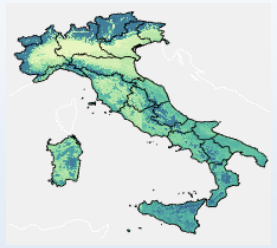


National Research Council of Italy, Institute of Atmospheric Sciences and Climate (CNR-ISAC)

NAPOLI PM 2026

26-29 May 2026

Long-term high-resolution mapping of air pollutants over Italy via machine learning (2002–2024)

Karam Mansour, Matteo Rinaldi, Marco Paglione, Paolo Stocchi,

Stefano Decesari, and Tony C. Landi

National Research Council of Italy, Institute of Atmospheric Sciences and Climate (CNR-ISAC), Bologna, 40129, Italy

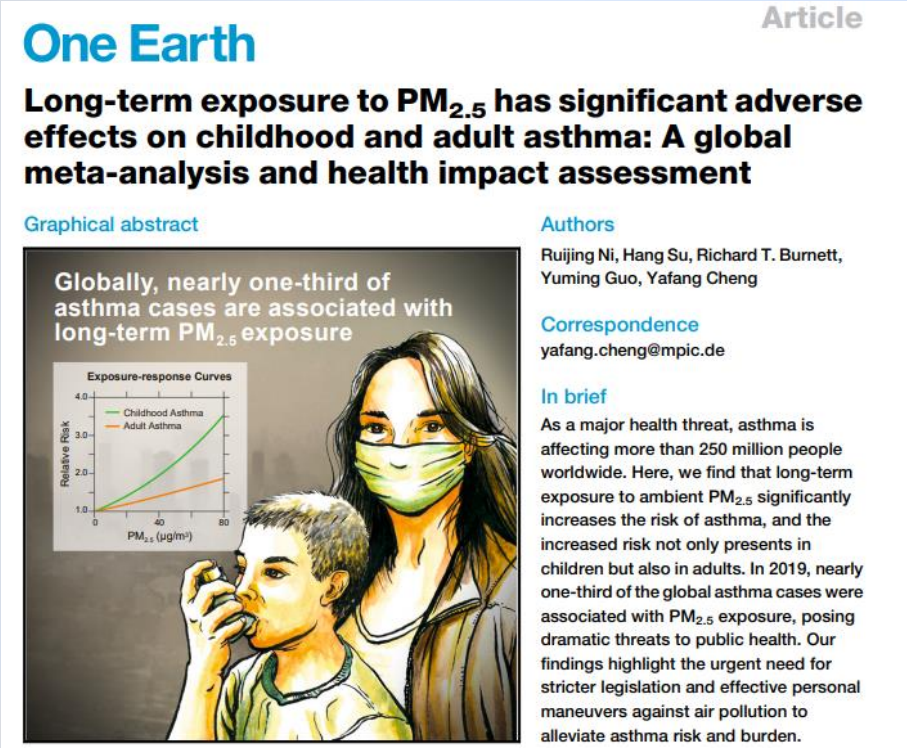
Correspondence: k.mansour@isac.cnr.it

Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani

Air pollution is one of the main environmental health risks



Article

One Earth

Long-term exposure to $PM_{2.5}$ has significant adverse effects on childhood and adult asthma: A global meta-analysis and health impact assessment

Graphical abstract

Globally, nearly one-third of asthma cases are associated with long-term $PM_{2.5}$ exposure

Exposure-response Curves

Childhood Asthma

Adult Asthma

Relative Risk

$PM_{2.5}$ (μg/m³)

Authors

Ruijing Ni, Hang Su, Richard T. Burnett, Yuming Guo, Yafang Cheng

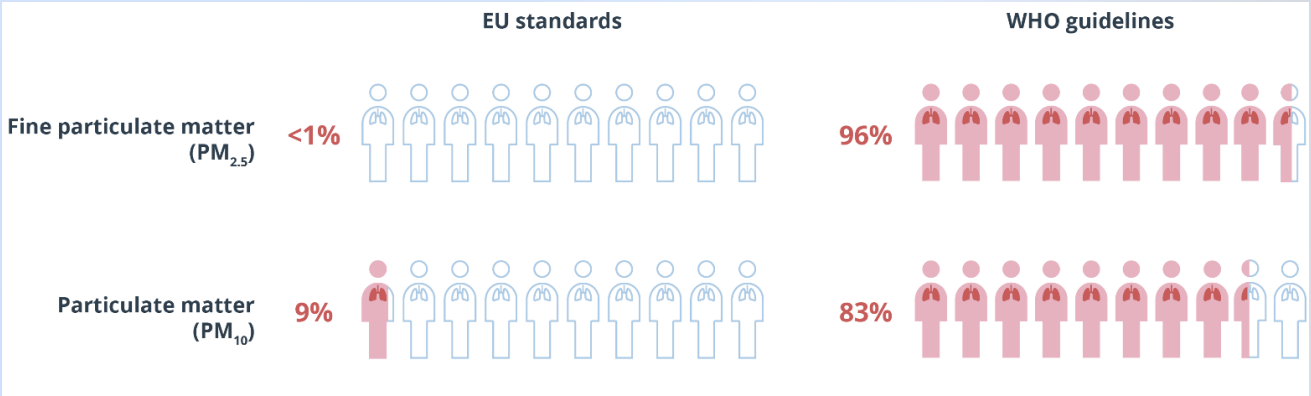
Correspondence

yafang.cheng@mpic.de

In brief

As a major health threat, asthma is affecting more than 250 million people worldwide. Here, we find that long-term exposure to ambient $PM_{2.5}$ significantly increases the risk of asthma, and the increased risk not only presents in children but also in adults. In 2019, nearly one-third of the global asthma cases were associated with $PM_{2.5}$ exposure, posing dramatic threats to public health. Our findings highlight the urgent need for stricter legislation and effective personal maneuvers against air pollution to alleviate asthma risk and burden.

- **Globally, nearly one-third of asthma cases are associated with long-term fine PM exposure.**



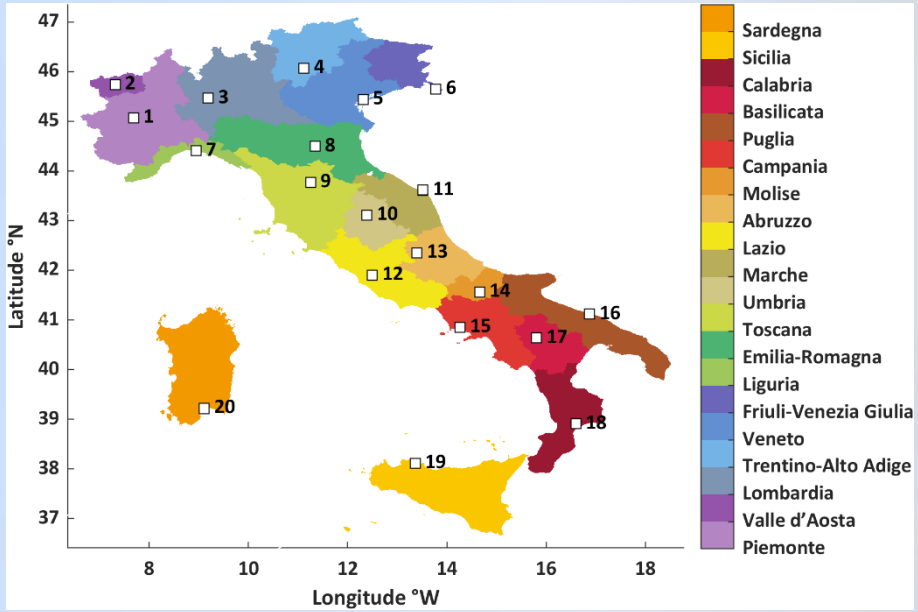
The EU urban population exposed to air pollutant concentrations above certain EU standards and WHO guidelines in 2022 (*Source: EEA, 2024*).

- **96% of the EU's urban population is exposed to unsafe levels of fine PM ($PM_{2.5}$) above WHO guidelines.**

Objectives

- To assess such impacts of air quality on different sectors (public health policies and environmental planning, etc.…), high-resolution maps of air pollutants concentrations are essential.
- Applying state-of-the-art machine learning algorithms to construct daily, long-term, 1-km gridded maps of near-surface concentrations of the main atmospheric pollutants:

1. PM_{10} (particles ≤ 10 µm aerodynamic diameter)
2. $PM_{2.5}$ (particles ≤ 2.5 µm aerodynamic diameter)
3. NO_2 (Nitrogen dioxide)
4. O_3 (Ozone)



Longitude = 1,192

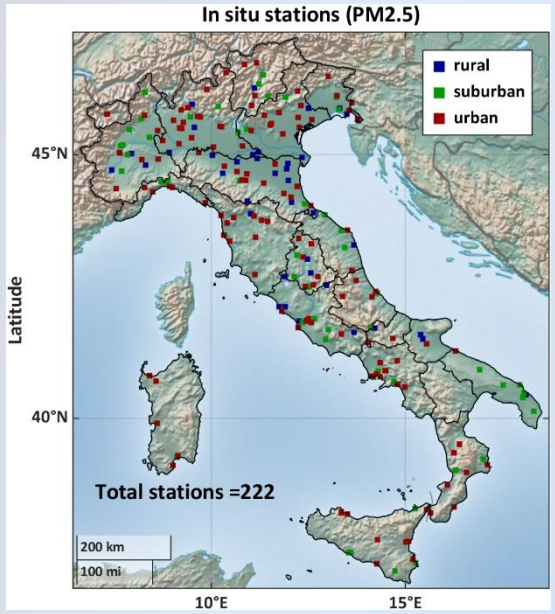
Latitude = 1,045

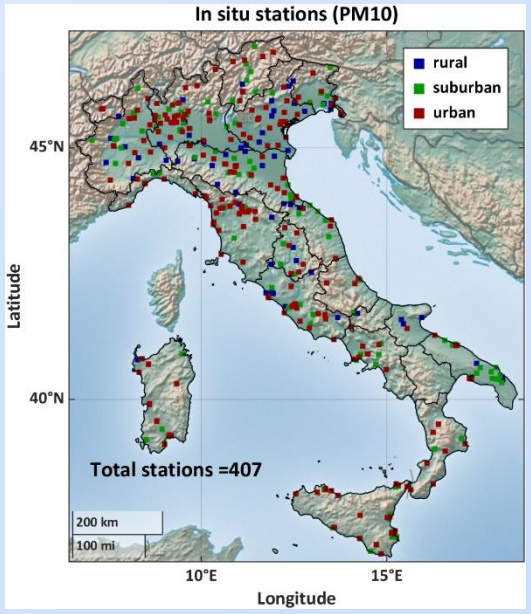
valid grid cells (Earth) = 332,346

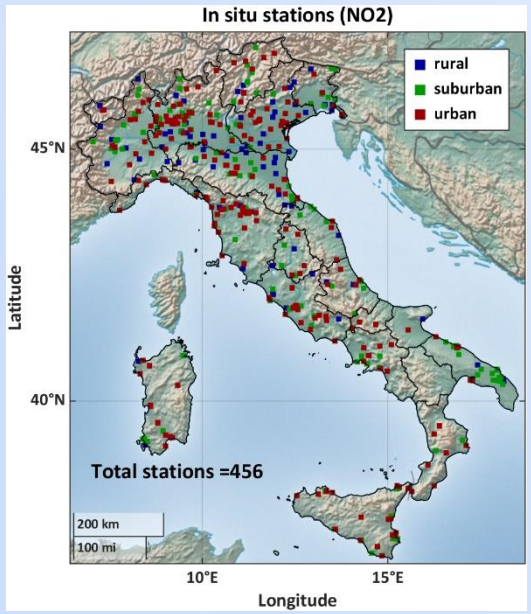
- Study Domain: The Italian territory
- Spatial resolution: 0.01° × 0.01° (~1 km)
- Covering period: 2002-2024 (*possible extension*)

