https://doi.org/10.5194/egusphere-2026-1109
Preprint. Discussion started: 6 March 2026

High-resolution mapping of air quality across Europe: an ensemble machine and deep learning framework integrating multi-scale spatial predictors (CHROMAP v1.0)

Antoine Guion[1], Alicia Gressent[1], Gaël Descombes[1], Yassine Janati[2], Elsa Real[1], Anthony Ung[1], Frédérik Meleux[1], Simone Schucht[1], Augustin Colette[1]

[1]French National Institute for Industrial Environment and Risk (Ineris), Verneuil-en-Halatte, 60550, France
[2]National Institute of Applied Sciences (INSA Rouen), Saint-Etienne-du-Rouvray, 76800, France

Correspondence to: Antoine Guion (antoine.guion@ineris.fr)

Abstract. This article presents a model for mapping air quality at high-resolution (called CHROMAP) based on the fusion of data from deterministic models, in-situ and satellite observations, and spatial proxies using an ensemble of ML and DL algorithms. Annual estimates of the SOMO35 indicator and the average concentrations of NO_2, PM2.5, PM10, and O_3 are produced and evaluated for the 2013-2023 period at a spatial resolution of 500 meters over the European domain. The methodology maintains consistency across all pollutant indicators while ensuring flexibility and transferability.

By including interpretable AI diagnostics, CHROMAP provides a quantitative assessment of the importance of the 26 features over 11 years for each air quality indicator. Integrating all types of stations into the regressions, the evaluation carried out reveals that the performance scores have been significantly improved compared to CAMS reanalyses (~10 km resolution) used for downscaling; with a reduction in RRMSE on average over the period of about -33% for NO_2, -21% for O_3, -10% for SOMO35, -22% for PM2.5 and -37% for PM10, and an increase in R^2 of 28%, 34%, 18%, 14% and 36%, respectively. In addition, a sensitivity analysis carried out on the static exposure of the population shows that significant differences can be found with values at high resolution, especially for NO_2, thus impacting the calculation of the health impact.

By ensuring sufficient availability of in-situ observations and concentration fields from CTMs for downscaling, this methodology could be extended to additional air quality indicators and applied at higher temporal frequency, opening new opportunities for comprehensive air quality assessment.

1 Introduction

The exposure to fine particulate matter (PM2.5), ozone (O_3) and nitrogen dioxide (NO_2) exceeding World Health Organization recommendations is estimated to be responsible for around 239,000, 70,000 and 48,000 premature deaths respectively, in Europe in 2022. Despite improvements in air quality over the last decades thanks to reductions in air pollutant emissions (the number of deaths in the European Union attributable to PM2.5 fell by 45% between 2005 and 2022), ambient air pollution remains the largest environmental health risk in Europe (European Environment Agency, 2024). In such context, accurate mapping of surface concentrations of regulated pollutants is essential to inform public authorities and to assess population exposure. Pollutant concentration fields are fundamental input data for assessing the burden of disease in epidemiological studies and, the health benefits that can be attained through additional air quality policies (Juginović et al., 2021; Orru et al., 2025; Sigsgaard & Hoffmann, 2024; World Health Organization, 2016).

Over recent decades, considerable progress has been made in air quality mapping, both in terms of the accuracy of concentration estimates and the level of resolution achieved, thanks to several factors: the improvements in modeling the physical and chemical processes of the atmosphere, the increase in the number of measurements, the advances in high-performance computing methods, and the development of geostatistical approaches including the use of artificial intelligence (AI) (e.g. Hamm et al., 2015; Kopczewska, 2022; Shen et al., 2022). However, spatio-temporal fields of pollutants remain complex and difficult to predict. Methodological developments such as those proposed in this article aim to reduce errors in spatial concentration predictions and their associated uncertainties.

There are many methods for mapping air quality at a fine scale (a few km or less) over a regional domain. A general distinction can be made between the process-based, geostatistical, and hybrid approaches. The determinist chemistry-transport models (CTMs) provide a comprehensive estimate of pollutant concentration fields by representing the physical and chemical processes at each grid cell of the domain, thus providing comprehensive coverage for a large number of species (e.g. Schaap et al., 2015). However, running CTMs at fine scale (or by downscaling using nested domains) requires significant computational resources in time and number of processor units. In addition, CTMs may be limited by the spatial resolution of their input data (e.g. meteorological data and emissions). Among the various statistical approaches, kriging is widely used for air quality mapping (e.g. Beauchamp et al., 2017). This is a method of spatial interpolation of measurement points that aims to minimise spatial variance. Kriging estimates are fundamentally constrained by the spatial distribution and reliability of available observations. In the absence of auxiliary variables that can serve as external drift (as in universal kriging), the interpolation process generates smooth spatial surfaces with attenuated gradients which may inadequately represent the sharp concentration variations and hotspots that characterize pollutants like NO_2 or PM10. Krigging can be also combined with CTM results (e.g. Real et al., 2022). Another important type of statistical models for mapping air quality are the land use regression (LUR) models (e.g. Ryan & LeMasters, 2007). They rely on a statistical relationship between air pollutant concentrations and, geographical features such as land use classes or distance to roads, initially through (multi-) linear regressions. Classical LUR models may be fast to run, but these purely statistical models do not consider input data related to the transformation and transport of chemical species, which can introduce a real bias in the estimation process. Finally, there is a growing trend toward hybrid methodologies in the state of the art, which combine various models or (geo)statistical approaches either simultaneously or in post-processing for downscaling, and generally with the use of AI. The uEMEP downscaling approach is based, for instance, on classical Gaussian plume modeling integrated into the EMEP MSC-W model and online calculation of the local emission fraction based on appropriate proxies to avoid double counting of emissions (Denby et al., 2020; Mu et al., 2022).

https://doi.org/10.5194/egusphere-2026-1109
Preprint. Discussion started: 6 March 2026

Machine learning (ML) and deep learning (DL) are used for various purposes in air quality modeling: forecast (e.g. Bertrand et al., 2023), source apportionment and scenarios (e.g. Colette et al., 2022), and mapping (e.g. Kopczewska, 2022). ML and DL algorithms are well suited for fast processing of large and diverse datasets, capable of capturing complex nonlinear relationships between variables. From these developments in AI, advanced LUR-type models have been developed, including not only land cover, but a whole series of predictors such as weather conditions, demographic characteristics or other model outputs. Based on around twenty predictors, Hough et al. (2021) trained three ML models (Linear Mixed Model, Gaussian Markov Random Field and Random Forest) to predict PM2.5 and PM10 concentrations with a spatial resolution of 1 km across France. Gouldsbrough et al. (2024) combined the concentration fields of the EMEP4UK model with weather and distance variables (e.g. distance to roads or coast) in a gradient-boosting tree model to downscale the EMEP4UK O3 maps from a resolution of 5 km × 5 km to 1 km × 1 km over the United Kingdom. Similar approaches can also be applied over regions covering several countries in Europe (e.g. Shen et al., 2022). The Regression-Interpolation-Merging Mapping (RIMM) method (Horálek et al., 2021) consists of kriging the residuals of a linear regression for different types of areas (rural, urban background and urban traffic) which are combined in a second step, reaching a spatial resolution of 1 km for annual concentrations over EEA member and cooperating countries. Finally, artificial neural networks (NNs), such as convolutional NNs or Autoencoders, are increasingly present in scientific literature used as downscaling method (e.g. Bessagnet et al., 2021; Baño-Medina et al., 2021). Trained on high and low resolution CTM outputs, the NNs process concentration fields as matrices on which they apply a series of transformations. However, NNs require a large training set for satisfactory performance which can be costly if the target domain is Europe.

While the literature presents numerous AI-based methods for high-resolution mapping, these approaches are typically developed and evaluated for specific countries, regions or individual pollutants. In addition, the multi-scale representation of pollutant concentration patterns, from background to hotspots levels with associated validation, can be limited. Further assessments of the uncertainty in predictions and interpretation of model outputs, especially for DL models, would be beneficial.

In this context, a consistent and transferable method using an ensemble of ML/DL algorithms has been developed to map air quality at high spatial resolution (~500 m x 500 m) in Europe. The targeted indicators are the annual average concentration of NO_2, O_3, PM2.5 and PM10 and the SOMO35 (yearly sum of the maximum daily 8 hour running mean concentrations greater than 35 parts per billion (ppb)), and the period covered is 2013-2023. This mapping model based on a hybrid approach, called CHROMAP, benefits from the synergy between observations, dynamic models and spatial predictors at multi-resolution using statistical fusion by AI. Moreover, it includes interpretable AI diagnostics and an estimation of the uncertainty based on the variability of the ensemble.

The purpose of this paper is to describe and evaluate CHROMAP. It is divided into 5 sections. Following this introduction (Section 1), Section 2 details the materials and methods underlying the CHROMAP model, including the presentation of predictor variables, observation data, and the ensemble of ML/DL models. Section 3 presents the results in four parts: mappings and temporal evolution assessment (3.1), feature importance analysis (3.2), prediction performance evaluation

https://doi.org/10.5194/egusphere-2026-1109
Preprint. Discussion started: 6 March 2026

compared to Copernicus Atmospheric Monitoring Service (CAMS) reanalysis (3.3) and sensitivity analysis of population exposure estimates to CHROMAP outputs (3.4). Section 5 provides a comprehensive discussion of the findings while Section 6 concludes with key insights and perspectives.

2 Description of the CHROMAP model

2.1 General design

CHROMAP (Computational model for High-ResOlution air quality MAPs based on data fusion using artificial intelligence) is as a model that maps air quality indicators at high spatial resolution based on the fusion of data from deterministic models, in-situ and satellite observations, and spatial proxies. It consists of an ensemble of supervised regression ML and DL algorithms that establish statistical relationships between a comprehensive set of explanatory variables (features) and pollutant concentrations or air quality indicators as target variables. Since CHROMAP relies primarily on a statistical approach using data fusion, and not on a process-based approach, the use of spatial data on atmospheric chemistry and physics as predictors (hereinafter referred to as “chemical feature,” see Sect. 2.2), such as concentration fields estimated by models or satellite observations, are fundamental inputs to the framework for consistent prediction of air quality indicators across the domain of the study. A specific section is dedicated in the article to the importance of each feature (see Sect. 3.2).

CHROMAP is evaluated in this article based on the annual average concentrations of NO_2, PM2.5, PM10 and O_3, and on the annual indicator SOMO35. Given that concentration fields from the Ensemble of CAMS regional reanalyses are used as “chemical feature” for each species mentioned above, the temporal and spatial coverage of the evaluation corresponds to that of the CAMS reanalyses. The period of the study is 2013-2023 and the boundaries of the European domain are 25.0°W, 45.0°E, 30.0°N and 72.0°N. CHROMAP’s grid resolution is set to 0.0064° x 0.0064° (~500 m at 45.5°N), which represents a grid of 10,934 x 6,560 (i.e., 71,727,040 grid cells). It depends on the initial resolution of its features used as proxies for air quality indicators. CHROMAP can therefore be described as a downscaling tool; from CAMS reanalyses at a horizontal resolution of 0.1° (~10 km) to spatial estimates produced at ~500 m. CHROMAP outputs are compared to CAMS reanalyses (see Sect. 3.3) to assess the gain in resolution and accuracy.

Figure 1 shows a schematic diagram of the processing steps in CHROMAP, organized as follows:

- The first step is to select the predictor variables based on scientific literature to ensure that they have a proven causal effect on the distribution of pollutants. The 26 features present in CHROMAP are detailed in Section 2.2.
- Step 2 involves feature preparation and transformation to align with the target grid. All predictor variables must be processed to match the resolution and spatial coverage of the CHROMAP grid (~500m) for integration as features into the ML and DL models. This processing involves several key operations: when the initial resolution is finer than the CHROMAP grid, data are spatially aggregated. Conversely, when the spatial resolution is coarser, the fields are interpolated using a bilinear approach. Certain features require transformation (e.g. the road length density in each grid cell based on the road network in vector data type) and other features are calculated internally (e.g. the

distance from the coast using a geographical information system). Finally, to prevent scale-induced bias in model training, features are normalized (e.g. standardization) using methods selected based on their statistical distribution (e.g. presence or not of outliers).

- Step 3 is the extraction of the features at observation points. The data is divided into training, validation, and test sets. Detailed information on in-situ observations is provided in Section 2.3.
- Each ML/DL model is trained in step 4 using the surface pollutant concentration from in-situ observations as the target. Then an Ensemble is computed as the median value of four ML/DL models that are detailed in Section 2.4. This approach aims to improve the average performance and robustness of predictions in case of individual model failure. The model tuning (hyper-parameterization) is performed using stratified K-Fold cross-validation. The data set is divided into 5 folds, 4 for training and 1 for validation, by stratified sampling according to the type of station (background, industrial and traffic) to have the most representative training set for validation. The distribution of training/validation folds is alternated 5 times until a complete prediction has been validated. The different hyper-parameters are found by the random search technique, from a list of predefined parameters. This approach is more effective for hyper-parameter optimization than trials on a grid, especially when using more complex models (Bergstra & Bengio, 2012). The hyperparameters used are described in the Supplement Material (SM) (Table S1).

