24}$$

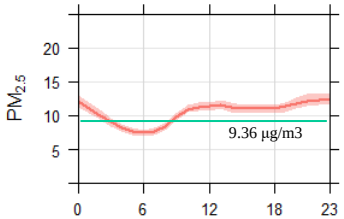
home work travel

µg/m³ d

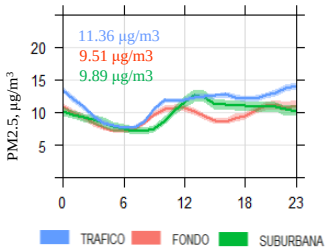
Estimated yearly PM2.5 concentrations



Pool of AQS



By types of AQS



Exposure 3/4

1) Assuming no daily variations in the PM2.5 concentrations

$$\overline{PM2.5} = \frac{\left(C_{rel_{global_{casa}}} * 13.5\,h + C_{rel_{global_{trabajo}}} * 9\,h + 2x45/60 \sum C_{rel_{global_{celda}}} \frac{time_s}{t_mov}\right)}{24\,h} C_global_mean$$

home work travel

µg/m³ d

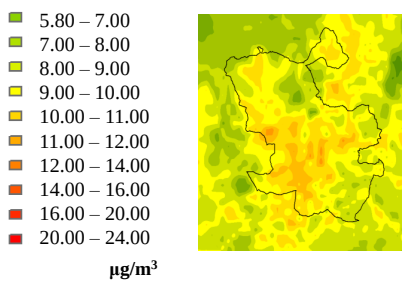
2) Introducing daily variations in the PM2.5 concentrations with ft

$$Exposure = \frac{\left[\sum_i^{1:9,21:24}(C_{PM2.5})(f_{t,i}) * 1h + \sum_i^{11:19}(C_{PM2.5})(f_{t,i}) * 1h + \left(\sum_{i=2}^{N-2}(C_{PM2.5}, i)(f_{t,10})\frac{time_{s,i}}{t_{mov}}\right)1h + \left(\sum_{i=2}^{N-2}(C_{PM2.5}, i)(f_{t,20})\frac{time_{s,i}}{t_{mov}}\right)1h\right]}{24}$$

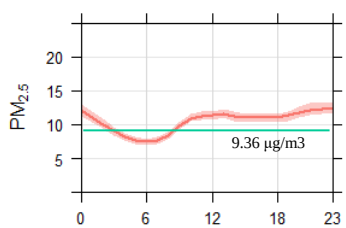
home work travel

µg/m³ d

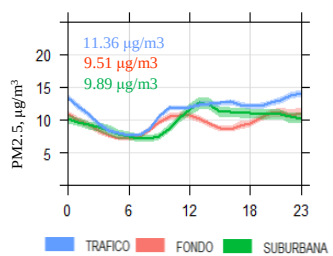
Estimated yearly PM2.5 concentrations



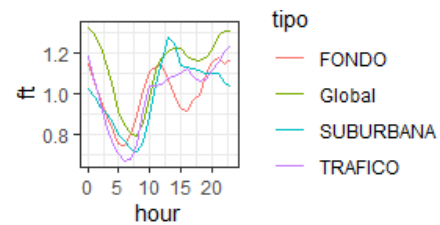
Pool of AQS



By types of AQS



$$ft = \frac{C_{PM2.5,Perfil}}{C_{global_{mean}}}$$



Exposure 4/4

PERSON A

PM2.5 in µg/m3 d				
Assuming no daily variations in the PM2.5 concentrations				
Global		Home		
9.10		8.97		
Introducing daily variations in the PM2.5 concentrations with ft				
Global	Traffic	Background	Suburban	Home
10.14	9.20	9.27	9.21	10.07

PERSON B

PM2.5 in µg/m3 d				
Assuming no daily variations in the PM2.5 concentrations				
Global		Home		
10.14		9.13		
Introducing daily variations in the PM2.5 concentrations with ft				
Global	Traffic	Background	Suburban	Home
10.01	8.88	8.92	8.91	10.25

Conclusion

- ✓ Methodology presented and validated
- ✓ Able to be used in every other place
- ✓ Future exposure can be assessed
- ✓ The 10-fold cross validation was performed to see the robustness of the models and the independency of the dataset used for training, obtaining an R2 of 0.95 ± 0.007, RMSE of 1.51 ± 0.14 and MAE of 1.06 ± 0.07. The mentioned results qualified the models as high accurate and their PM2.5 estimations as high reliable.

- ✓ Try other ML approachs
- ✓ Use the obtained results to assess people's exposure

UNIÓN EUROPEA
Fondos estructurales
Invertimos en su futuro

Comunidad de Madrid

Programas de Actividades de I + D entre grupos de investigación de la Comunidad de Madrid en Tecnologías 2018

PAIDIR TEC 2018

AIRTEC-CM

EVALUACIÓN INTEGRAL DE LA CALIDAD DEL AIRE URBANO Y CAMBIO CLIMÁTICO

P2018/EMT-4329

3ª Reunión del Comité Cientifico-Técnico

23 Noviembre 2021, CIEMAT, Madrid / zoom