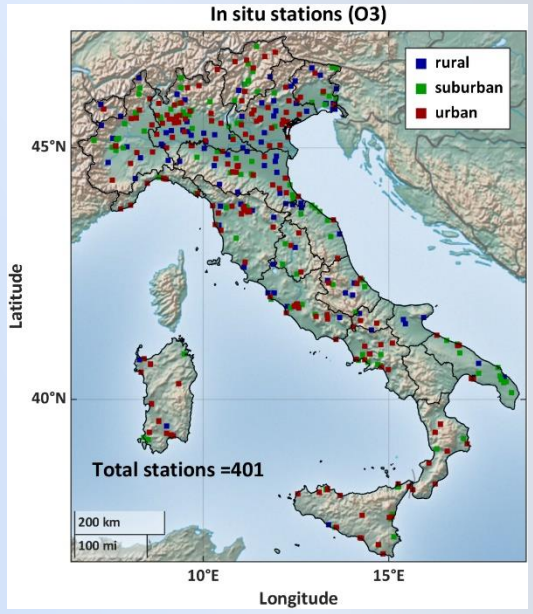
In Situ Observations (Target Responses)

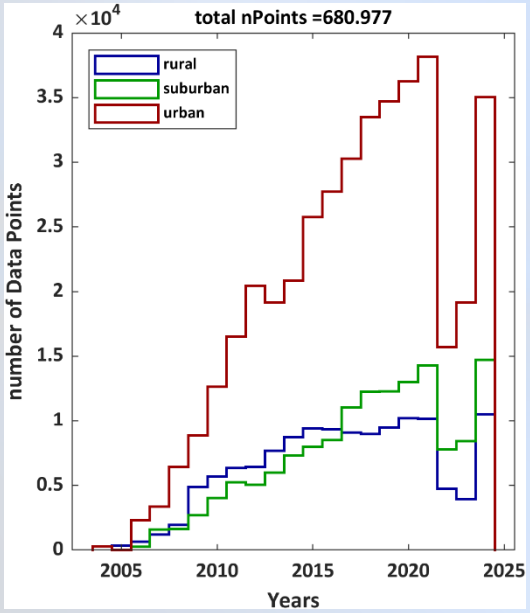
Source: European Air Quality Portal (AQP), managed by the European Environmental Agency (EEA)

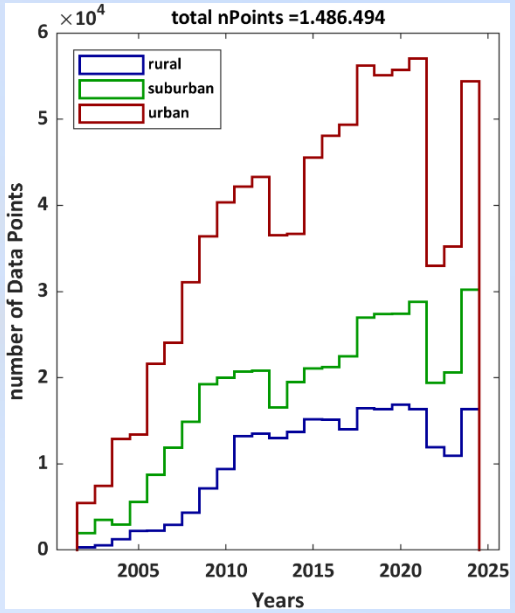


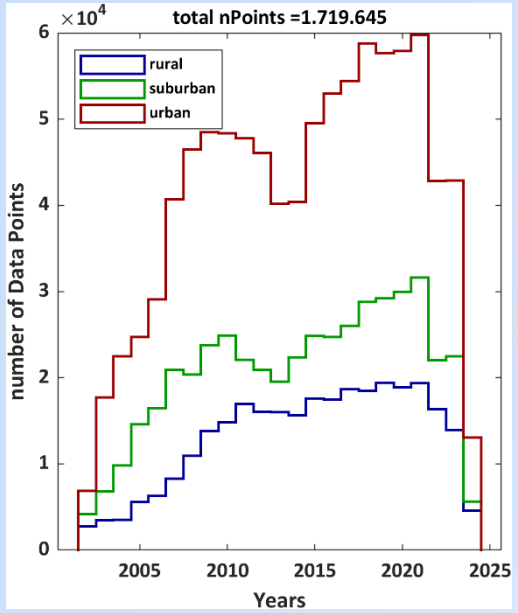


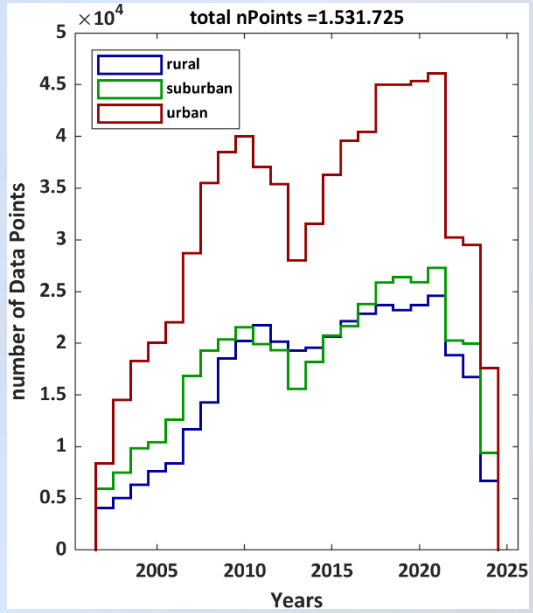






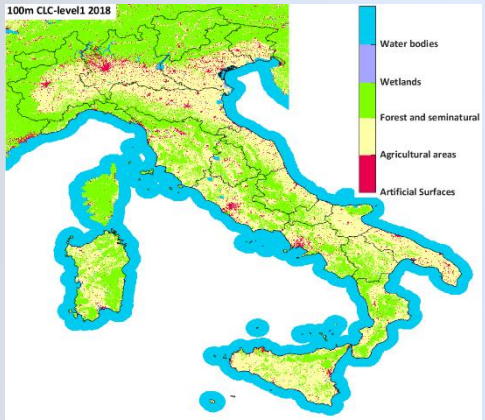




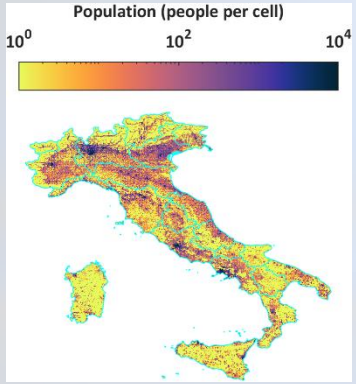


Predictors

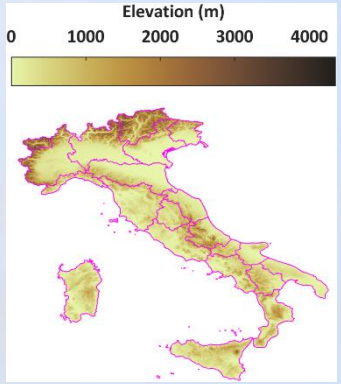
Spatial



land cover/use with 44 thematic classes
Source: CORINE Land Cover
6-Yearly - 100m (1990-onging)

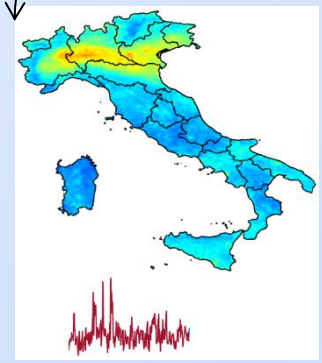


GHSL - Global Human Settlement Layer
6-Yearly– 100m (1990-ongoing)

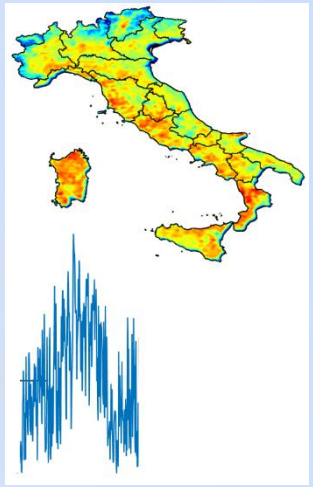


NASA Digital Elevation Map (DEM) (100m)

Spatiotemporal



LGHAP v2 - AOD
Daily - 1 km (2000-ongoing)
https://doi.org/10.5194/essd-16-2425-2024



Normalized Difference Vegetation Index (NDVI)
Source: NOAA
Daily-0.05° (1981-ongoing)

https://doi.org/10.5194/essd-2025-789

Abstract | Assets | Discussion | Metrics

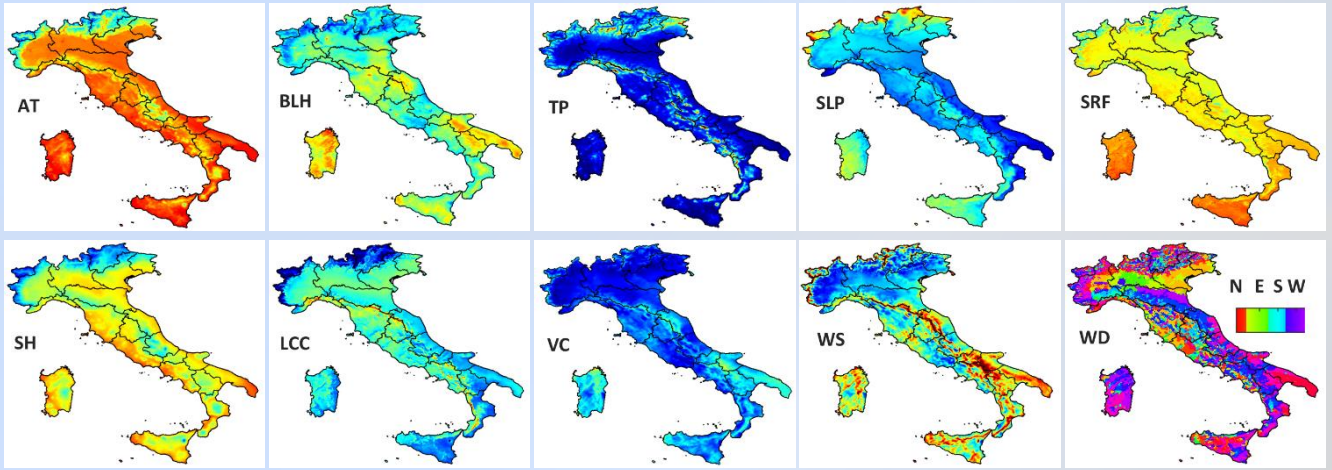
31 Mar 2026

Status: this preprint is currently under review for the journal ESSD.