  In addition, as part of an interpretable machine learning approach (e.g. Molnar et al., 2020), two diagnostics are performed to better understand the decisions of the models. First, the permutation feature importance (e.g. Kaneko, 2022) which consists of randomly shuffling the values of individual features one after another and evaluating the possible degradation of the model score. Second, the Shapley Additive exPlanations (SHAP) framework (Lundberg & Lee, 2017) which allows to evaluate the contribution of each feature to the prediction.

  Finally, the performance scores are calculated with an independent test set for a chosen year which is 2021 in this study, for each air quality indicator (see Sect. 3.3).
- Step 5 is the inference over all grid cells to produce maps. The maps are assessed to identify potential problems of inconsistent values or discontinuity issues. This may result in adjustments to the feature selection and preparation steps. Finally, standard deviation maps are produced to give an estimate of the uncertainty, related to the deviation between the models of the Ensemble over specific regions.

These steps are repeated for each year. As a result, CHROMAP does not need to integrate a time dimension into its predictions, as it is intrinsically integrated with dynamic predictors and observations. Furthermore, the results of Shen et al. (2022) showed that multiple-year models do not show better performance scores than single-year models. Temporal consistency will be evaluated regarding observations (Sect. 3.1).

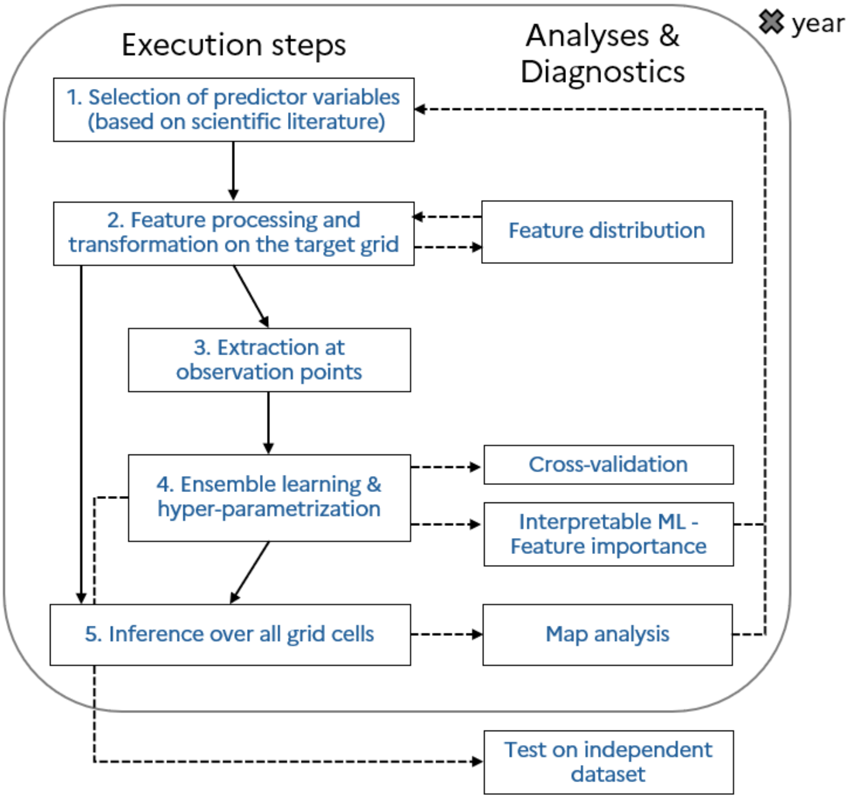


Figure 1: Schematic diagram of the CHROMAP workflow applied for each air quality indicators.

2.2 Spatiotemporal predictor variables

2.2.1 Chemical features

Several predictor variables have been identified to provide optimal estimates for air quality mapping. They are processed to serve as features in CHROMAP for use as explanatory variables in the ensemble of supervised regression algorithms. The framework incorporates 26 features in total: 6 chemical, 7 meteorological and 13 geographic. Table 1 presents these 26 features, including their initial spatial resolution and temporal coverage. The spatial extent of all features encompasses the entire European domain as defined in Section 2.1. All features used in CHROMAP come from open-access databases.

As mentioned in the Introduction (see Sect. 1), incorporating CTM outputs into data fusion enhances both prediction accuracy and spatial consistency. In this study, European air quality reanalysis from CAMS are used (Colette et al., 2025; Peuch et al., 2022). CAMS is the core global and regional atmospheric environmental service provided as part of the atmosphere component of the European Copernicus program. The regional production system integrates 11 CTMs from

https://doi.org/10.5194/egusphere-2026-1109
Preprint. Discussion started: 6 March 2026

various institutes and countries, providing daily analyses and forecasts of regional atmospheric composition in near-real-time, alongside comprehensive reanalysis datasets for past years. Reanalysis data are produced by running simulations that replay specific years using an observation assimilation system. These simulations incorporate either the most recent observations for "interim" reanalysis or validated observations for "validated" reanalysis (generally available with a 2-year delay). In both cases, background stations are used as they are representative of CAMS model resolution.

The ensemble of the reanalysis (median of the 11 CTMs calculated at each grid cell) was used for two periods: 2013-2021 (validated reanalysis) and 2022-2023 (interim reanalysis). Annual averages of NO_2, PM2.5, PM10 and O_3, and the SOMO35 indicator were calculated from hourly concentration fields to provide the "CHEM_CTM_no2", "CHEM_CTM_o3", "CHEM_CTM_pm2_5", "CHEM_CTM_pm10", and "CHEM_CTM_somo35" features, respectively.

In addition to CTM outputs, satellite observations of NO_2 columns are included in the list of features, as this product tends to improve NO_2 predictions (Horálek et al., 2021; Shen et al., 2022). The "CHEM_Sat_no2" feature represents the total column of NO_2 in the atmosphere. It is based on data from the TROPOspheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor (S5P) satellite in low orbit, as part of the European Union's Copernicus Program, which aims to monitor the Earth's atmosphere at high resolution. Since April 2018, the S5P sun-synchronous orbit has been measuring the tropospheric vertical column densities of NO_2 for daily global coverage with a resolution of approximatively 5.5 km x 3.5 km (7.5 km x 3.5 km between April 2018 and August 2019). The evolution of products and algorithms, as well as the auxiliary data used, are constantly referenced in the Algorithm Theoretical Basis Document (van Geffen et al., 2024). Reprocessed data (version 2.4) are used to ensure consistent NO_2 product quality throughout the study period. In addition, raw NO_2 columns are filtered using a quality assurance value greater than 0.75 as recommended in the ATBD. Finally, the "CHEM_Sat_no2" feature is used from 2019, the first complete data available, until 2023. This feature also contributes to O_3 concentration predictions as NO_2 acts as both a precursor and scavenger of O_3, influencing O_3 production and spatial distribution patterns (Jenkin & Clemitshaw, 2000).

2.2.2 Meteorological features

Weather conditions play a key role in the distribution of pollutant emissions (Grange et al., 2019; Guion et al., 2025; Yang et al., 2019), but also in their concentration in the atmosphere through chemical transformation and deposition mechanisms (e.g. Guion et al., 2023; Kavassalis & Murphy, 2017; Makar et al., 2015). A set of meteorological features has been implemented in CHROMAP based on two datasets: the fifth generation ECMWF atmospheric reanalysis of the global climate and weather (ERA5, Hersbach et al. (2023)) and the Copernicus European Regional ReAnalysis (CERRA, Schimanke et al. (2021)). ERA5 provides hourly estimates for many atmospheric variables based on physics-based model that assimilates observation data. The global grid has a spatial resolution of approximately 27.8 x 27.8 km. CERRA is a high-resolution regional dataset for Europe that employs ERA5 for lateral boundary conditions (serving as prior estimates of the atmospheric conditions) and assimilates observations to enable reconstruction at a finer grid resolution (~5.5 x 5.5 km).

This enhanced spatial resolution provides, among other benefits, improved representation of topography and physiographic features.

The list of meteorological features used in CHROMAP includes wind speed at 10 meters (“MET_Winds”), temperature at 2 meters (“MET_Temp”), surface pressure (“MET_SurfP”), and total cloud cover (“MET_CloudC”) from CERRA, supplemented by total precipitation (“MET_Precip”), short-wave radiation downwards (“MET_Rad”) and the boundary layer height (“MET_Pbl”) from ERA5. All meteorological variables, from both ERA5 and CERRA, were annually averaged and regridded to the CHROMAP domain at a 500 m resolution using bilinear interpolation.

2.2.3 Geographical features

Geographical features are fundamental for high-resolution mapping (e.g. Gouldsbrough et al., 2024; Hough et al., 2021; Shaddick et al., 2020), especially for LUR-based approaches like CHROMAP. High spatial resolution data on population density, land cover, altitude, and other geographic variables, enable effective downscaling applications.

Land cover features are derived from the Global Land Cover product (version 3) of the Copernicus Land Service component (Buchhorn et al., 2020). This global dataset (~100 m x 100 m spatial resolution), provides spatial information on different classes of physical coverage of the Earth’s surface. Based on long-term satellite observations (PROBA-V and Sentinel-2 satellites), the land cover classification demonstrates an overall accuracy of about 80%, validated against 28,000 independent reference points. The fraction layers of crops that are considered in this study, in particular forest/tree, grass, urban/built-up, shrub and bare soil, have been aggregated on the CHROMAP grid to provide the “GEO_Crop”, “GEO_Tree”, “GEO_Grass”, “GEO_Builtup”, “GEO_Shrub”, “GEO_Bare”, features respectively.

Predictor variables related to anthropogenic emission sources are also important for predicting the spatial distribution of pollutant concentration fields at high spatial resolution. CHROMAP therefore includes features such as population density (“GEO_PopDens”) and degree of urbanization (“GEO_UrbDeg”) derived from the Global Human Settlement product (GHS, (Pesaresi et al., 2024)). The GHS population spatial raster dataset (GHS-POP R2023A) describes the number of people per grid cells of 100 m x 100 m spatial resolution worldwide. It redistributes data from the population census by administrative entity, harmonized by the Center for Integrated Earth System Information (CIESIN), such as gridded data with information on buildings observed by satellite (Sentinel-2 and Landsat). The population density is aggregated at the CHROMAP grid level. The GHS Settlement Model layers (GHS-SMOD R2023A product) describe different degrees of urbanisation. The total population, cluster size of population and built-up area densities are taken as classification criteria to define the following settlement typologies: urban centre, dense urban cluster, semi-dense urban cluster, sub-urban, rural cluster, low density rural and, very low-density rural grid. From a spatial resolution of 1 km, GHS-SMOD was interpolated at 500 m by the method of the nearest neighbour. The years 2015 and 2020 covered by GHS-POP and GHS-SMOD are used, applying temporal interpolation based on the nearest neighbour method for the missing years from 2013 to 2023.

The road network is an indicator of the pollutant emissions related to the traffic sector, especially nitrogen oxides (NO_x) and PM. It is commonly used as feature in ML/DL models (e.g. Yu & Liu, 2021). The “Geo_RoadNet” feature represents the total road length in each grid cell of CHROMAP. It is calculated using OpenStreetMap (OSM) data (OpenStreetMap, 2024). OSM provides mapping data for thousands of websites, mobile applications and devices. The dataset is in open access and provides geographical features at global scale. Extraction of OSM data is possible for specific regions. In this study, OSM data were extracted over Europe through Geofabrik (Geofabrik Download Server). The extracted data files contain several types of elements: nodes (defining spatial points), ways (defining linear features and area boundaries), and relations (defining interactions between elements). All data elements can include a tag that describes their associated meaning. The "highway" key identifies any kind of road, street or path. Four highway categories were considered at the European level: motorways (defined as a restricted access major divided highway, normally with two or more running lanes plus emergency hard shoulder, which is equivalent to the Freeway, Autobahn, etc.), trunk roads (the most important roads in a country's system that are not motorways and are not necessarily divided highway), primary roads (the next most important roads in a country's system that often link larger towns), and secondary roads (the next most important roads in a country's system that often link towns). Using these elements, the total road length within each grid cell was computed by summing the lengths of all road segments intersecting that cell.

Finally, purely spatial variables are also included in CHROMAP's features. Longitude ("GEO_Lon"), latitude ("GEO_Lat"), and distance from the coast on the land surface ("GEO_DistCoast") have been calculated for each grid cell of the domain. Altitude is also included ("GEO_Alt") and derived from the Copernicus Digital Elevation Model product (Copernicus, 2022). This product, with an initial spatial resolution of 90 m, was calculated by "Distribution Airbus Defence and Space GmbH" based on the TanDEM-X mission.

Table 1: List of features used in CHROMAP, with European spatial coverage.