MORE, a new convection-permitting reanalysis dataset over Italy and Alpine region: Validation and application in weather, climate and hydrology

Paolo Stocchi, Francesco Cavalleri, Mohsin Tariq, Michele Brunetti, Stefania Camici, Daniele Mastrangelo, Fabio Di Sante, and Silvio Davolio

Abstract. This study presents a new convection-permitting reanalysis dataset over Italy and Alpine region, produced through dynamical downscaling of ERA5 reanalysis with the non-hydrostatic mesoscale model MOLOCH. MORE (MOloch-downscaled ERA5 REanalysis) is a very high spatial resolution (~1.7 km) gridded dataset, covering the 1990–present period. The dataset includes hourly outputs of a wide range of variables at the surface and on pressure-levels.



Methodological Framework

Input

Predictors

Spatial

Land Cover (LC)
(*Agriculture + Artificial + Forest*)
Population Density
Digital Elevation Map (DEM)

Spatio-temporal

- Normalized Difference Vegetation Index (NDVI)
- Aerosol Optical Depth (AOD)
- Air Temperature (AT)
- Boundary Layer Height (BLH)
- Total Precipitation (TP)
- Surface Pressure (PMSL)
- Solar Radiation Flux (SSF)
- Specific Humidity (SH)
- Low Cloud Cover (LCC)
- Ventilation Coefficient (VC)
- Wind Speed (WS)
-Wind Speed (WD)

Temporal

Day of year
Weekdays
Years

Response(s)

- $PM_{2,5}$
- PM_{10}
- NO_2
- O_3

Training/ cross-validation approach

All Data

Training (80%) | Validation (20%)

Training #1 (subsets A, B, C & D)	**Test #1** (subset E)
Training #2 (subsets E, A, B & C)	**Test #2** (subset D)
Training #3 (subsets D, E, A & B)	**Test #3** (subset C)
Training #4 (subsets C, D, E & A)	**Test #4** (subset B)
Training #5 (subsets B, C, D & E)	**Test #5** (subset A)

We implement three types of CV:

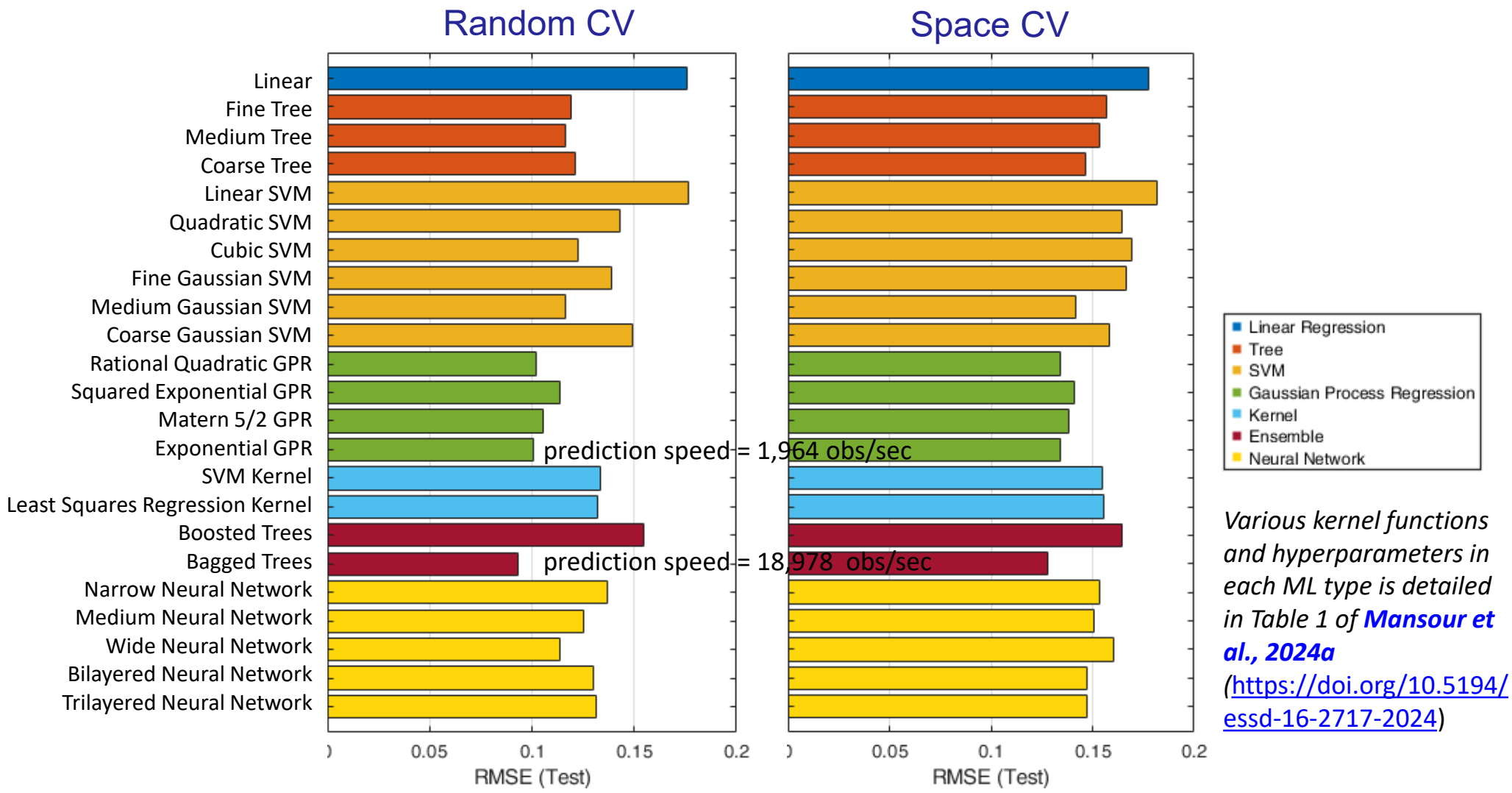
- **Random 5-fold CV:** this method randomly divides all data into five equal-sized subsets.
- **Space-block 5-fold CV:** the data are divided based on spatial location (stations).
- **Time-block 5-fold CV:** the data are divided based on measurement time (years).

Info on cross-validation and Training/Test Split found in:
Mansour et al., 2023 (https://doi.org/10.1016/j.scitotenv.2023.162123)
Mansour et al., 2024a (https://doi.org/10.5194/essd-16-2717-2024)

Info on Nested cross-validation for ML evaluation on limited datasets found in:
Mansour et al., 2024b (***https://www.nature.com/articles/s41612-024-00830-y***)

List of Trained ML Models

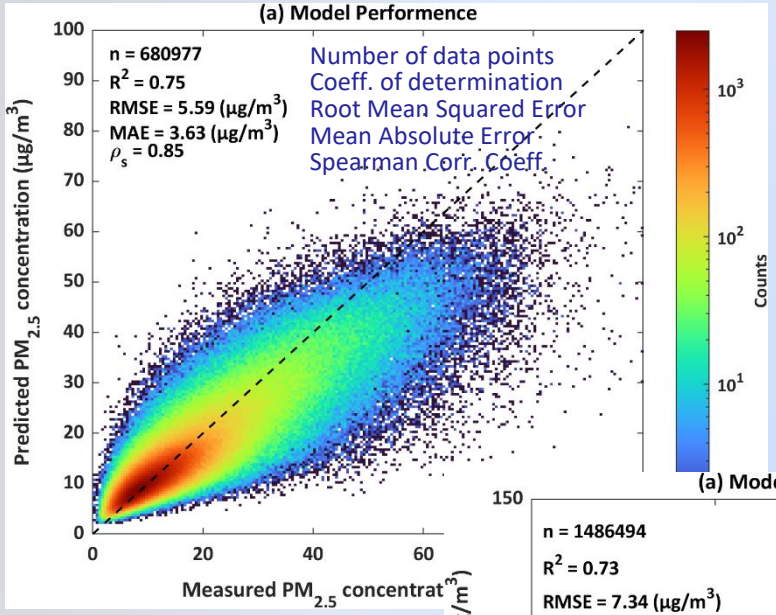
Example of training 23 ML models on PM_{10} in 2022. Based on the RMSE, the best model is Ensemble Bagged Trees (EBT). It offered strong performance (based on both Random and Space CV) and higher predictive speed which is important when considering Big data analysis. Therefore, EBT was selected as the optimal regressor.



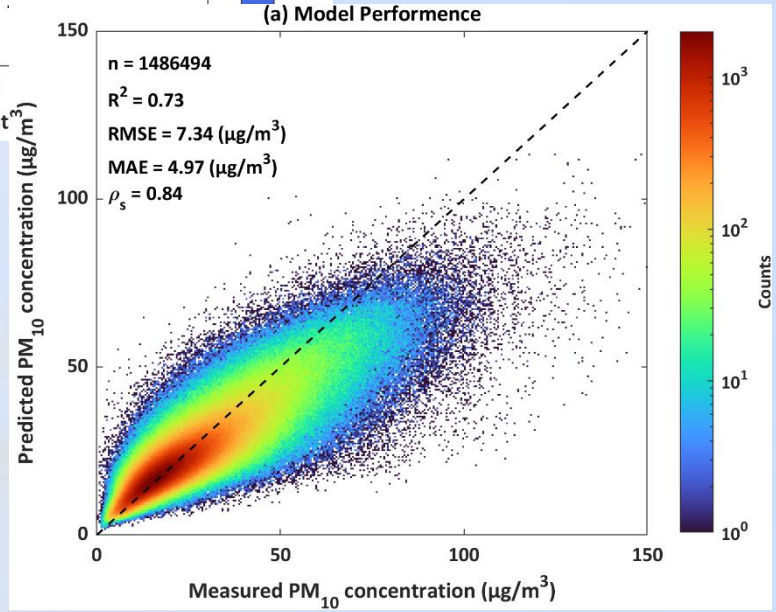
Various kernel functions and hyperparameters in each ML type is detailed in Table 1 of ***Mansour et al., 2024a*** (https://doi.org/10.5194/essd-16-2717-2024)

Results: performance of EBT (Random CV)