Feature ID	Description	Spatial resolution	Temporal coverage	Data source
CHEM_CTM_no2	Background concentration fields of NO_2 [µg/m^3]	~ 10.0 km x 10.0 km	2013-2023	Median of CAMS reanalysis (Colette et al., 2024)
CHEM_CTM_o3	Background concentration fields of O_3 [µg/m^3]			
CHEM_CTM_somo35	Background concentration fields of SOMO35 [µg/m^3]			
CHEM_CTM_pm2_5	Background concentration fields			

	of PM2.5 [μg/m^3]			
CHEM_CTM_pm10	Background concentration fields of PM10 [μg/m^3]			
CHEM_Sat_NO2	NO_2 total column [mol/m2]	~ 7.5 km x 5.5 km until August 2019 and ~ 3.5 km x 5.5 km thereafter until 2023	2019-2023	Copernicus Sentinel-5P (van Geffen et al., 2024)
MET_Precip	Reanalysis of total precipitation [m]	~ 27.8 km x 27.8 km	2013-2023	ERA (Hersbach et al., 2023)
MET_Rad	Reanalysis of surface short-wave (solar) radiation downwards [J/m^2]			
MET_Pbl	Reanalysis of boundary layer height [m]			
MET_Winds	High-resolution reanalysis of 10 metre wind speed [m/s]	~ 5.5 km x 5.5 km		CERRA (Schimanke et al., 2021)
MET_Temp	High-resolution reanalysis of 2 metre temperature [K]			
MET_SurfP	High-resolution reanalysis of surface pressure [Pa]			
MET_CloudC	High-resolution reanalysis of total cloud cover [ratio between 0 and 1]			
GEO_Crop	Cover fraction of crops [%]	~100 m x 100 m	2018	Global Dynamic Land Cover

https://doi.org/10.5194/egusphere-2026-1109
Preprint. Discussion started: 6 March 2026

GEO_Tree	Cover fraction of forest/tree [%]			(Buchhorn et al., 2020)
GEO_Grass	Cover fraction of grass [%]			
GEO_Builtup	Cover fraction of urban/built-up [%]			
GEO_Shrub	Cover fraction of shrub [%]			
GEO_Bare	Cover fraction of bare soil [%]			
GEO_Lat	Latitude [degree]	~500 m x 500 m	/	Calculated
GEO_Lon	Longitude [degree]			
GEO_Alt	Topographic altitude [m]	~90 m x 90 m	2022	Copernicus Digital Elevation Model (Copernicus, 2022)
GEO_DistCoast	Distance from the coast [m]	~500 m x 500 m	/	Calculated
GEO_RoadNet	The total road length [km]	Original road width (linear vector)	2024	OSM (OpenStreetMap, 2024)
GEO_PopDens	Population density [no. of inhabitants]	~100 m x 100 m	2015, 2020	GHSL (Pesaresi et al., 2024)
GEO_UrbDeg	Degree of urbanization [classes]	~1.0k m x 1.0 km		

2.3 In-situ pollutant observations

In-situ pollutant observations are from the European surface network Air Quality e-Reporting (AQ e-Reporting, (EEA, 2025)), which gathers air quality measurements from EU members states and other EEA collaborating countries through an automated quality control process. Annual averages (based on calendar years) for NO_2, O_3, PM2.5 and PM10 along with the SOMO35 indicator, were extracted from the annual air quality statistics dataset covering 2013-2023. EEA calculates directly these aggregate air quality values based on observation time series using the « E1a/validated data AQ e-Reporting » data flow source. All monitoring station types (background, industrial and traffic) and areas (rural, suburban, urban) were included in the analysis. Following quality control procedures, monitoring stations with annual concentration exceeding the 99th percentile of the distribution for each pollutant were excluded as outlier.

The surface concentration observation dataset used in CHROMAP therefore includes between 2917 and 3549 monitoring stations (depending on the year) for NO_2, 1973 and 2121 for O_3, 1962 and 2109 for SOMO35, 1163 and 2164 for PM2.5, and 2348 and 3006 for PM10. The change over time in the total number of monitoring stations and by type between 2013 and 2023 is shown in SM (see from Fig. S1 to Fig. S5). The number increased by 86% for PM2.5, 28% for PM10, 22% for NO_2, and less than 10% for SOMO35 and O_3. The spatial distribution of stations is also heterogeneous, both within and between countries. Northern and Eastern Europe are less well covered by observations for all considered pollutants (see from Fig. S1 to Fig. S5).

2.4 Ensemble of regression models

For each year and air quality indicator, an ensemble of four ML and DL regression models is trained on surface observations from the EEA. The ensemble comprises Ridge regression (RI), Random Forest (RF), XGBoost (XGB) and Multi-layer Perceptron (MLP) models.

RI is a linear regression model that includes a correction factor, known as L2 regularization (Hoerl & Kennard, 1970). By penalizing large coefficients, the regularization term prevents overfitting on training data. RI is specifically designed to address multicollinearity in the dataset. This statistical model can be particularly relevant in this study, given the numerous predictor variables that can exhibit multicollinearity.

RF is a tree-based ensemble algorithm (Breiman, 2001). RF works in three steps: bootstrap sampling, feature sampling and aggregation. First, RF randomly selects training datasets for each tree. Then, random subsets of features are used to build each tree. Finally, a prediction is performed for all the individual trees, and the results are aggregated into a single output prediction. This ensemble approach allows to produce a more accurate and stable prediction by reducing the variance that may be high on some individual decision trees. This prevents overfitting and improves the generalization of the model.

XGB is a scalable end-to-end tree boosting system (Chen & Guestrin, 206). Unlike RF, which calculates all decision trees in parallel, XGB relies on sequential tree building based on a gradient boosting approach where each new tree corrects errors made by the previous ones, minimizing the loss when adding new models. In addition, the algorithm is optimized including regularization terms (L1 and L2) and dedicated handling of missing data. XGB is known as specifically designed for speed processing and performance predictions (notably due to its ability to avoid overfitting). However, its associated structures of decision tree random forests, gradient descent and regularization involve many parameters to be tested during a tuning step (hyper-parameterization).

MLP, classified as a deep learning algorithm, is a type of artificial neural network comprising multiple layers of interconnected neurons (McCulloch & Pitts, 1943; Rumelhart et al., 1986; Werbos, 1990). MLP performs particularly well at learning complex data, especially with non-linear relationships. The MLP workflow can be described in successive layers: the first layer composed of inputs (predictive variables), the hidden intermediate layers composed of a defined number of layers and nodes (neurons), and finally the output layer (predicted variable). Data are combined in the hidden layers to create weights on which activation functions are applied, allowing for non-linear transformations of the data. Weights that are randomly set in the first step (giving a stochastic characteristic to the prediction of the algorithm), are updated by gradient descent and adjusted during backpropagation steps, aiming to minimize the cost function.

Each model undergoes supervised learning using identical features (see Sect. 2.2) (except for NO_2 and O_3 that also includes "CHEM_Sat_no2" from 2019). The implementation relies on the Python module "Scikit-learn", which provides a wide range of state-of-the-art machine learning algorithms for medium-scale supervised and unsupervised applications (Pedregosa et al., 2011).

The prediction $\hat{y}$ of each model m can be considered as a function of n variables x (defined here by the features) plus an error term ε (Eq. 1). Predictions are made for each cell of index i on the grid. The CHROMAP prediction $\hat{Y}_i$ is equal to the median of the model predictions, calculated for each grid cell (Eq. 2).

$$\hat{y}_m = f(x_n) + \varepsilon \,, \tag{1}$$

$$\hat{Y}_i = Med(\hat{y}_{m,i}) \,, \tag{2}$$

The cost function in ML/DL models used to quantify the error between predicted and expected values is based on the mean squared error (MSE) score (Eq. 3). The scores used in this article to evaluate predictions are the relative square root of MSE (i.e. the RRMSE, Eq. 4) and the coefficient of determination (R^2) (Eq. 5).

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{j=1}^{n} (y_j - \hat{y}_j)^2 \,, \tag{3}$$

$$RRMSE\,(y, \hat{y}) = \frac{\sqrt{MSE}}{\bar{y}} \,, \tag{4}$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{j=1}^{n} (y_j - \hat{y}_j)^2}{\sum_{j=1}^{n} (y_j - \bar{y})^2} \,, \tag{5}$$

where $\hat{y}$ is the predicted value of the sample j, y_j the corresponding true value (the targeted observation) and $\bar{y}$ the average value of y_j for the total n samples.

3 Results

3.1 Annual maps over the 2013-2023 period

The CHROMAP model generates 11 years of annual maps between 2013 and 2023 at 500 m spatial resolution across Europe. The mapped and evaluated air quality indicators include annual average concentrations of NO_2, O_3, PM2.5 and PM10, along with the SOMO35 indicator. Following the workflow described in Section 2.1, CHROMAP is executed independently for each year and each indicator using the consistent methodology. It is therefore essential to ensure that its outputs maintain a consistent temporal evolution throughout the 2013-2023 period and align with observed trends. Since CHROMAP performs spatial downscaling of CAMS Ensemble (CAMS_ens) outputs, the high-resolution maps are compared against CAMS_ens in the results section to assess enhancement performance.

Figure 2 presents the annual evolution of average observed values with standard deviation in Europe (see Sect. 2.3 for the description of the observations) compared to co-located CHROMAP and CAMS_ens estimates for the above-mentioned air quality indicators. The annual evolution of the median values can be found in SM (Fig. S6). CHROMAP predictions (produced by 5-fold cross-validation) are very close to the observed concentrations, with an average relative difference in mean value of less than 1% for each species. The average relative difference in median value is 8.3% for NO_2, 1.3% for O_3, 5.2% for SOMO35, 1.9% for PM2.5 and, 2.4% for PM10. The standard deviation of CHROMAP's predictions is mainly lower than observations (by –2.84 µg/m^3 for NO_2, –3.01 µg/m^3 for O_3, –535.22 µg/m^3 for SOMO35, –0.87 µg/m^3 for PM2.5 and, –1.60 µg/m^3 for PM10). Furthermore, despite independent learning between years, CHROMAP shows high covariance with observations in time (Pearson's coefficient R close to 0.99 for all the indicators), higher than for CAMS_ens (0.77 for NO_2, 0.96 for O_3, 0.91 for SOMO35, 0.78 for PM2.5 and, 0.70 for PM10). Due to the regressions, performed by the Ensemble of ML/DL models, CHROMAP decreases biases and improves interannual variability compared to CAMS_ens. Trends that are observed in Europe are also present in CHROMAP: significant negative ones for NO_2, PM2.5, and PM_{10}, and a positive one for O_3. The consistency of temporal evolution of the estimates produced by CHROMAP over the domain of the study (not only at co-located observation points) is ensured using dynamic predictor variables that vary over time, such as the meteorological fields and the pollutant concentration fields by CTM or satellite in the case of NO_2.

https://doi.org/10.5194/egusphere-2026-1109
Preprint. Discussion started: 6 March 2026

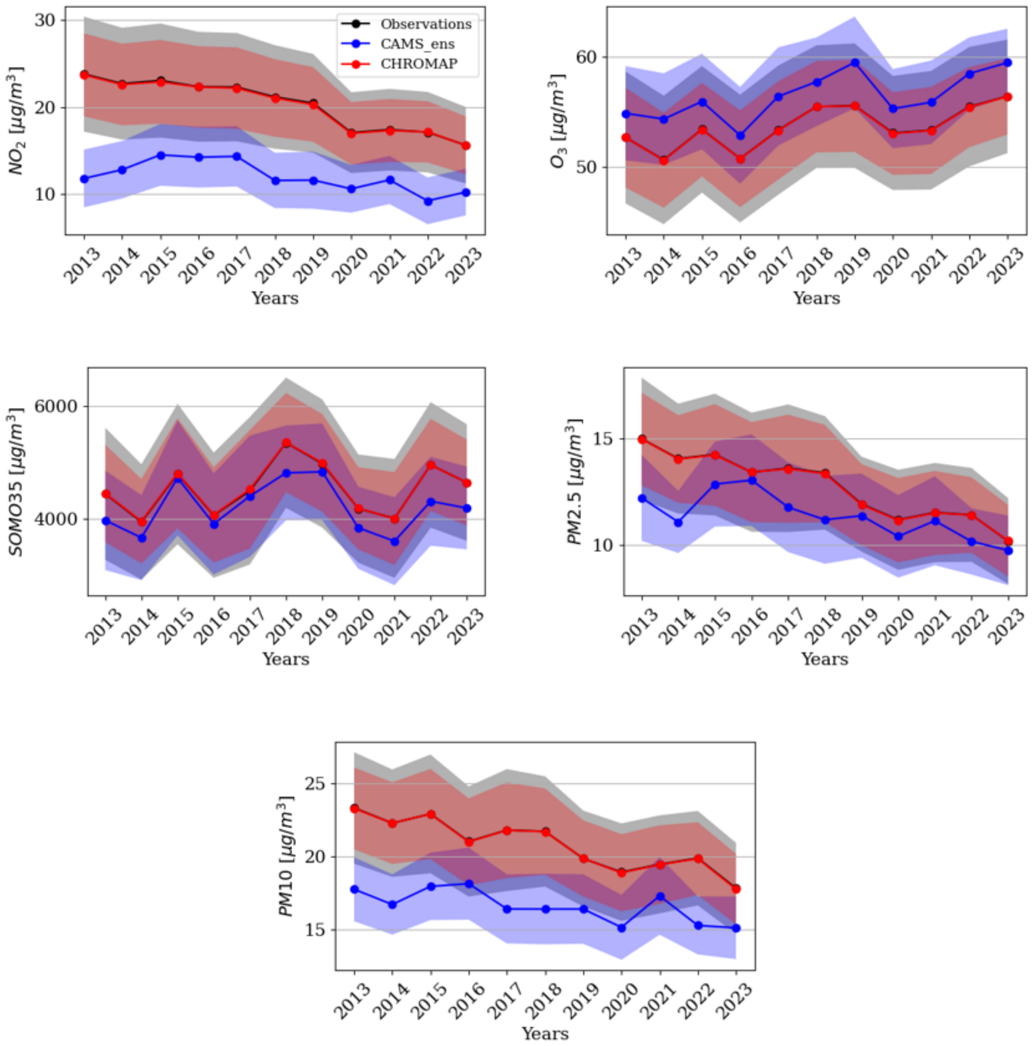


Figure 2: Annual evolution of average concentrations over Europe: observations (in black), compared to co-located CHROMAP (in red) and CAMS_ens (in blue) estimates for NO$_2$ (top left), O$_3$ (top right), SOMO35 (middle left), PM2.5 (middle right) and PM10 (bottom). The shaded curves are the standard deviation range.