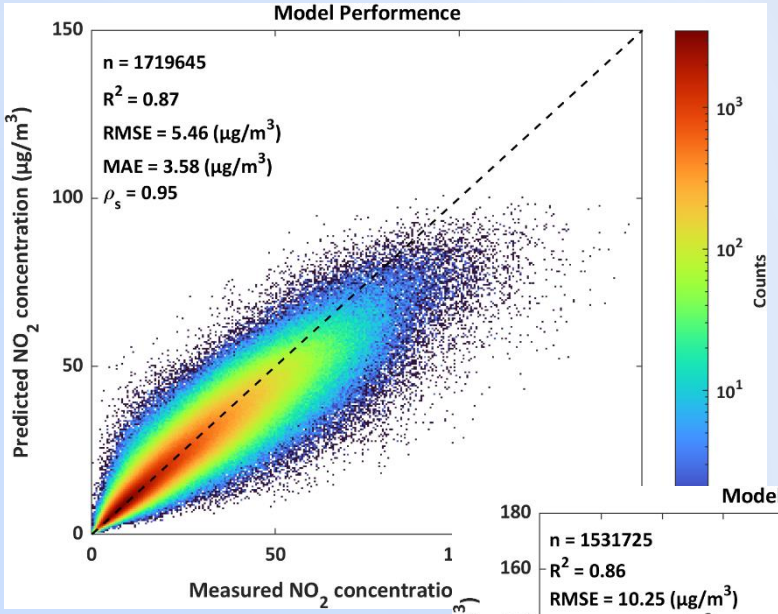
$PM_{2.5}$



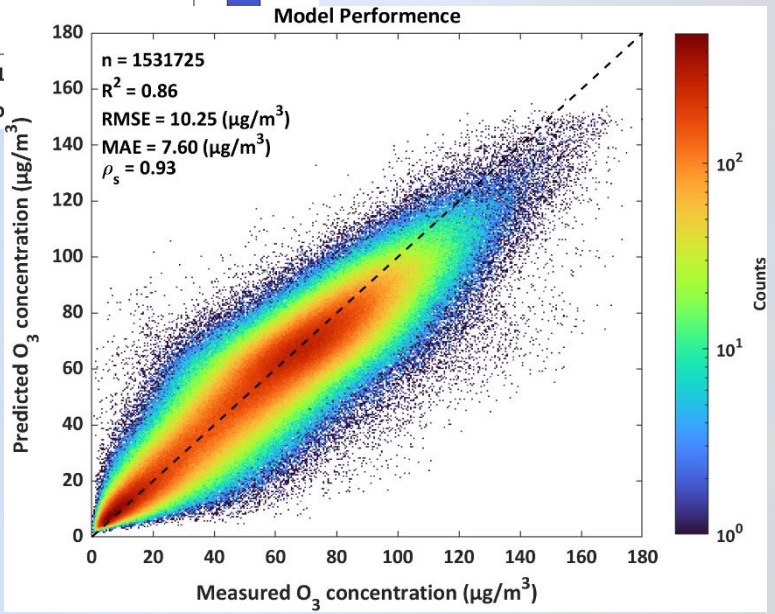
PM_{10}



NO_2

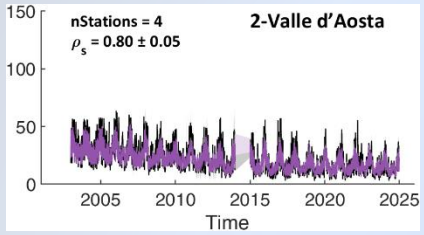


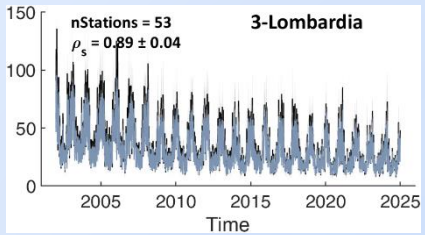
O_3

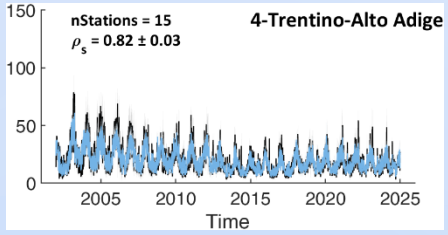


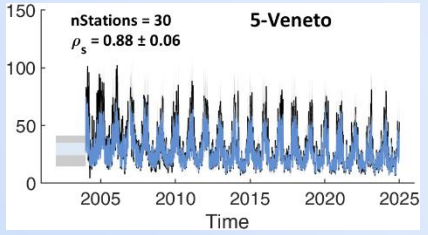
Comparison of predicted and measured daily air pollutants based on EBT model. The validation subsets are used to compute the model's performance.

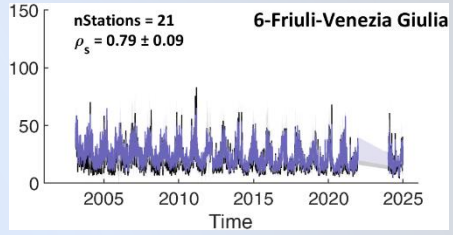
Results: performance of EBT (Random CV)