An example of a PM2.5 map produced by CHROMAP for the year 2021 is presented in Figure 3. The European domain is covered with a resolution of 500 m, including sparsely monitored regions such as Scandinavia and unmonitored areas such as Eastern Europe, seas and oceans. This demonstrates CHROMAP's capacity to interpolate and extrapolate pollutant concentrations across the domain of the study using the available training set. This capability is due to the high quality of the features employed throughout the domain, particularly the physics-based CTM fields, and for NO$_2$, satellite-derived total column observations. It should be noted that the CHROMAP maps exhibit concentration fields without spatial discontinuities, mainly due to the careful feature selection and processing. Furthermore, CHROMAP generates standard deviation maps of the different predicted values derived from the members of the ensemble of ML/DL models. This allows to evaluate the uncertainty of certain regions in Europe based on the variability of regression models. An example is given for the corresponding PM2.5 map in SM (Fig. S7). Standard deviation values remain relatively low in central Europe (about 2 µg/m^3) but increase over the Atlantic Ocean (about 4 µg/m^3) and the Mediterranean Sea (about 4 µg/m^3), reaching maximum values at the southern (over Algeria) and eastern boundary (over Iraq) of the domain (about 10 µg/m^3).

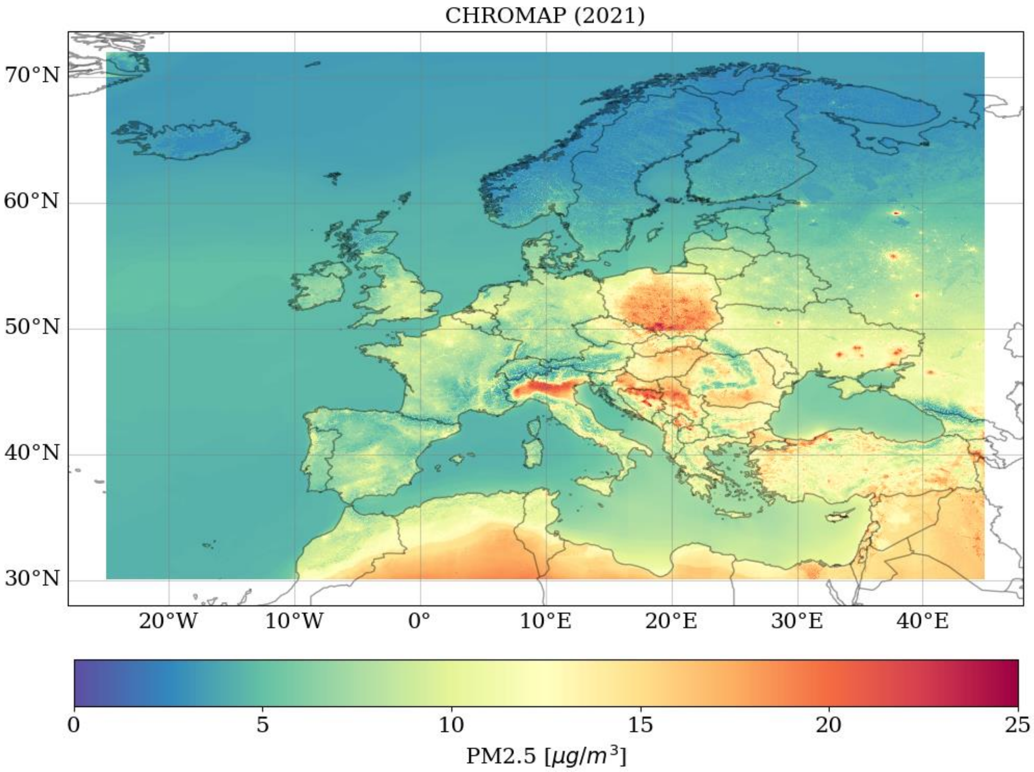


Figure 3: Example of CHROMAP estimate at 500 m spatial resolution throughout Europe for PM2.5 in 2021.

Figure 4 compares CHROMAP's 2021 concentration maps against CAMS_ens in different regions of interests with marked concentration gradients: northern Italy/southern Germany for NO_2, in the Balkans for O_3, and in the Iberian Peninsula region for PM10. CHROMAP provides more details for both urban and rural areas using high spatial resolution predictor variables. Indeed, CHROMAP captures territorial elements that affect surface concentration spatial patterns. For instance, the size and shape of urban centres are represented by the features "Geo_PopDens", "Geo_UrbDeg", "Geo_RoadNet" (e.g. higher PM10 concentrations over Barcelona, Porto and its surroundings), and geomorphological boundaries by the feature "Geo_Alt" (e.g. pronounced spatial gradients at the boundaries of the Sierra Nevada in Andalusia and the Transylvanian Alps in Romania, respectively for PM10 and O_3). Regarding NO_2 distribution, CHROMAP highlights the density of the road network (e.g. roads crossing valley floors in the Alps) as a fundamental predictor (see Sect. 2.2).

Overall, CHROMAP preserves the spatial patterns of CAMS_ens (the dominant feature, see Section 3.2), while overlaying high-resolution geographic features that enhance spatial gradients. Nevertheless, by fitting to observed pollutant concentrations, CHROMAP corrects the biases of CAMS_ens which can lead to concentration differences of several µg/m^3.

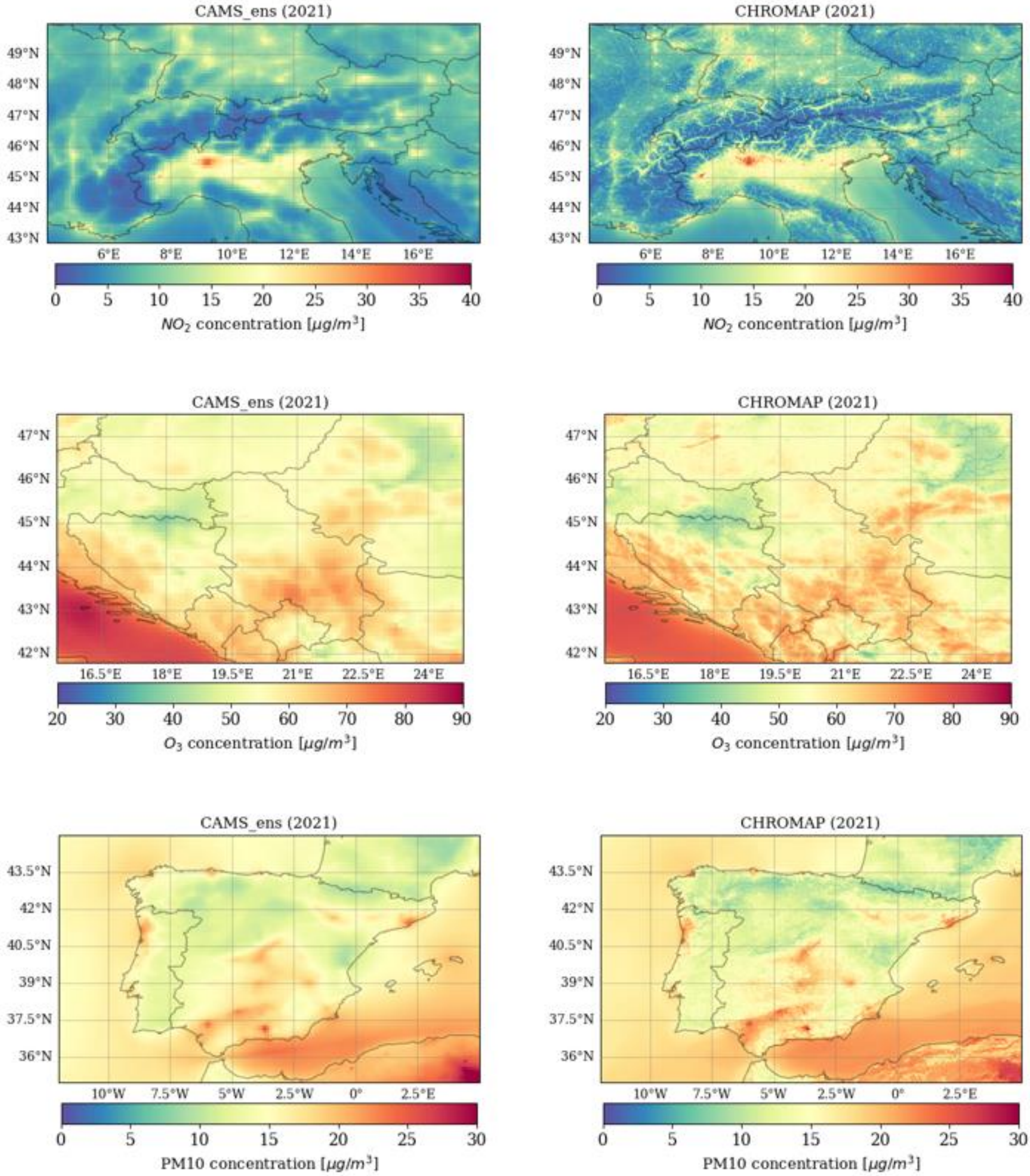


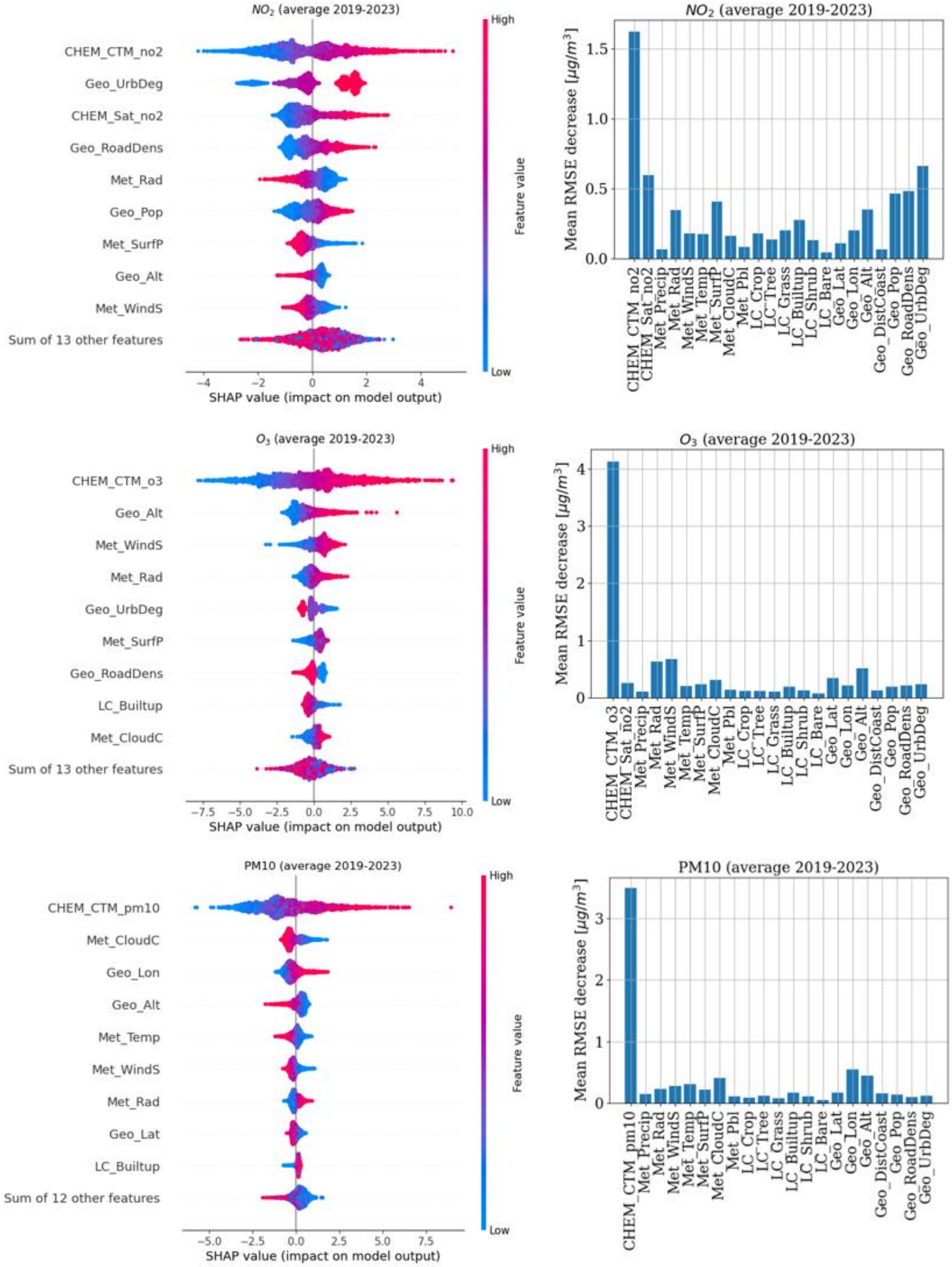
Figure 4: Comparison of annual average concentration of NO_2 (top panels), O_3 (middle panels), and PM10 (bottom panels) in 2021 between CAMS_ens, at 10 km spatial resolution, (left panels) and CHROMAP, at 500 m spatial resolution (right panels).

3.2 Assessment of feature importance

Interpretable machine learning diagnostics are included in the CHROMAP model to better assess the influence of each feature. Figure 5 shows the feature importance expressed with RMSE decrease, and the SHAP values that represent how much each feature contributes to the model outputs (positive SHAP value meaning a positive impact on predictions), for NO_2, O_3 and PM10, on average over the period 2019-2023 (due to feature coverage “CHEM_Sat_NO2”). The same figures for PM2.5 and SOMO35 are presented in SM (see Fig. S8). Details on diagnostic methods and associated references are given in Section 2.1. For all indicators, “CHEM_CTM_$poll” is the most decisive feature for CHROMAP’s predictions. SHAP values are highest for this feature (ranging from -7.5 to +10 depending on the species which is 2 to 4 times higher than other features). The attributed decrease in RMSE is 1.8 $\mu g/m^3$ for NO_2, 4.6 $\mu g/m^3$ for O_3, 1585.5 $\mu g/m^3$ for SOMO35, 3.4 $\mu g/m^3$ for PM2.5 and 4.0 $\mu g/m^3$ for PM10.

Apart for “CHEM_CTM_$poll”, the order and magnitude of feature importance vary depending on the predicted indicator. The most important features for NO_2 concentration predictions are "Geo_UrbDeg", " CHEM_Sat_NO2", "Geo_RoadDens" and "Geo_Pop" (after “CHEM_CTM_$poll”). Weather conditions are also determinant for CHROMAP’s predictions, especially the "Met_Rad" feature, i.e. the solar radiation, which, according to the SHAP values, is anti-correlated with NO_2 concentrations. This can be explained with the conversion of NO_2 into O_3 by photolysis (Jenkin & Clemitshaw, 2000). Altitude (“Geo_Alt”) is another important feature for the prediction of NO_2 and O_3 concentrations. Low-altitude areas (e.g. valley floors and plains) tend to be more urbanised than high-altitude areas (e.g. mountains) and, are therefore associated with considerable NO_x emissions. Titration of O_3 by NO producing NO_2 can lead to a decrease in O_3 at low altitudes densely populated areas. “Met_Winds” is another key feature of O_3 maps which suggests the importance of long-range transport in the spatial distribution of average concentrations (e.g. Safieddine et al., 2014). Concerning PM2.5 and PM10, the geographical features (“Geo_Alt”, “Geo_Lon” and “Geo_Lat”) and the “Met_CloudC” feature are the most decisive after “CHEM_CTM_$poll”. As for NO_2, dense urban areas (generally at low altitudes) are associated with significant primary aerosol emissions (e.g. from traffic and the residential sector). Latitude by the distance to the Sahara, and longitude by the distance to Eastern Europe or the Middle East which present higher average concentrations in both CTM outputs and observations, also determine the spatial distribution of particles. Through complex chemical and physical processes of aerosol formation and depositions, cloud cover impact also the PM surface concentrations (e.g. Belle et al., 2017). SOMO35 maps are mainly driven by the “CHEM_CTM_$poll” feature (mean RMSE decrease of 1586 $\mu g/m^3$) with low influence of the others features (e.g. mean RMSE decrease of about 170 $\mu g/m^3$ for “Geo_Alt” and 58 $\mu g/m^3$ for “Geo_Pop”).

Finally, it should be noted that the estimated importance values (from both SHAP and permutation framework) are sensitive to inter-feature correlations, meaning that results can vary if the analysis exclude any feature as this would alter the regression model. All features presented in this study are consistently applied to produce the maps, especially since they all contribute to a reduction in RMSE to varying degrees.



415 **Figure 5: Permutation feature importance expressed as a decrease in RMSE (right panel) and distribution of SHAP values (left panel) for NO_2 (top), O_3 (middle) and PM10 (bottom). The distribution of SHAP values is represented only for the first 9 features in order of importance. The same figures for PM2.5 and SOMO35 are presented in SM (Fig. S8).**

3.3 Performance evaluation at monitoring stations

3.3.1 Independent validation

CHROMAP performance is assessed in execution mode using independent observations that were not used during the model training. The year 2021 was chosen due to a considerable number of observations while including validated reanalyses in the predictor variables. The test set represents 20 % of the total number of observations for this year. The monitoring stations were selected by stratified random sampling according to the type of station (background, industrial and traffic). To ensure robust representativeness, the distribution of values for training is similar to that of the test set for all air quality indicators. Histograms of pollutant concentration values of this test set are available in SM (See Fig. S9). RRMSE and R^2 scores are presented in Table for each air quality indicator. Scatter plot of concentrations predicted by CHROMAP compared to those observed can be found in SM (see Fig. S10). For all indicators, R^2 is at least greater than or equal to 0.61, with the best values for PM10 (0.68) and PM2.5 (0.78). PM2.5 presents the lowest RRMSE with 18.95 % (for a mean concentration of 11.77 µg/m^3 in the observations). The scores are lower for SOMO35 with an R^2 of 0.61 and an RRMSE of 33.29 % (for a mean value of 3889.49 µg/m^3).

Table 2: Performance scores based on the test set composed of 20% of the pollutant observations from the year 2021 by stratified random sampling according to the type of monitoring stations.

Metric	RRMSE [%]	R^2	Observed average value (and standard deviation) [µg/m^3]	Predicted average value (and standard deviation) [µg/m^3]	Number of stations in the test set
Average NO_2	32.59	0.66	16.94 (9.41)	17.06 (7.23)	710
Average O_3	11.48	0.62	52.96 (9.86)	53.26 (7.88)	425
SOMO35	33.29	0.61	3889.49 (2049.87)	4017.67 (1578.70)	421
Average PM2.5	18.95	0.78	11.77 (4.67)	11.39 (3.93)	399
Average PM10	17.79	0.68	19.50 (6.09)	19.55 (5.04)	597

3.3.2 Cross-validation and comparison to CAMS performance over the 2013-2023 period

Previous analysis demonstrated that CHROMAP maps are more detailed than CAMS_ens with a considerable increase in resolution (see Sect. 3.1). This subsection evaluates the quantitative improvements in prediction scores (R^2 and RRMSE). Table summarizes the average score differences between CHROMAP and CAMS_ens over the 2013-2023 period and for all observation stations (all types combined), based on a 5-fold cross-validation. For all indicator, RRMSE decreases and R^2 increases with CHROMAP. The most substantial improvements occur for PM10 showing a 37.2 % RRMSE reduction and

+36.0 % R^2 increase, and for NO_2 showing a 32.7 % RRMSE reduction and +28.1 % R^2 increase. Improvements are also obtained for O_3, SOMO35 and PM2.5; with RRMSE reduction of –21.2 %, 9.8 % and 21.5 % respectively, and R^2 increases of +33.7 %, 18.0 % and 14.4 %.

Table 3: Average relative difference [%] between CHROMAP and CAMS_ens performance scores (calculated as CHOM-CAMS/CAMS) over the 2013-2023 period for all stations.

Metric	**Relative RRMSE difference between CHROMAP and CAMS_ens**	**Relative R^2 difference between CHROMAP and CAMS_ens**
Average NO_2	-32.7 %	+28.1 %
Average O_3	-21.2 %	+33.7 %
SOMO35	-9.8 %	+18.0 %
Average PM2.5	-21.5 %	+14.4 %
Average PM10	-37.2 %	+36.0 %

Additionally, an evaluation of the CHROMAP predictions (5-fold cross-validation) for each year between 2013 and 2023 is conducted by type of monitoring station: background, industrial, and traffic. The performance scores from CHROMAP estimates are compared to those of CAMS_ens in Figure 6, Figure 7 and Figure 8 for NO_2, O_3 and PM10 respectively. The same figures are available for SOMO35 and PM2.5 in SM (see Fig. S11 and Fig. S12). For all years and species, CHROMAP allows to reduce the error (RRMSE decrease) and to improve the variance of predictions (R^2 increase) at each monitoring station type. The decrease in RRMSE with CHROMAP compared to CAMS_ens is generally larger for industrial and traffic stations. With a spatial resolution of ~10 km x 10 km, CAMS_ens is more suited to represent background concentrations, while CHROMAP with a spatial resolution of ~500 m x 500 m, is expected to be more representative of industrial areas and large traffic sections, especially since ML/DL regressions are trained using the three types of monitoring stations.

CHROMAP estimates present for NO_2 a R^2 between 0.70 and 0.77 for background stations (depending on the year), 0.62 and 0.69 for industrial stations and, between 0.42 and 0.52 for traffic stations. The RRMSE related to traffic stations is almost divided by two with the application of CHROMAP (compared to CAMS_ens). For both CAMS_ens and CHROMAP, the RRMSE (for any station type) decrease between 2013 and 2023. This improvement can be explained by both better modeling capacities in the CAMS reanalysis and the increase of the observation number. Concerning O_3 (SOMO35), CHROMAP estimates present a R^2 between 0.56 and 0.68 (0.64 and 0.77) for background stations, 0.51 and 0.65 (0.41 and 0.63) for industrial stations and, between 0.27 and 0.55 (0.34 and 0.48) for traffic stations. Although it is reduced by about 30% with CHROMAP, the RRMSE at traffic stations remains higher than in other types of stations. Finally, CHROMAP estimates present for PM10 (PM2.5) a R^2 between 0.70 and 0.77 (0.71 and 0.83) for background stations, 0.48 and 0.59 (0.56 and 0.65) for industrial stations and, between 0.58 and 0.73 (0.61 and 0.81) for traffic stations. The analysis reveals that industrial monitoring stations have the highest RRMSE values for $PM_{2.5}$ and PM_{10} in all years, while traffic stations have the highest RRMSE for the remaining air quality indicators.

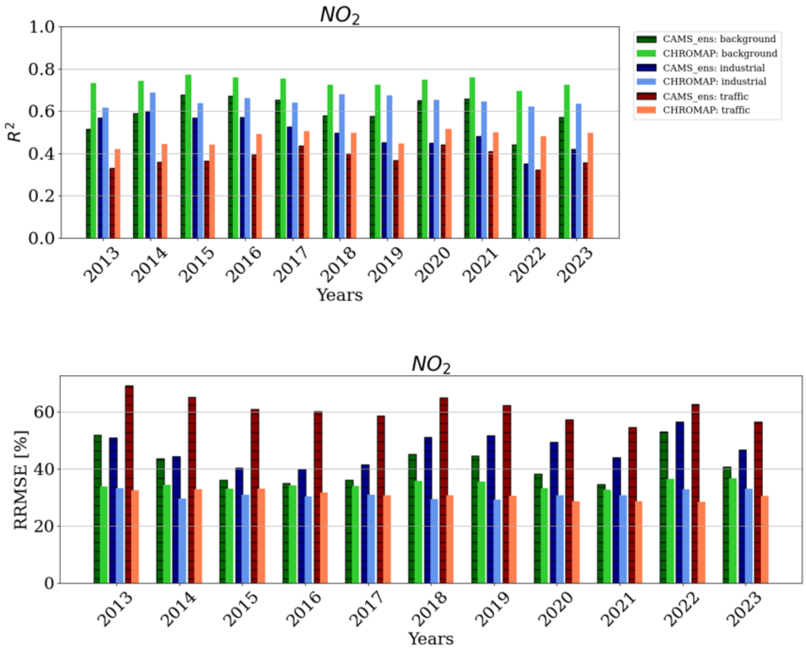


Figure 6: Comparison of performance scores (R^2 in the upper panel and RRMSE in the lower panel) between CHROMAP at 500 m spatial resolution (light bars) and CAMS_ens at 10 km spatial resolution (dark bars) for NO_2. The scores are presented by monitoring stations type: background in green, industrial in blue and traffic in red.