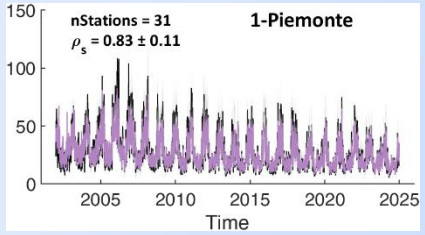


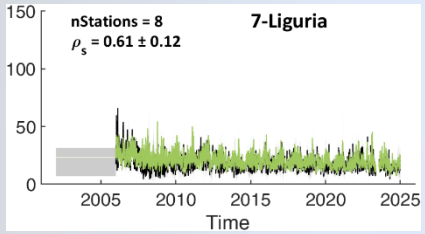


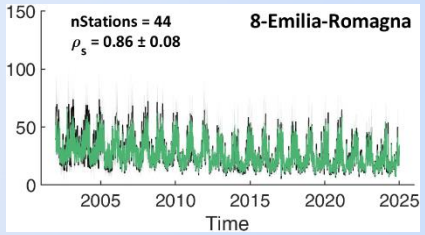


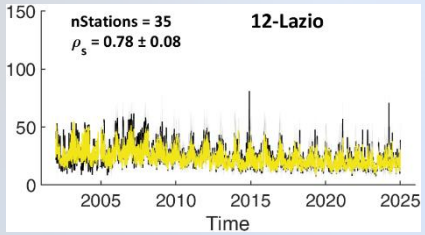


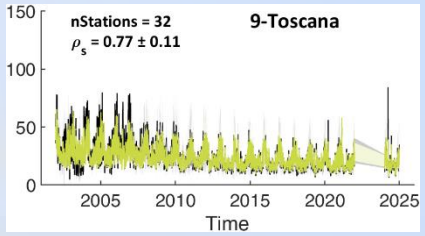


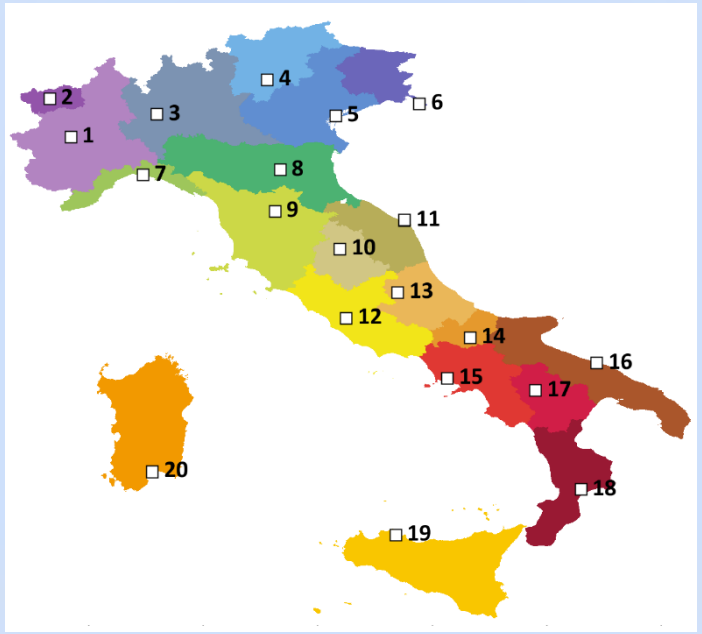


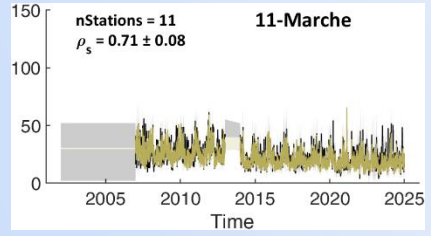


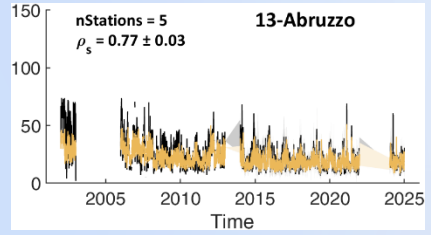


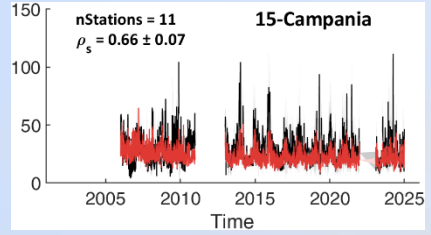


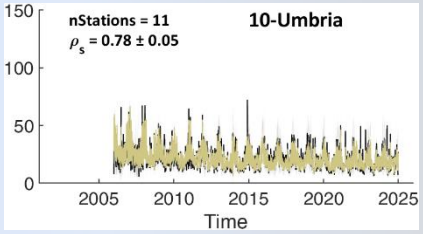


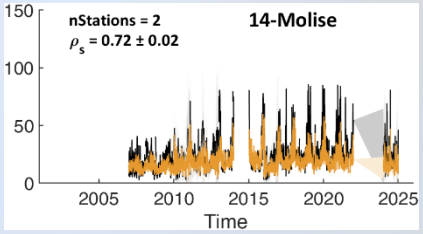


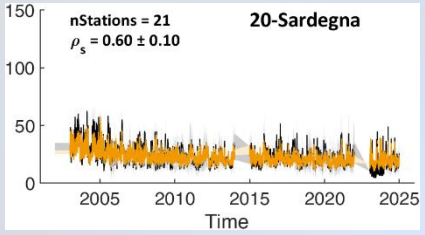


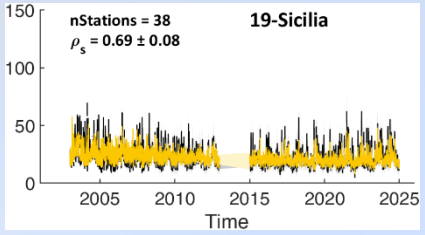


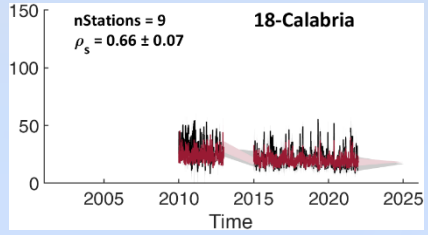


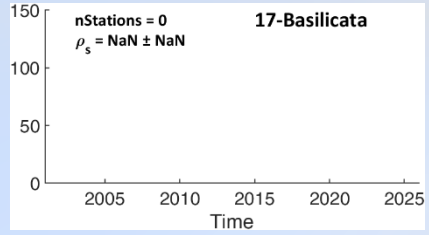


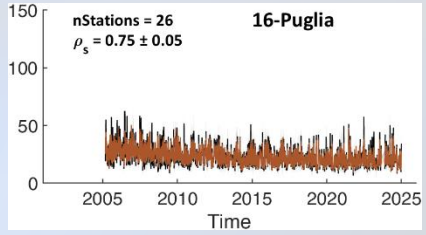




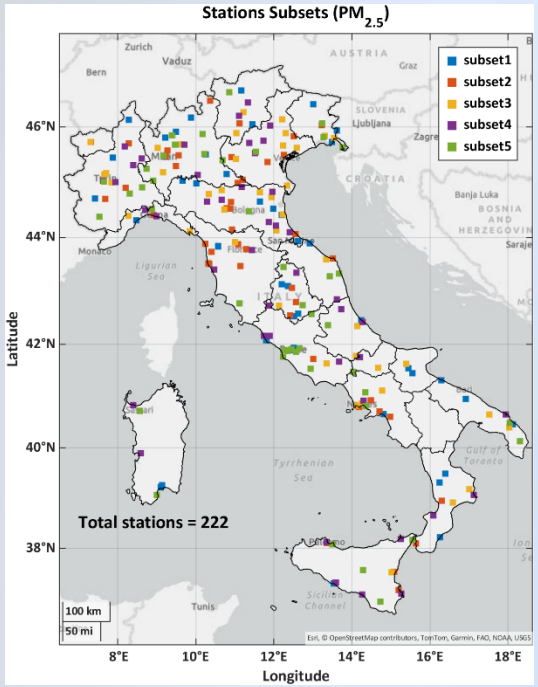


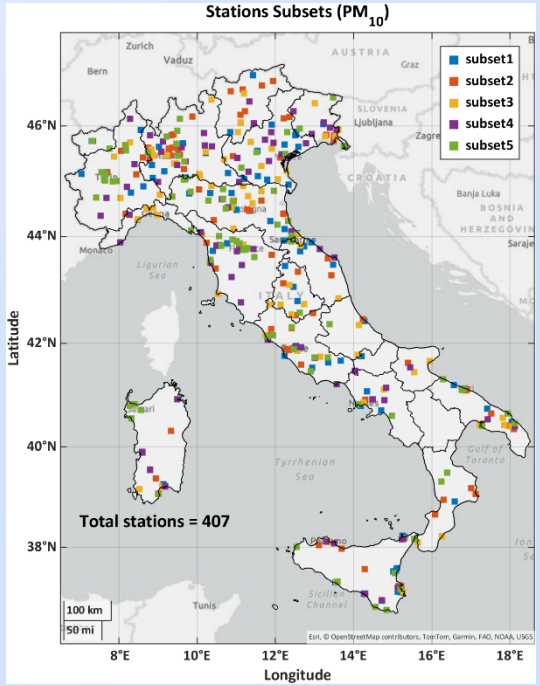


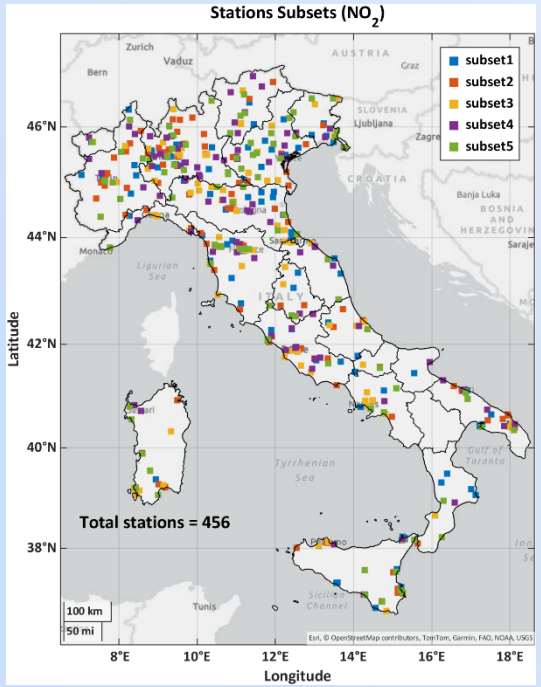


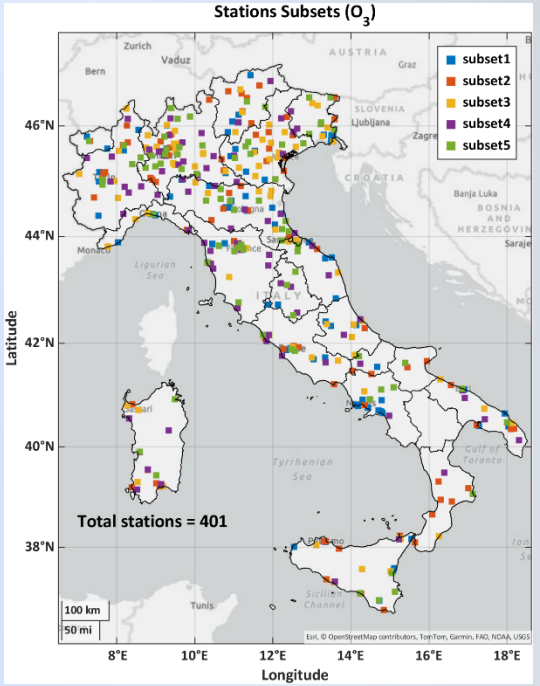


Results: performance of EBT (Space CV)





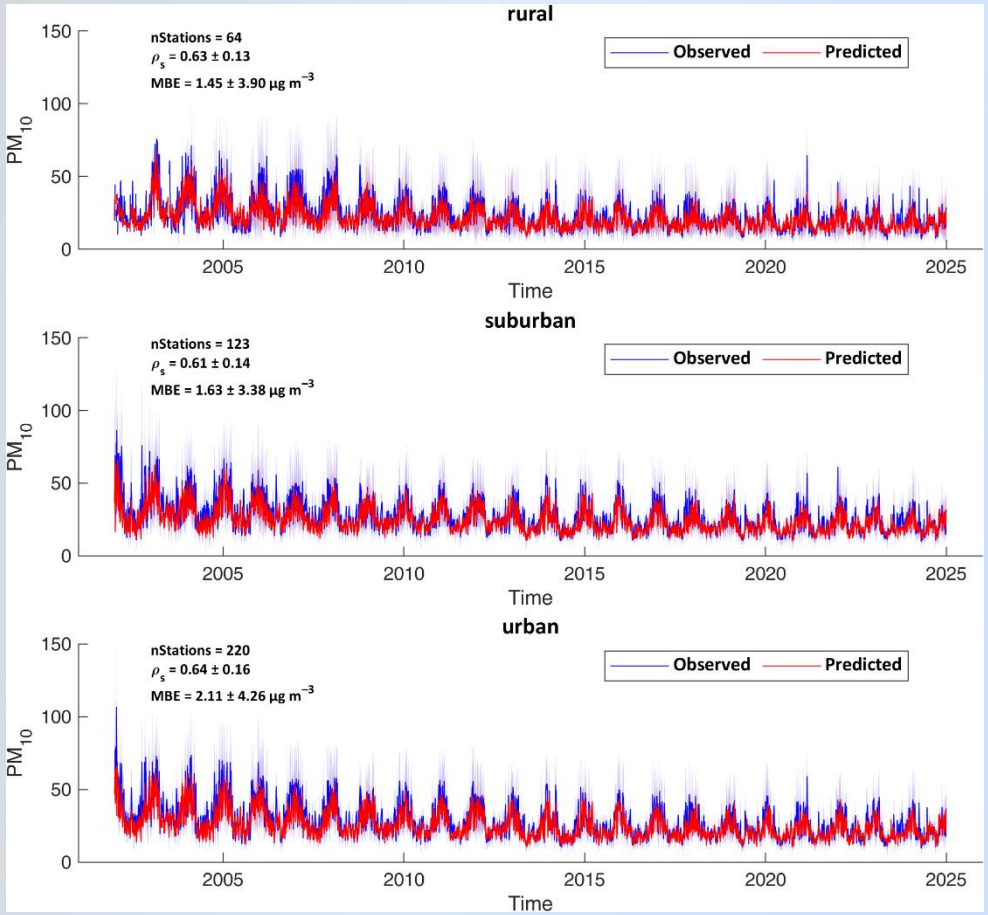




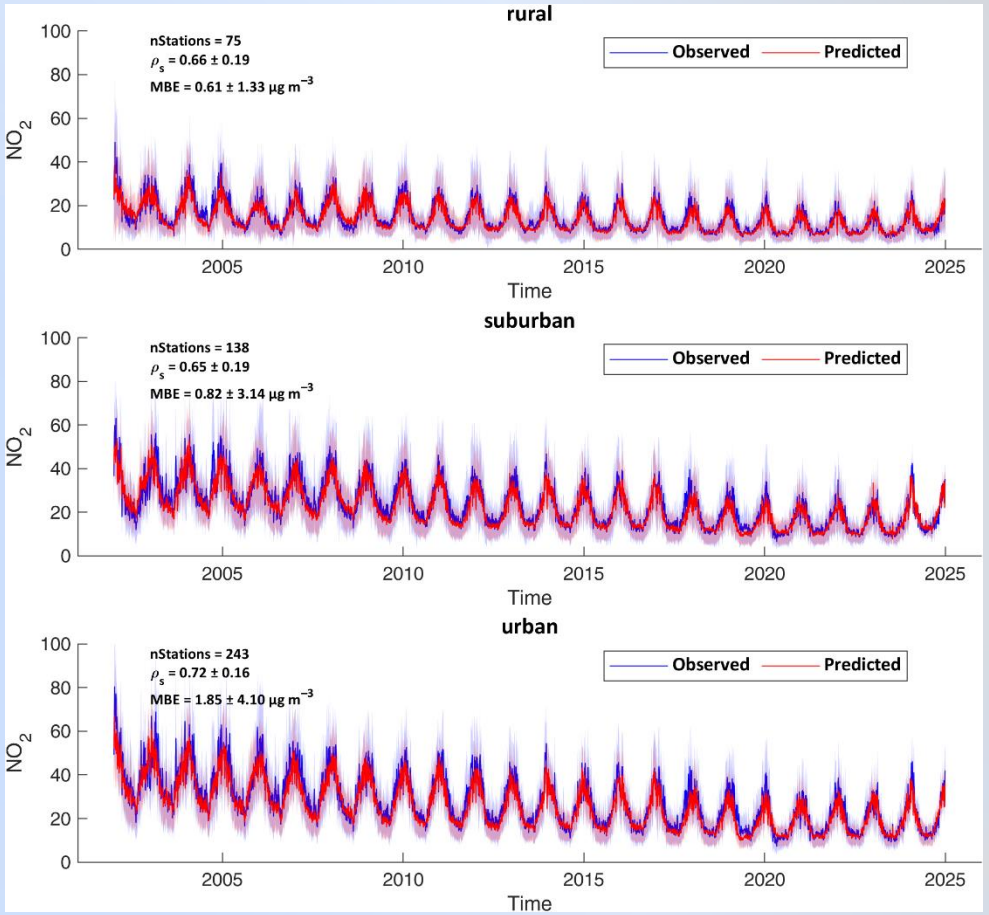
$PM_{2.5}$		PM_{10}		NO_2		O_3	
MAE	ρ_s	MAE	ρ_s	MAE	ρ_s	MAE	ρ_s
4.52 ± 0.21	0.76 ± 0.04	5.95 ± 0.36	0.77 ± 0.02	8.13 ± 0.22	0.68 ± 0.02	11.05 ± 0.75	0.84 ± 0.02

When using space CV, evaluated on unseen stations, the EBT still performs well. These values confirm the model's ability to generalize beyond the measuring sites.