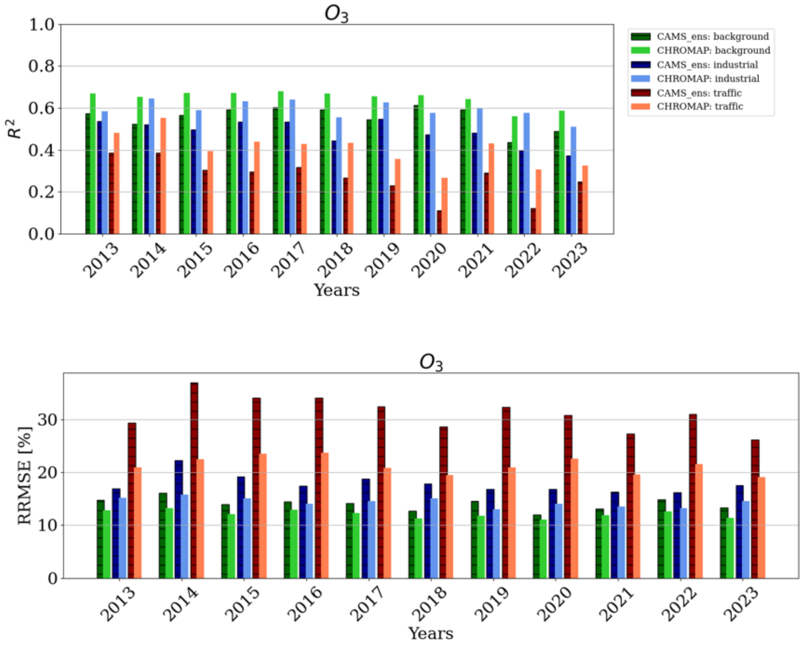


Figure 7: Same as Figure 6 for O_3.

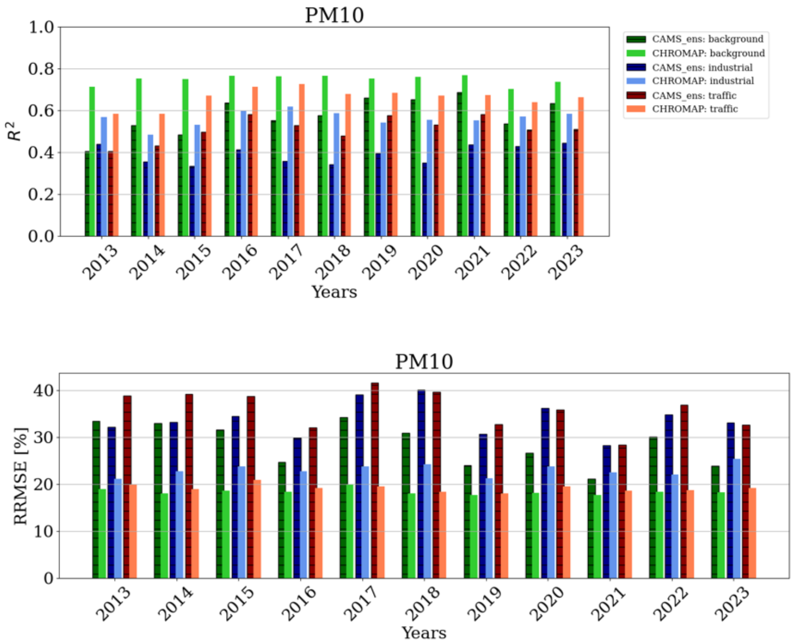


Figure 8: Same as Figure 6 for PM10.

3.4 Impact on population exposure to air pollution

In studies assessing the impact of air pollution on health, health indicators are generally directly proportional to exposure data. Analysing the impact of using CHROMAP downscaled fields rather than CAMS_ens values on population exposure data is therefore a good indicator of the impact this will have on health impact assessment.

Exposure is calculated in this paper as the annual average atmospheric concentration of pollutants weighted by population density on each grid cell of the domain. By applying a mask for each country, exposure is calculated as follows:

$$\sum_{i=1}^{n} \frac{p_i \times conc_i}{tot\_p}, \qquad (6)$$

where n is the number of grid cells in the corresponding country, i the index of the respective cell, p_i the number of inhabitants per cell, $conc_i$ the average annual concentration per cell, and tot_p the total population of the country.

The concentration fields from the CAMS reanalyses dataset (initially at 10 km spatial resolution) were re-interpolated using a bilinear approach at 500 m to compare exposure with that calculated from CHROMAP results. The population density dataset is derived from the Global Human Settlement Layer product from the Joint Research Centre (Pesaresi et al., 2024) with an initial spatial resolution of 100 m. The population has been aggregated on a 500 m grid. The population is fixed (year 2015 taken) in the calculation of the annual exposure for comparison over the period. The average exposure with CAMS_ens and CHROMAP is calculated for each European country. To avoid significant differences compared to countries or principalities characterized by a small population, Liechtenstein was merged with Austria, Monaco with France, San Marino and the Vatican with Italy, Andorra and Gibraltar with Spain, and Guernsey, Jersey and the Isle of Man with the UK. The total population per country can be found in SM (Table S2).

Figure 9 shows scatter plots comparing exposure values averaged at a national level for each European country (between 2013 and 2023) calculated using CHROMAP versus CAMS_ens concentration fields with identical population data for NO_2, O_3, SOMO35, PM2.5 and PM10. Results show that the exposure can be significantly different based on high spatial resolution estimates of pollutant concentrations. On average, exposure varies by +120 % for NO_2, -7% for O_3, +12% for SOMO35, +17% for PM2.5 and +39% for PM10 when using concentration fields from CHROMAP compared to CAMS_ens. These differences can be explained by two phenomena: first, it has been shown that, compared to CAMS data, CHROMAP data allow for additional correction of concentration biases relative to measurements (see RRMSE reduction in Sect. 3.3). CAMS_ens tends to underestimate the concentrations of NO_2, $PM_2.5$ and PM10, and to overestimate O_3 concentration compared to in situ observations. Second, with a resolution of 500 m, CHROMAP data make it possible to locate high concentrations of pollutants closer to their source. This is particularly true for NO_2, which is emitted by traffic and then decreases fairly rapidly. In large cities, NO_2 concentrations are therefore associated with higher population densities, which explains the very large differences between the exposure data calculated with CHROMAP and with CAMS_ens for NO_2 data population exposure in particular.

It should be emphasized that the calculated exposure differences between CHROMAP and CAMS_ens have decreased for recent years. For example, the exposure differences for PM2.5 varied from +36 % with CHROMAP in 2013 to +9 % in 2023.

This may correspond to the improvement in pollutant concentration modeling with CAMS_ens over the past 10 years (see RRMSE reduction in Sect. 3.3). Finally, as exposure is calculated from the population-weighted concentrations, densely populated areas generally characterized by high concentrations of NO_2, PM2.5 and PM10 account for a significant weight in the national average exposure calculation. Being specifically designated to make high resolution estimates and including features related to anthropogenic activities, CHROMAP shows higher concentrations on these areas than CAMS_ens.

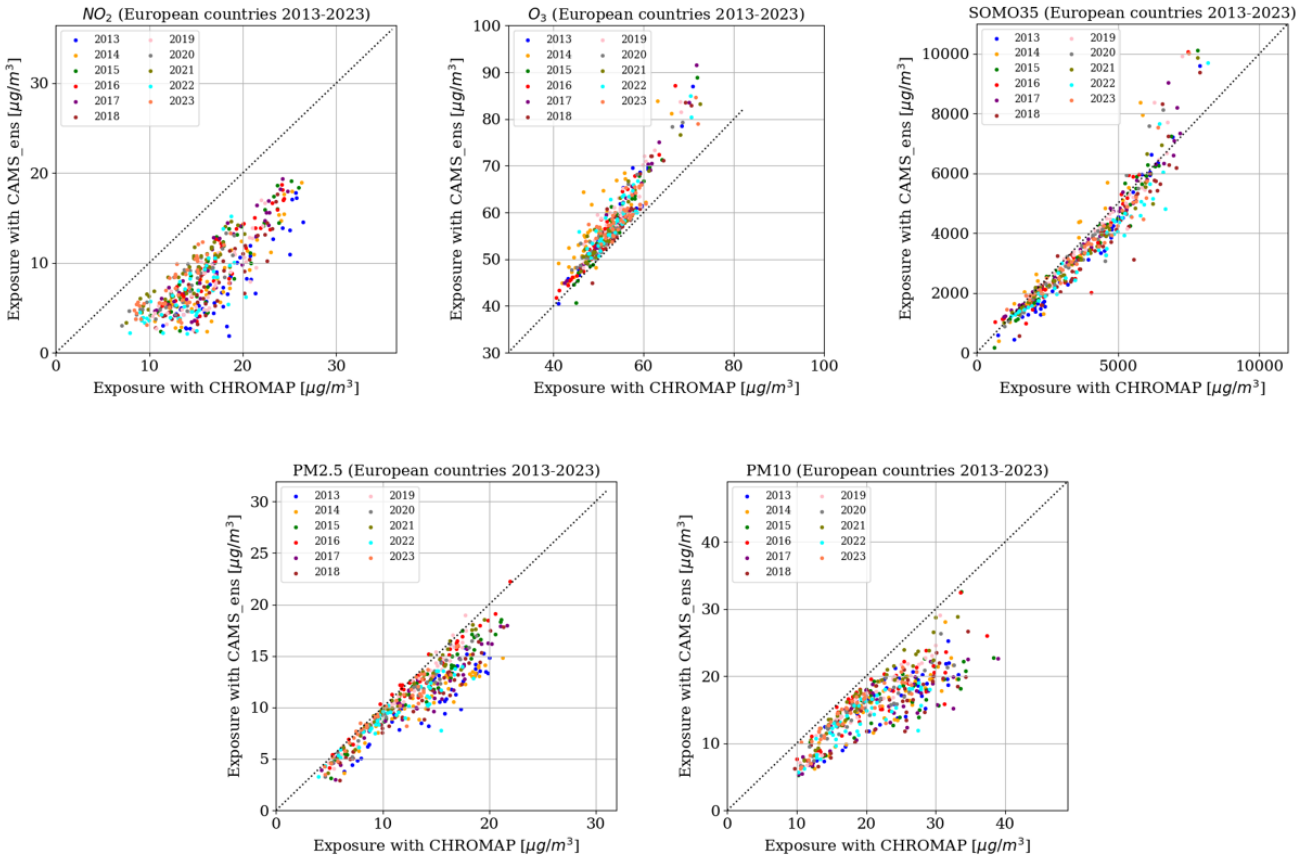


Figure 9: Scatter plot of population exposure values calculated using pollutant concentration fields generated by CHROMAP (500 m spatial resolution) and CAMS_ens (10 km spatial resolution), using the same population data. Each point represents the average value for a European country for a year between 2013 and 2023.

4. Discussions

The quality of CHROMAP's predictions depends considerably on the quality of its input features. A key challenge in this work is identifying high-spatial resolution predictor variables that provide a complete coverage of the European domain in a spatial homogeneous manner ensuring accurate and coherent production of air quality maps. There are databases at the national level that could improve our set of features, but this would require significant work to aggregate or reconstruct a complete set across Europe with a significant risk of spatial discontinuity. For instance, emissions inventories could improve prediction accuracy, but their application is typically limited to specific regions. The issue of spatial inconsistency when spatializing the annual totals reported by different European countries is often raised (Kuenen et al., 2022; López-Aparicio et al., 2017). Since CHROMAP includes CTM outputs that already incorporate spatialised emission inventories in their workflow, this feature is not included within CHROMAP, at least not at European level. More generally, further research identifying the key factors and processes controlling the spatial distribution of air pollutants would be beneficial to further improving this approach.