Results: performance of EBT (Time CV)



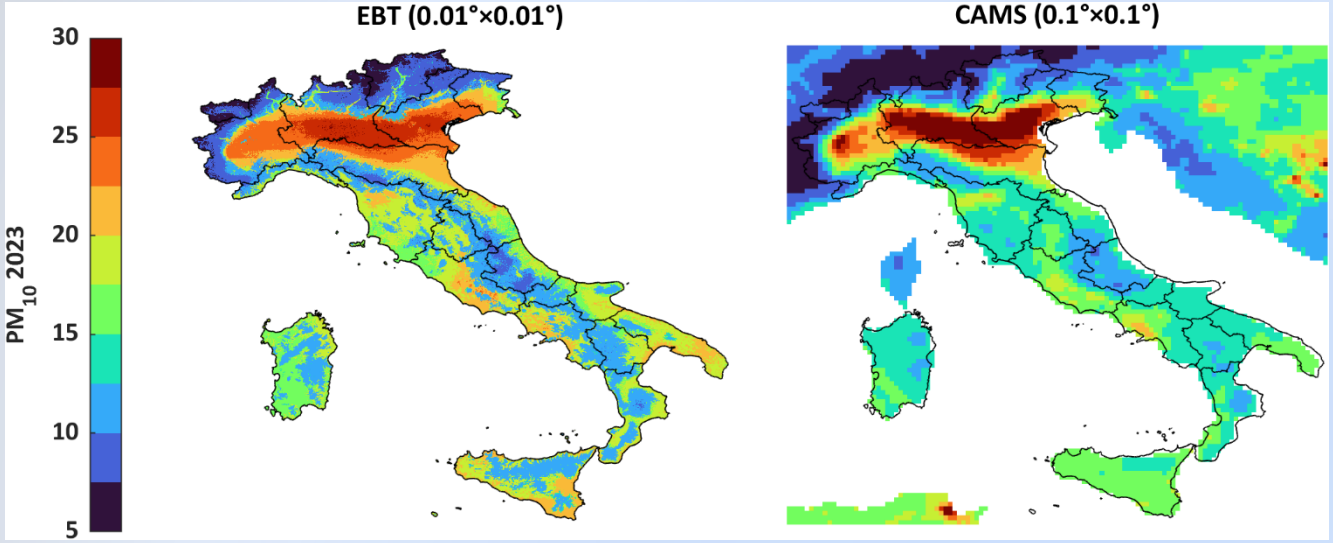
Why Important!



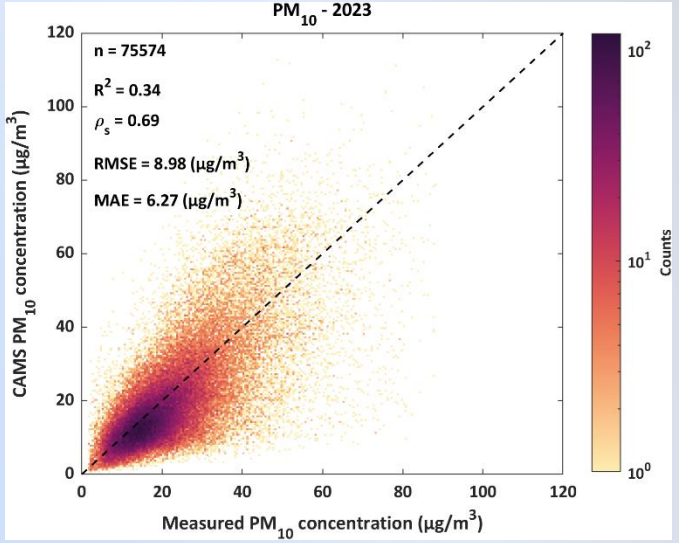
Comparison of predicted and measured daily air pollutants based on EBT model. The validation subsets are the left-out years not included in the model.

Importantly, the model demonstrates robust performance across different spatial clusters (urban, suburban, rural).
Performance based on station-wise in different types.

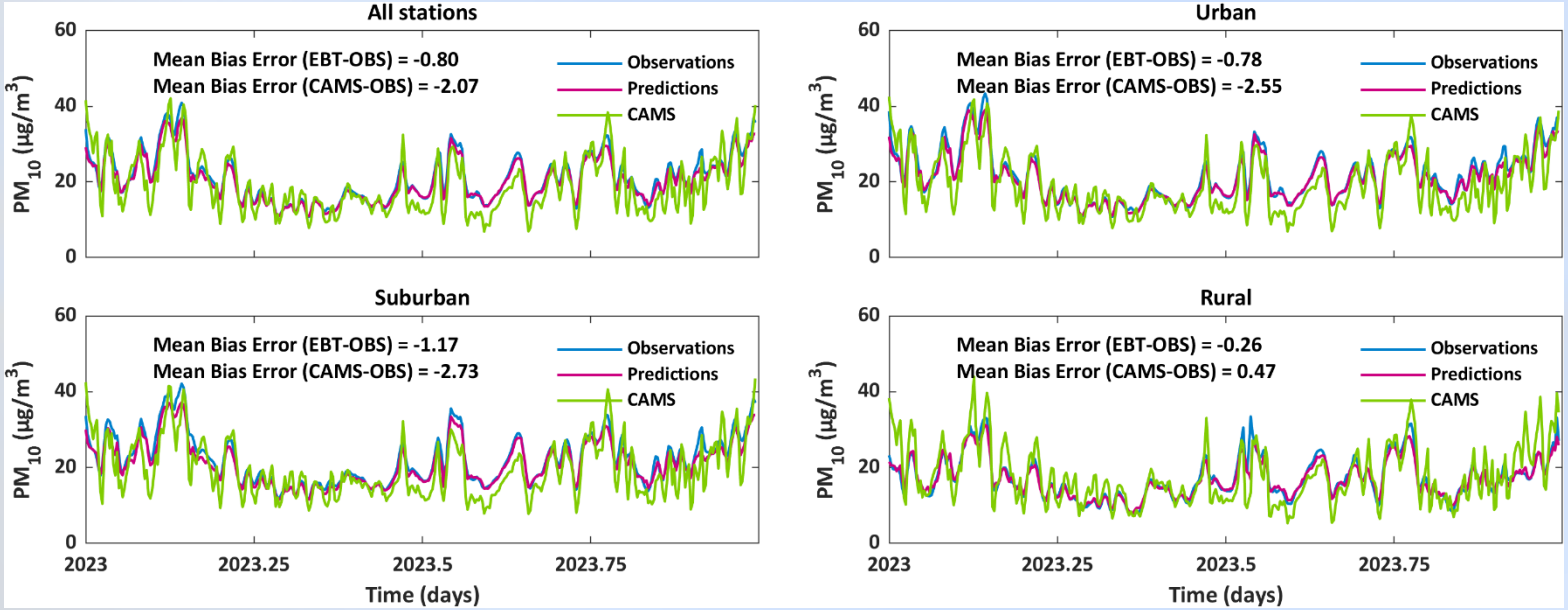
Comparison with CAMS European air quality analysis (PM_{10})



The 1km resolution outputs from EBT offer significantly improved spatial detail.

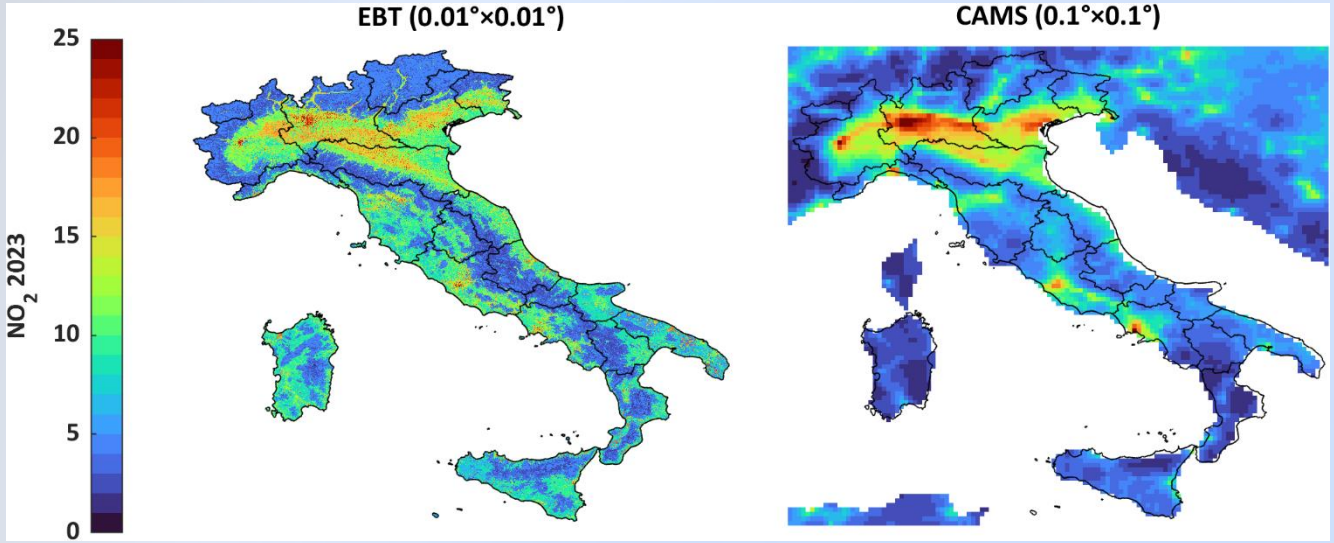


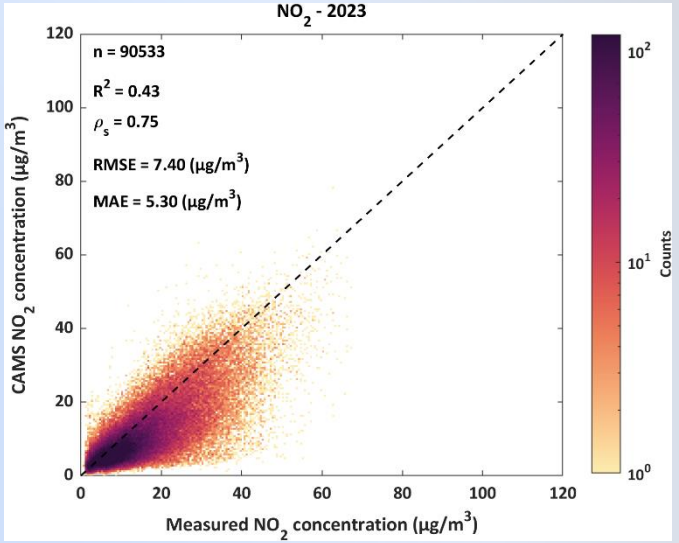
EBT explains a greater percentage (82%) of the observed variance.

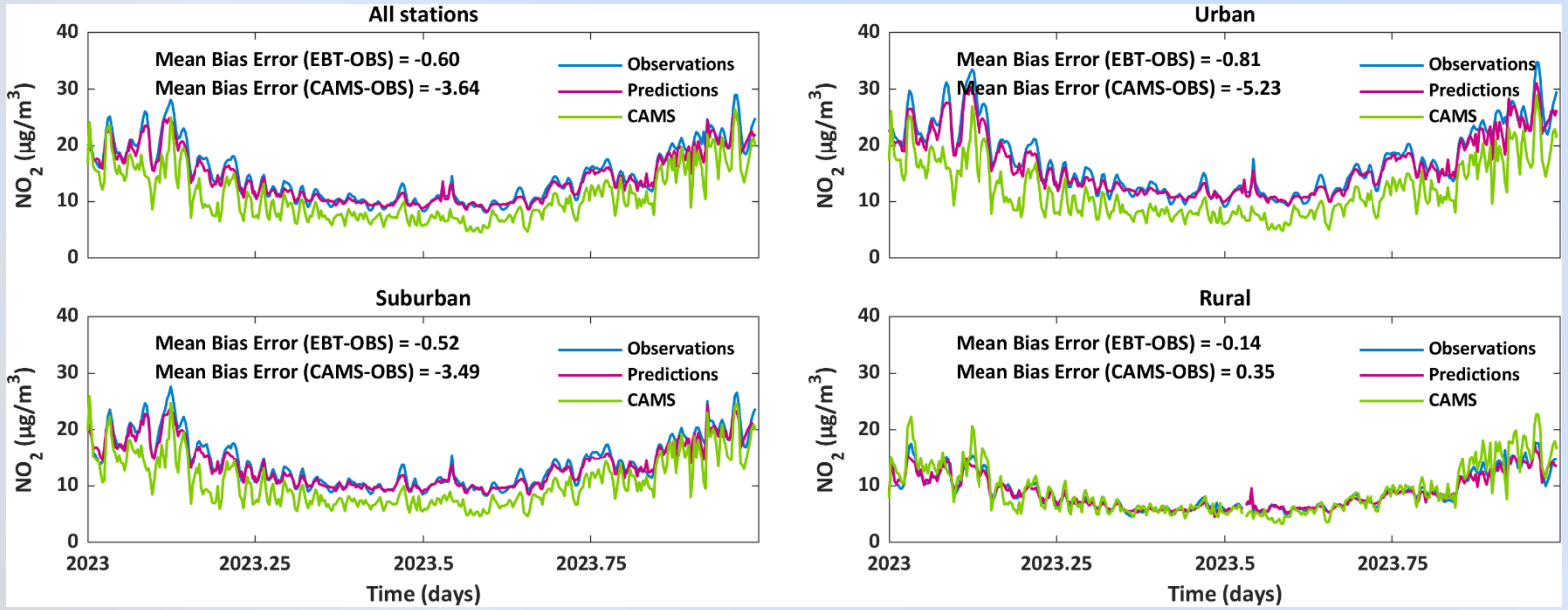


And in terms of time series, EBT has a lower biased error.

Comparison with CAMS European air quality analysis (NO_2)

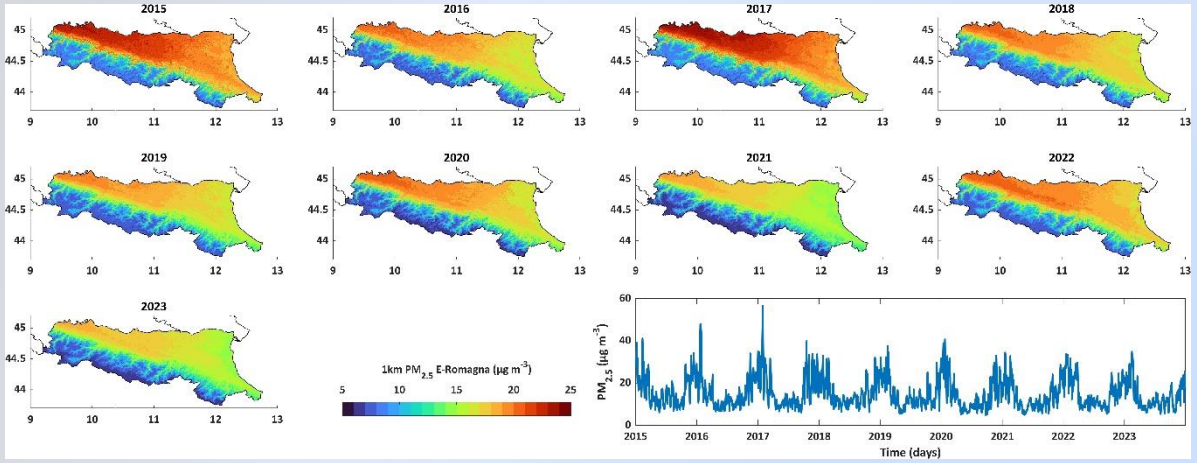


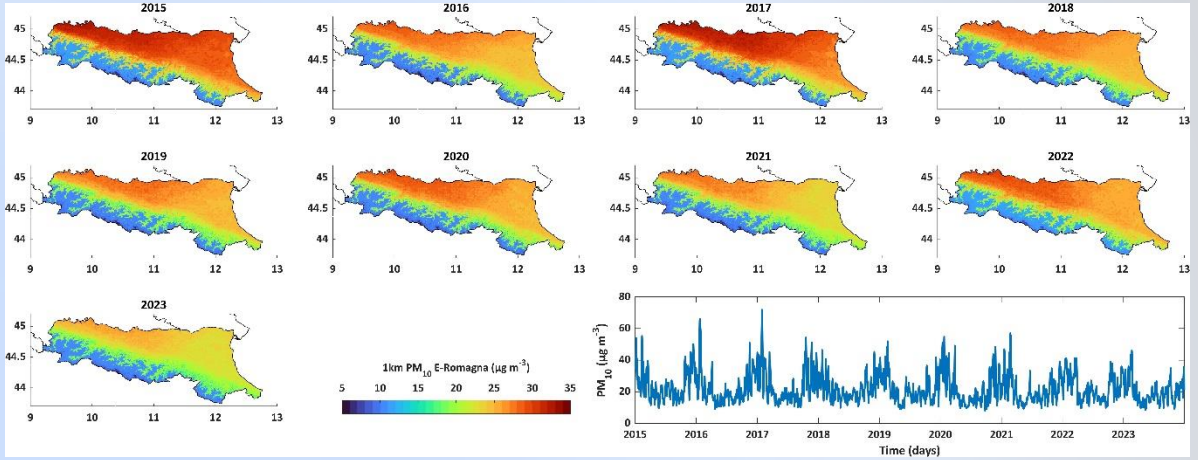


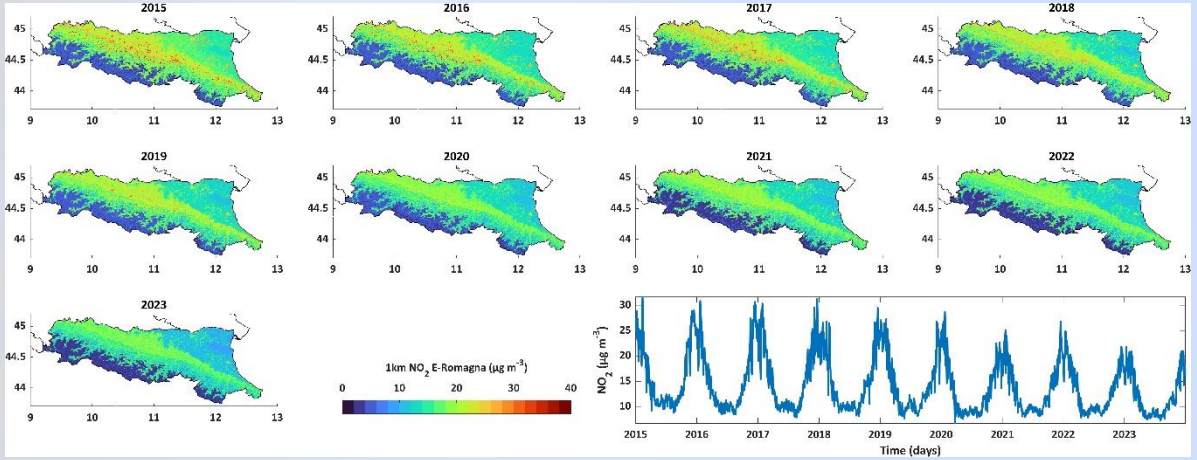


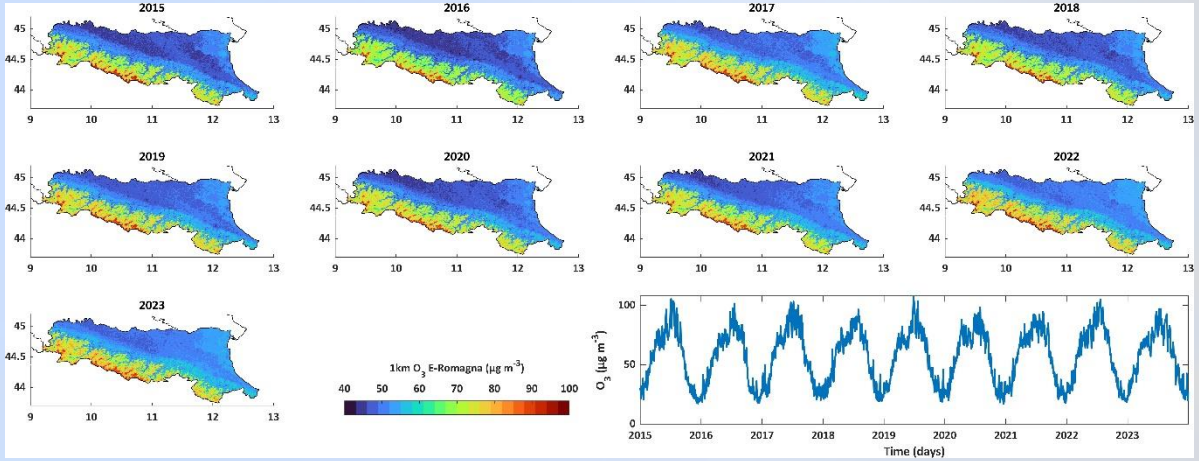
CAMS shows NO_2 underestimated concentrations.