Another important discussion point concerns the number and spatial distribution of observations used for training the ML/DL models. While the dataset includes stations across Europe with sufficient observations for robust training over the period of the study, RRMSE scores for NO_2, PM2.5 and PM10 improved considerably between 2013 and 2023, partly due to the increased number of monitoring stations. However, the heterogeneous spatial distribution and absence of observations in certain regions/countries present limitations. For Eastern Europe, North Africa or the Middle East, CHROMAP must extrapolate from training data primarily derived from Western, Southern and Northern Europe to these other regions characterized by different emission source distributions and meteorological conditions. The potential of low-cost sensor data also needs to be considered. With rigorous preprocessing, these datasets could improve high-resolution predictive performance, although most applications remain at the local scale (e.g. Liang et al., 2023; Lim et al., 2019). Enhanced spatial coverage through additional observations, representative of diverse environmental conditions, would improve CHROMAP's prediction performance while reducing associated uncertainty.

Nevertheless, the use of satellite data across the study domain significantly reduces uncertainty in extrapolation zones. Although satellite observations, typically expressed in atmospheric columns, do not directly correspond to surface concentrations, they still provide a valuable spatial representation of the distribution of pollutants. In the case of CHROMAP, the NO_2 columns were used as features. Several studies used the Aerosol Optical Depth as a predictor variable (e.g. Li et al., 2018; Shetty et al., 2025), but this requires processing steps to fill gaps and remove noise from the data in order to be used for air quality mapping, especially at high spatial resolution or temporal frequency (e.g. Hough et al., 2021).

Our performance scores (calculated by 5-fold cross-validation) are in line with and, in some cases, even outperform those reported in the literature, both in terms of average value and by type of monitoring station. For instance, on average over the comparable period of 2013-2019, CHROMAP's R^2 values are close to those of Shen et al. (2022) at a resolution of 25m (by 5-fold cross-validation, as in our method) for NO_2, PM2.5, and PM10 and about 0.05 higher for O_3. Compared to the RIMM scores (leave-one-out cross-validation) performed in Horálek et al. (2023) on the final maps of 2020 at 1 km resolution, the CHROMAP RMSE decreases by -0.1 $\mu g/m^3$ and -2.9 $\mu g/m^3$ for PM2.5 and NO_2 respectively at the traffic stations, but is slightly higher by +0.2 $\mu g/m^3$ and +1.5 $\mu g/m^3$ respectively at the background stations.

Our results demonstrate that population exposure averaged at the national level varies significantly when using high-spatial resolution maps from CHROMAP compared to low-resolution regional maps, with the magnitude of differences depending on the specific pollutant considered. This finding aligns with Denby et al. (2024) who also identify significant impact of sub grid variability on exposure assessments. It is important to acknowledge that the sensitivity analysis presented in this article is based on a static (residential) exposure calculation methodology. More sophisticated methods that incorporate dynamic exposure patterns, accounting for individual mobility and varying pollutant concentrations across different environments (workplace, home, transport, or leisure), especially at 500 m spatial resolution. However, the current analysis focuses on annual indicators to initially assess the average impact. Supporting this approach, Marquet et al. (2023) compared static and dynamic exposure methodologies (by GPS tracking) for air pollution assessment in Barcelona. Their findings revealed significant differences only for PM2.5 (by a variation of 6% on average). They suggest that unless going down to the street level, static exposure calculations provide sufficiently accurate estimates for several pollutants including NO_2 and PM10, in the context of their study.

Finally, this work is part of a transparency initiative with free access to code and data, and version and update tracking. The annual production of high-resolution maps for Europe is planned to follow the release of validated CAMS reanalyses (see sections Data and Code Availability).

5. Conclusions

This article presents CHROMAP, a high spatial resolution air quality mapping model based on the fusion of data from deterministic models, in-situ and satellite observations, and spatial proxies using an ensemble of ML and DL algorithms. Yearly estimates of the SOMO35 indicator and the average concentrations of NO_2, PM2.5, PM10, and O_3 are produced and evaluated for the 2013-2023 period at a spatial resolution of 500 meters over the European domain. The methodology maintains consistency across all pollutant indicators while ensuring flexibility and transferability. In addition, CHROMAP includes interpretable AI diagnostics and provides uncertainty estimation from ensemble variability.

The results of the evaluation show that CHROMAP exhibits consistent temporal evolution throughout 2013-2023, accurately capturing inter-annual variability and trends in observations. Feature importance analysis highlighted that concentration fields from CTMs are essential for coherent and accurate map production. At high spatial resolution, CHROMAP significantly improves prediction performance for background, industrial and traffic monitoring station types compared to CAMS reanalysis (R^2 increase up to +36 % for PM10 and RMSE decrease by -37 %). Furthermore, applying CHROMAP's high spatial resolution maps for static population exposure calculations can result in significant differences, particularly for NO_2, compared to CAMS reanalysis.

The development of CHROMAP establishes promising methodological foundations for high spatial resolution pollutant concentration fields through data fusion approaches. By ensuring sufficient availability of in-situ observations and concentration fields from CTMs for downscaling, this methodology could be extended to additional air quality indicators and applied at higher temporal frequency, opening new opportunities for comprehensive air quality assessment.

Code and data availability

The CHROMAPv1.0 code is available at https://zenodo.org/records/18846210 (Guion, 2026). Gridded files of annual concentrations of NO_2, PM2.5, PM10, O_3, and the SOMO35 indicator, as well as their associated standard deviation, produced with CHROMAPv1.0 for the period 2013-2023 are available upon request.

Author contributions

AGU: Conceptualization, Methodology, Data collection, Software development, Analysis, Writing-original draft and Supervision; AGR: Methodology, Data collection, Writing-Reviewing and Editing; GDE: Data collection, Writing-Reviewing and Editing; YJA: Methodology and Analysis; ERE, AUN, FME, SSC and ACO: Writing-Reviewing and Editing.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

"Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies

Financial support

This work was supported by the "CARTOKRIF" project funded by the "Astreinte Qualité de l'Air" decided by the French Council of State, and the "VALESOR" project (grant No. 101095611) funded by the European health and digital executive agency (HADEA).

References

Baño-Medina, J., Manzanas, R., and Gutiérrez, J. M.: On the suitability of deep convolutional neural networks for continental-wide downscaling of climate change projections, Clim. Dyn., 57, 2941–2951, https://doi.org/10.1007/s00382-021-05847-0, 2021.

Beauchamp, M., de Fouquet, C., and Malherbe, L.: Dealing with non-stationarity through explanatory variables in kriging-based air quality maps, Spat. Stat., 22, 18–46, https://doi.org/10.1016/j.spasta.2017.08.003, 2017.

Belle, J. H., Chang, H. H., Wang, Y., Hu, X., Lyapustin, A., and Liu, Y.: The Potential Impact of Satellite-Retrieved Cloud Parameters on Ground-Level PM2.5 Mass and Composition, Int. J. Environ. Res. Public. Health, 14, 1244, https://doi.org/10.3390/ijerph14101244, 2017.

Bergstra, J. and Bengio, Y.: Random search for hyper-parameter optimization, J. Mach. Learn. Res., 1st February, 281–305, 2012.

Bertrand, J.-M., Meleux, F., Ung, A., Descombes, G., and Colette, A.: Technical note: Improving the European air quality forecast of the Copernicus Atmosphere Monitoring Service using machine learning techniques, Atmospheric Chem. Phys., 23, 5317–5333, https://doi.org/10.5194/acp-23-5317-2023, 2023.

Bessagnet, B., Beauchamp, M., Menut, L., Fablet, R., Pisoni, E., and Thunis, P.: Deep learning techniques applied to super-resolution chemistry transport modeling for operational uses, Environ. Res. Commun., 3, 085001, https://doi.org/10.1088/2515-7620/ac17f7, 2021.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Li, L., and Tarko, A.: Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual, Zenodo, https://doi.org/10.5281/ZENODO.3938963, 2020.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794, https://doi.org/10.1145/2939672.2939785, 2016.

Colette, A., Rouïl, L., Meleux, F., Lemaire, V., and Raux, B.: Air Control Toolbox (ACT_v1.0): a flexible surrogate model to explore mitigation scenarios in air quality forecasts, Geosci. Model Dev., 15, 1441–1465, https://doi.org/10.5194/gmd-15-1441-2022, 2022.

Colette, A., Collin, G., Besson, F., Blot, E., Guidard, V., Meleux, F., Royer, A., Petiot, V., Miller, C., Fermond, O., Jeant, A., Adani, M., Arteta, J., Benedictow, A., Bergström, R., Bowdalo, D., Brandt, J., Briganti, G., Carvalho, A. C., Christensen, J. H., Couvidat, F., D'Elia, I., D'Isidoro, M., Denier Van Der Gon, H., Descombes, G., Di Tomaso, E., Douros, J., Escribano, J., Eskes, H., Fagerli, H., Fatahi, Y., Flemming, J., Friese, E., Frohn, L., Gauss, M., Geels, C., Guarnieri, G., Guevara, M., Guion, A., Guth, J., Hänninen, R., Hansen, K., Im, U., Janssen, R., Jeoffrion, M., Joly, M., Jones, L., Jorba, O., Kadantsev, E., Kahnert, M., Kaminski, J. W., Kouznetsov, R., Kranenburg, R., Kuenen, J., Lange, A. C., Langner, J., Lannuque, V., Macchia, F., Manders, A., Mircea, M., Nyiri, A., Olid, M., Pérez García-Pando, C., Palamarchuk, Y., Piersanti, A., Raux, B., Razinger, M., Robertson, L., Segers, A., Schaap, M., Siljamo, P., Simpson, D., Sofiev, M., Stangel, A., Struzewska, J., Tena, C., Timmermans, R., Tsikerdekis, T., Tsyro, S., Tyuryakov, S., Ung, A., Uppstu, A., Valdebenito, A., Van Velthoven, P., Vitali, L., Ye, Z., Peuch, V.-H., and Rouïl, L.: Copernicus Atmosphere Monitoring Service – Regional Air Quality Production System v1.0, https://doi.org/10.5194/egusphere-2024-3744, 12 December 2024.

Colette, A., Collin, G., Besson, F., Blot, E., Guidard, V., Meleux, F., Royer, A., Petiot, V., Miller, C., Fermond, O., Jeant, A., Adani, M., Arteta, J., Benedictow, A., Bergström, R., Bowdalo, D., Brandt, J., Briganti, G., Carvalho, A. C., Christensen,

J. H., Couvidat, F., D'Elia, I., D'Isidoro, M., Denier Van Der Gon, H., Descombes, G., Di Tomaso, E., Douros, J., Escribano, J., Eskes, H., Fagerli, H., Fatahi, Y., Flemming, J., Friese, E., Frohn, L., Gauss, M., Geels, C., Guarnieri, G., Guevara, M., Guion, A., Guth, J., Hänninen, R., Hansen, K., Im, U., Janssen, R., Jeoffrion, M., Joly, M., Jones, L., Jorba, O., Kadantsev, E., Kahnert, M., Kaminski, J. W., Kouznetsov, R., Kranenburg, R., Kuenen, J., Lange, A. C., Langner, J., Lannuque, V., Macchia, F., Manders, A., Mircea, M., Nyiri, A., Olid, M., Pérez García-Pando, C., Palamarchuk, Y., Piersanti, A., Raux, B., Razinger, M., Robertson, L., Segers, A., Schaap, M., Siljamo, P., Simpson, D., Sofiev, M., Stangel, A., Struzewska, J., Tena, C., Timmermans, R., Tsikerdekis, T., Tsyro, S., Tyuryakov, S., Ung, A., Uppstu, A., Valdebenito, A., Van Velthoven, P., Vitali, L., Ye, Z., Peuch, V.-H., and Rouïl, L.: Copernicus Atmosphere Monitoring Service – Regional Air Quality Production System v1.0, Geosci. Model Dev., 18, 6835–6883, https://doi.org/10.5194/gmd-18-6835-2025, 2025.

Copernicus: Copernicus Digital Elevation Model, 2022.

Denby, B. R., Gauss, M., Wind, P., Mu, Q., Grøtting Wærsted, E., Fagerli, H., Valdebenito, A., and Klein, H.: Description of the uEMEP_v5 downscaling approach for the EMEP MSC-W chemistry transport model, Geosci. Model Dev., 13, 6303–6323, https://doi.org/10.5194/gmd-13-6303-2020, 2020.

Denby, B. R., Kiesewetter, G., Nyiri, A., Klimont, Z., Fagerli, H., Wærsted, E. G., and Wind, P.: Sub-grid Variability and its Impact on Exposure in Regional Scale Air Quality and Integrated Assessment Models: Application of the uEMEP Downscaling Model, Atmos. Environ., 333, 120586, https://doi.org/10.1016/j.atmosenv.2024.120586, 2024.

EEA: Air Quality e-Reporting - Annual AQ statistics, 2025.

European Environment Agency: Harm to human health from air pollution in Europe: burden of disease status, Publications Office, LU, 2024.

van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., and Veefkind, J. P.: TROPOMI ATBD of the total and tropospheric NO2 data products, 2024.