Examples of the 1km Data Product – Emillia Romagna



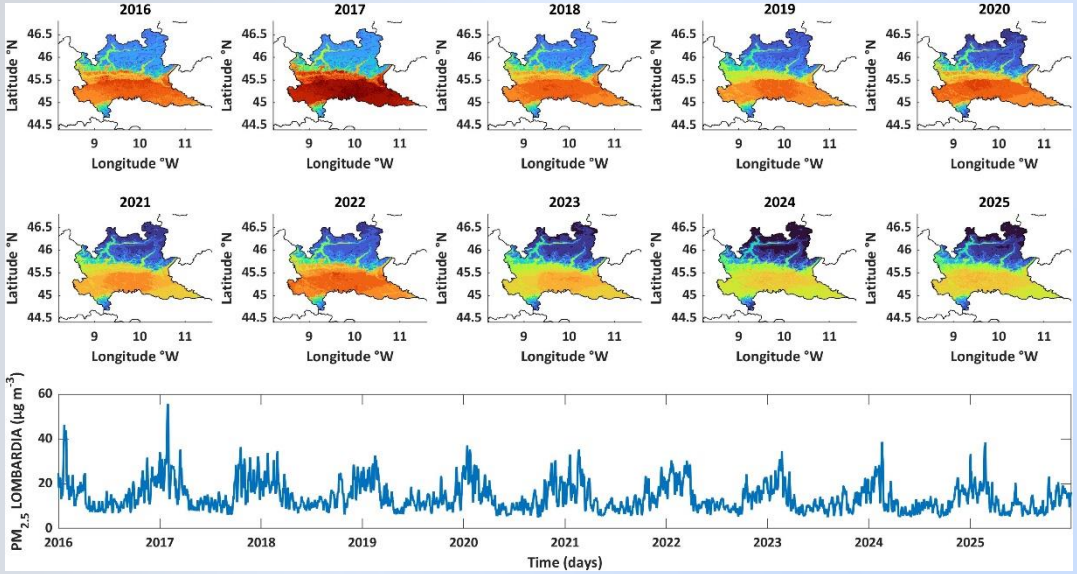


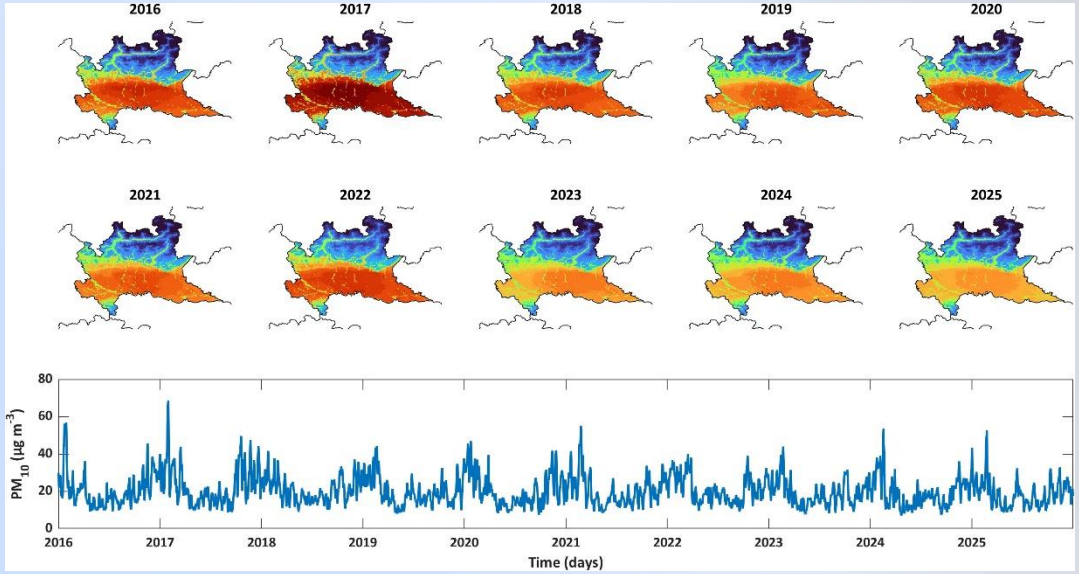


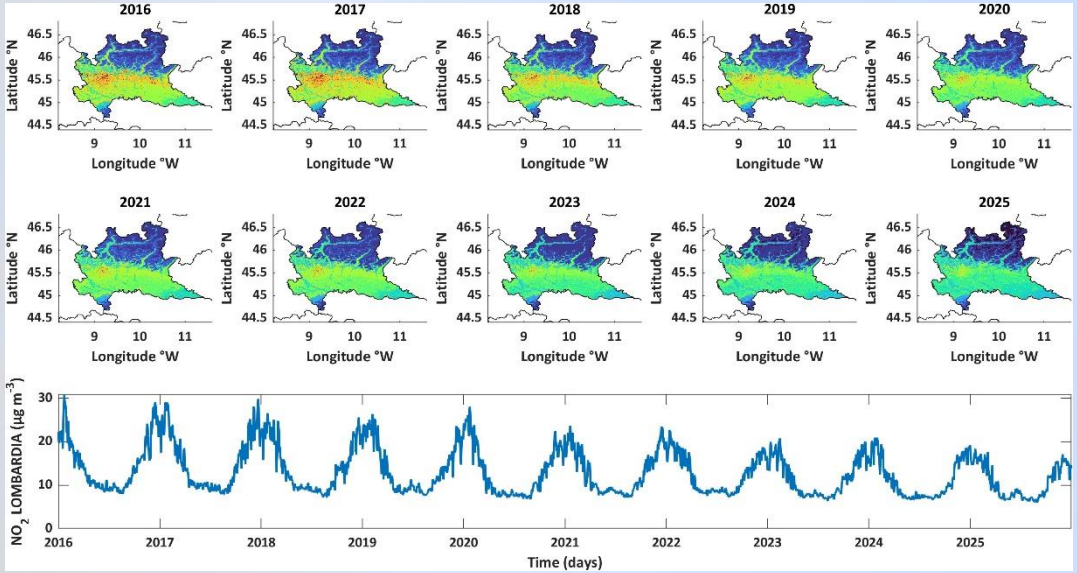


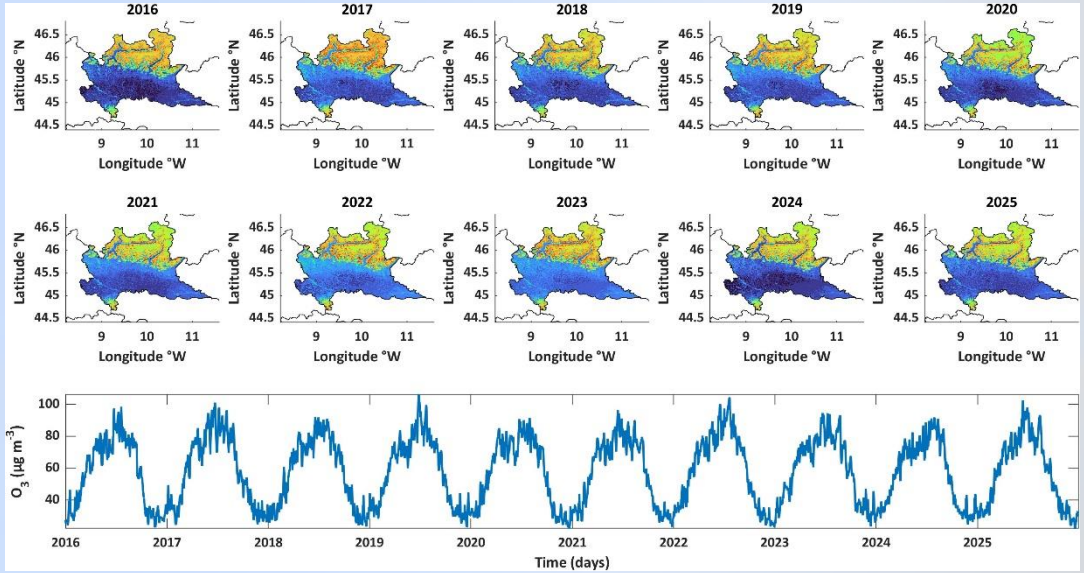
The maps present the annual mean of each pollutant concentration predicted by EBT at a daily – 1 km resolution

Examples of the 1km Data Product – Lombardia



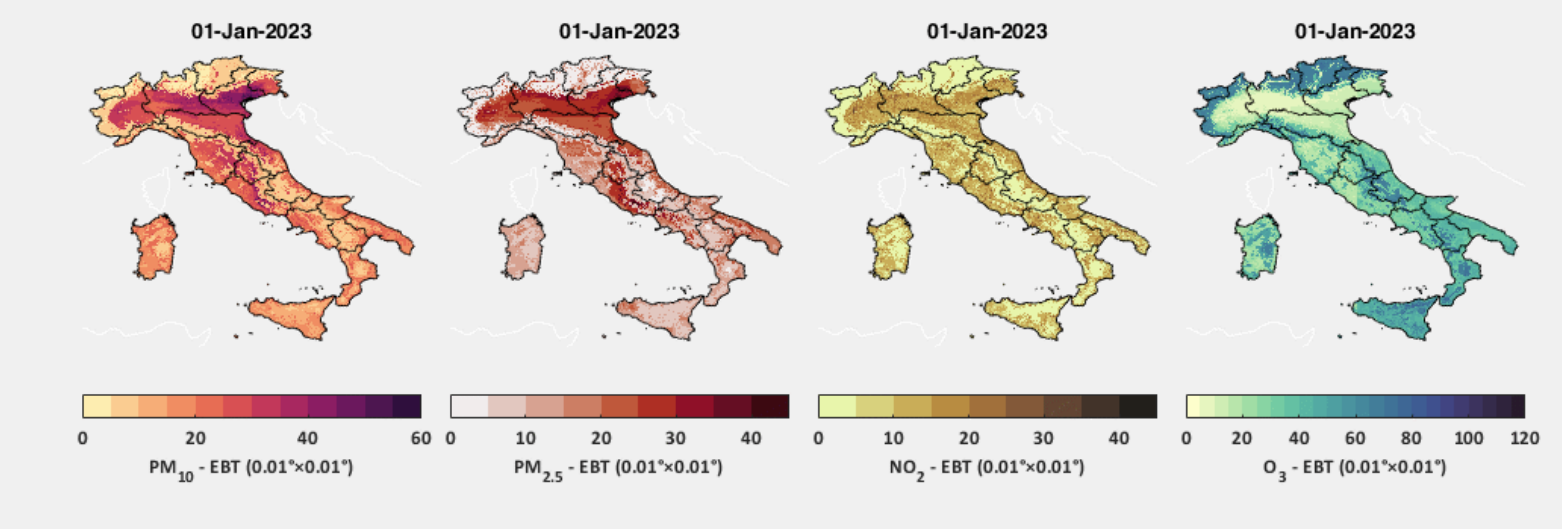






The maps present the annual mean of each pollutant concentration predicted by EBT at a daily – 1 km resolution

Summary



- We developed daily, long term, 1 km maps of $PM_{2.5}$, PM_{10}, NO_2, and O_3 over Italy from 2002-ongoing.
- The EBT model performs strongly across random, spatial, and temporal CV and across different Regions/Types.
- Our product outperforms CAMS in both resolution and accuracy.
- These maps provide a unique resource for public health, urban planning, and climate air quality co benefit studies.

Consiglio Nazionale delle Ricerche — Istituto di Scienze dell'Atmosfera e del Clima

The Institute of Atmospheric Sciences and Climate

Promotes and develops an integrated scientific understanding of the atmosphere, the ocean and their processes, through a multidisciplinary approach.



CAMEO
Climate and Meteorology, Modelling, and Earth Observation

IMPEACH
Impacts on environment, cultural heritage and human health

CAFCA
Atmospheric composition, climate forcing and air quality

Thank you for your attention!

Questions?

Correspondence: k.mansour@isac.cnr.it

Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani

Project funded under the National Recovery and Resilience Plan (NRRP), Mission 04 Component 2 Investment 1.5 – NextGenerationEU, Call for tender n. 3277 dated 30/12/2021. Award Number: 0001052 dated 23/06/2022 (ECS_00000033_ECOSISTER).