Gouldsbrough, L., Hossaini, R., Eastoe, E., Young, P. J., and Vieno, M.: A machine learning approach to downscale EMEP4UK: analysis of UK ozone variability and trends, Atmospheric Chem. Phys., 24, 3163–3196, https://doi.org/10.5194/acp-24-3163-2024, 2024.

Grange, S. K., Farren, N. J., Vaughan, A. R., Rose, R. A., and Carslaw, D. C.: Strong Temperature Dependence for Light-Duty Diesel Vehicle NOx Emissions, Environ. Sci. Technol., 53, 6587–6596, https://doi.org/10.1021/acs.est.9b01024, 2019.

Guion, A., Turquety, S., Cholakian, A., Polcher, J., Ehret, A., and Lathière, J.: Biogenic isoprene emissions, dry deposition velocity, and surface ozone concentration during summer droughts, heatwaves, and normal conditions in southwestern Europe, Atmospheric Chem. Phys., 23, 1043–1071, https://doi.org/10.5194/acp-23-1043-2023, 2023.

Guion, A., Couvidat, F., Guevara, M., and Colette, A.: Country- and species-dependent parameters for the heating degree day method to distribute NO_x and PM emissions from residential heating in the EU 27: application to air quality modeling and multi-year emission projections, Atmospheric Chem. Phys., 25, 2807–2827, https://doi.org/10.5194/acp-25-2807-2025, 2025.

Guion, A. CHROMAPv1.0. Zenodo. https://doi.org/10.5281/zenodo.18846210, 2016.

Hamm, N. A. S., Finley, A. O., Schaap, M., and Stein, A.: A spatially varying coefficient model for mapping PM10 air quality at the European scale, Atmos. Environ., 102, 393–405, https://doi.org/10.1016/j.atmosenv.2014.11.043, 2015.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, A., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.adbb2d47, 2023.

Hoerl, A. E. and Kennard, R. W.: Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics, 12, 55–67, https://doi.org/10.1080/00401706.1970.10488634, 1970.

Horálek, J., Hamer, P., Schreiberová, M., and Schneider, P.: ETC/ATNI Report 2020/11: European air quality interim mapping under ETC/ATNI. Evaluation of AQ mapping using UTD measurement and CAMS forecast modeling data: an approach for more timely European AQ annual maps?, Zenodo, https://doi.org/10.5281/ZENODO.6368242, 2021.

Horálek, J., Vlasáková, L., Schreiberová, M., Marková, J., Schneider, P., Kurfürst, P., Tognet, F., Schovánková, J., Vlček, O., and Damašková, D.: European air quality maps for 2020. PM10, PM2.5, Ozone, NO2, NOx and Benzo(a)pyrene spatial estimates and their uncertainties, European Topic Centre on Human Health and the Environment, 2023.

Hough, I., Sarafian, R., Shtein, A., Zhou, B., Lepeule, J., and Kloog, I.: Gaussian Markov random fields improve ensemble predictions of daily 1 km PM2.5 and PM10 across France, Atmos. Environ., 264, 118693, https://doi.org/10.1016/j.atmosenv.2021.118693, 2021.

Jenkin, M. E. and Clemitshaw, K. C.: Ozone and other secondary photochemical pollutants: chemical processes governing their formation in the planetary boundary layer, Atmos. Environ., 34, 2499–2527, https://doi.org/10.1016/S1352-2310(99)00478-1, 2000.

Juginović, A., Vuković, M., Aranza, I., and Biloš, V.: Health impacts of air pollution exposure from 1990 to 2019 in 43 European countries, Sci. Rep., 11, 22516, https://doi.org/10.1038/s41598-021-01802-5, 2021.

Kaneko, H.: Cross-validated permutation feature importance considering correlation between features, Anal. Sci. Adv., 3, 278–287, https://doi.org/10.1002/ansa.202200018, 2022.

Kavassalis, S. C. and Murphy, J. G.: Understanding ozone-meteorology correlations: A role for dry deposition: Ozone-Meteorology Correlations: Dry Dep, Geophys. Res. Lett., 44, 2922–2931, https://doi.org/10.1002/2016GL071791, 2017.

Kopczewska, K.: Spatial machine learning: new opportunities for regional science, Ann. Reg. Sci., 68, 713–755, https://doi.org/10.1007/s00168-021-01101-x, 2022.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier van der Gon, H.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modeling, Earth Syst. Sci. Data, 14, 491–515, https://doi.org/10.5194/essd-14-491-2022, 2022.

Li, R., Ma, T., Xu, Q., and Song, X.: Using MAIAC AOD to verify the PM2.5 spatial patterns of a land use regression model, Environ. Pollut., 243, 501–509, https://doi.org/10.1016/j.envpol.2018.09.026, 2018.

Liang, L., Daniels, J., Bailey, C., Hu, L., Phillips, R., and South, J.: Integrating low-cost sensor monitoring, satellite mapping, and geospatial artificial intelligence for intra-urban air pollution predictions, Environ. Pollut., 331, 121832, https://doi.org/10.1016/j.envpol.2023.121832, 2023.

Lim, C. C., Kim, H., Vilcassim, M. J. R., Thurston, G. D., Gordon, T., Chen, L.-C., Lee, K., Heimbinder, M., and Kim, S.-Y.: Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea, Environ. Int., 131, 105022, https://doi.org/10.1016/j.envint.2019.105022, 2019.

López-Aparicio, S., Guevara, M., Thunis, P., Cuvelier, K., and Tarrasón, L.: Assessment of discrepancies between bottom-up and regional emission inventories in Norwegian urban areas, Atmos. Environ., 154, 285–296, https://doi.org/10.1016/j.atmosenv.2017.02.004, 2017.

Lundberg, S. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, https://doi.org/10.48550/ARXIV.1705.07874, 2017.

Makar, P. A., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Milbrandt, J., Im, U., Balzarini, A., Baró, R., Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M. D., Pabla, B., Pérez, J. L., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., Zhang, J., and Galmarini, S.: Feedbacks between air pollution and weather, part 2: Effects on chemistry, Atmos. Environ., 115, 499–526, https://doi.org/10.1016/j.atmosenv.2014.10.021, 2015.

Marquet, O., Tello-Barsocchini, J., Couto-Trigo, D., Gómez-Varo, I., and Maciejewska, M.: Comparison of static and dynamic exposures to air pollution, noise, and greenness among seniors living in compact-city environments, Int. J. Health Geogr., 22, https://doi.org/10.1186/s12942-023-00325-8, 2023.

McCulloch, W. S. and Pitts, W.: A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys., 5, 115–133, https://doi.org/10.1007/bf02478259, 1943.

Molnar, C., Casalicchio, G., and Bischl, B.: Interpretable Machine Learning -- A Brief History, State-of-the-Art and Challenges, https://doi.org/10.48550/ARXIV.2010.09337, 2020.

Mu, Q., Denby, B. R., Wærsted, E. G., and Fagerli, H.: Downscaling of air pollutants in Europe using uEMEP_v6, Geosci. Model Dev., 15, 449–465, https://doi.org/10.5194/gmd-15-449-2022, 2022.

OpenStreetMap: Planet OSM, 2024.

Orru, H., Raza, W., Forastiere, F., Spadaro, J., Olstrup, H., Dragic, N., Radic, I., Harhaji, S., Bijelovic, S., Modig, L., Holland, M., Andres, S., Troise, A., Guion, A., Gressent, A., Schullehner, J., and Sommar, J. N.: A Review of the Evidence of the Toxicity of Chemical Substances Included in the European Union Ambient Air Quality and Drinking Water Directives: Perspectives for Health Impact Assessments, Environ. Health, https://doi.org/10.1021/envhealth.4c00277, 2025.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 2825–2830, 2011.

https://doi.org/10.5194/egusphere-2026-1109
Preprint. Discussion started: 6 March 2026

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J. H., Carioli, A., Corbane, C., Dijkstra, L., Florio, P., Friedrich, H. K., Gao, J., Leyk, S., Lu, L., Maffenini, L., Mari-Rivero, I., Melchiorri, M., Syrris, V., Van Den Hoek, J., and Kemper, T.: Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data, Int. J. Digit. Earth, 17, https://doi.org/10.1080/17538947.2024.2390454, 2024.

Peuch, V.-H., Engelen, R., Rixen, M., Dee, D., Flemming, J., Suttie, M., Ades, M., Agustí-Panareda, A., Ananasso, C., Andersson, E., Armstrong, D., Barré, J., Bousserez, N., Dominguez, J. J., Garrigues, S., Inness, A., Jones, L., Kipling, Z., Letertre-Danczak, J., Parrington, M., Razinger, M., Ribas, R., Vermoote, S., Yang, X., Simmons, A., Garcés De Marcilla, J., and Thépaut, J.-N.: The Copernicus Atmosphere Monitoring Service: From Research to Operations, Bull. Am. Meteorol. Soc., 103, E2650–E2668, https://doi.org/10.1175/BAMS-D-21-0314.1, 2022.

Real, E., Couvidat, F., Ung, A., Malherbe, L., Raux, B., Gressent, A., and Colette, A.: Historical reconstruction of background air pollution over France for 2000–2015, Earth Syst. Sci. Data, 14, 2419–2443, https://doi.org/10.5194/essd-14-2419-2022, 2022.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by back-propagating errors, Nature, 323, 533–536, https://doi.org/10.1038/323533a0, 1986.

Ryan, P. H. and LeMasters, G. K.: A Review of Land-use Regression Models for Characterizing Intraurban Air Pollution Exposure, Inhal. Toxicol., 19, 127–133, https://doi.org/10.1080/08958370701495998, 2007.

Safieddine, S., Boynard, A., Coheur, P.-F., Hurtmans, D., Pfister, G., Quennehen, B., Thomas, J. L., Raut, J.-C., Law, K. S., Klimont, Z., Hadji-Lazaro, J., George, M., and Clerbaux, C.: Summertime tropospheric ozone assessment over the Mediterranean region using the thermal infrared IASI/MetOp sounder and the WRF-Chem model, Atmospheric Chem. Phys., 14, 10119–10131, https://doi.org/10.5194/acp-14-10119-2014, 2014.

Schaap, M., Cuvelier, C., Hendriks, C., Bessagnet, B., Baldasano, J. M., Colette, A., Thunis, P., Karam, D., Fagerli, H., Graff, A., Kranenburg, R., Nyiri, A., Pay, M. T., Rouïl, L., Schulz, M., Simpson, D., Stern, R., Terrenoire, E., and Wind, P.: Performance of European chemistry transport models as function of horizontal resolution, Atmos. Environ., 112, 90–105, https://doi.org/10.1016/j.atmosenv.2015.04.003, 2015.

Schimanke, S., Ridal, M., Le Moigne, P., Berggren, L., Undén, P., Randriamampianina, R., Andrea, U., Bazile, E., Bertelsen, A., Brousseau, P., Dahlgren, P., Edvinsson, L., El Said, A., Glinton, M., Hopsch, S., Isaksson, L., Mladek, R., Olsson, E., Verrelle, A., and Wang, Z. Q.: CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.622a565a, 2021.

Shaddick, G., Salter, J. M., Peuch, V.-H., Ruggeri, G., Thomas, M. L., Mudu, P., Tarasova, O., Baklanov, A., and Gumy, S.: Global Air Quality: An Inter-Disciplinary Approach to Exposure Assessment for Burden of Disease Analyses, Atmosphere, 12, 48, https://doi.org/10.3390/atmos12010048, 2020.

Shen, Y., De Hoogh, K., Schmitz, O., Clinton, N., Tuxen-Bettman, K., Brandt, J., Christensen, J. H., Frohn, L. M., Geels, C., Karssenberg, D., Vermeulen, R., and Hoek, G.: Europe-wide air pollution modeling from 2000 to 2019 using geographically weighted regression, Environ. Int., 168, 107485, https://doi.org/10.1016/j.envint.2022.107485, 2022.

Sigsgaard, T. and Hoffmann, B.: Assessing the health burden from air pollution, Science, 384, 33–34, https://doi.org/10.1126/science.abo3801, 2024.

Werbos, P. J.: Backpropagation through time: what it does and how to do it, Proc. IEEE, 78, 1550–1560, https://doi.org/10.1109/5.58337, 1990.

World Health Organization: Ambient air pollution: a global assessment of exposure and burden of disease, Clean Air J., 26, 6, https://doi.org/10.17159/2410-972X/2016/v26n2a4, 2016.

Yang, J., Ji, Z., Kang, S., Zhang, Q., Chen, X., and Lee, S.-Y.: Spatiotemporal variations of air pollutants in western China and their relationship to meteorological factors and emission sources, Environ. Pollut., 254, 112952, https://doi.org/10.1016/j.envpol.2019.07.120, 2019.

Yu, M. and Liu, Q.: Deep learning-based downscaling of tropospheric nitrogen dioxide using ground-level and satellite observations, Sci. Total Environ., 773, 145145, https://doi.org/10.1016/j.scitotenv.2021.145145, 2021.