Environment International 209 (2026) 110143

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Full length article

High-resolution modelling of organic aerosol over Europe: exploring spatial and temporal variability and drivers

Daniel Trejo Banos [a,1], Abhishek Upadhyay [b,1,*], Yun Cheng [a], Jianhui Jiang [c], Petros Vasilakos [b], Andrea Nava [a], Pavol Ševera [a,d], Benjamin Flueckiger [e,f], Aikaterini Bougiatioti [g], Ana Maria Sanchez De La Campa Verdona [h], Andrea Schemmel [i], Andrés Alastuey [j], Anikó Vasanits [k], Anna Font [l,m], Anna Tobler [n], Aude Bourin [l], Attila Machon [o], Benjamin Chazeau [p], Benjamin Bergmans [q], Célia A. Alves [r], Céline Voiron [s], Christoph Hueglin [t], Chunshui Lin [u], Claudio A. Belis [v], Cristina Colombi [w], Cristina Reche [j], Daniel Alejandro Sanchezrodas Navarro [h], Dario Massabò [x], David C. Green [m,y], Eleonora Cuccia [w], Evelyn Freney [z], Fabio Giardi [aa], Francesco Canonaco [n], Gaëlle Uzu [s], Gang I. Chen [m], Hannes Keernik [ab,ac], Harald Flentje [ad], Hartmut Herrmann [ae], Hasna Chebaicheb [l,af], Hilkka Timonen [ag,ah], Hugo Denier van der Gon [ai], Iasonas Stavroulas [g], Imre Salma [k], Jaroslav Schwarz [aj], Jaroslaw Necki [ak], Jean Sciare [al], Jean-Eudes Petit [am], Jean-Luc Jaffrezo [s], Jeni Vasilescu [an], Jesús D. De La Rosa [h], Julija Pauraite [ao], Jurgita Ovadnevaite [u], Karl Espen Yttri [ap], Konstantinos Eleftheriadis [aq], Laurent Poulain [ae], Livio Belegante [an], Lucas Alados-Arboledas [aw], Manousos-Ioannis Manousakas [aq], Marco Paglione [ar], Marek Maasikmets [as], María Cruz Minguillón [j], Maria I. Gini [aq], Matteo Rinaldi [ar], Michael Pikridas [al,at], Minna Aurela [au], Nicolas Marchand [p], Olga Zografou [aq], Olivier Favez [af], Petr Vodička [aj], Petra Pokorná [aj], Radek Lhotka [aj], Samira Atabakhsh [ae], Sébastien Conil [av], Sonia Castillo [aw], Stefania Gilardoni [ax], Stephen M. Platt [ap], Stuart K. Grange [t,ay], Vanes Poluzzi [az], Varun Kumar [ba], Véronique Riffault [l], Wenche Aas [ap], Xavier Querol [j], Yulia Sosedova [n], Nicole Probst-Hensch [e,f], Danielle Vienneau [e,f], André S.H. Prévôt [b], Kees de Hoogh [e,f], Kaspar R. Daellenbach [b,*], Ekaterina Krymova [a,*], Imad El Haddad [b,bb,*]

[a] Swiss Data Science Center, EPFL and ETH Zürich, Zürich, Switzerland
[b] PSI Center for Energy and Environmental Sciences, 5232 Villigen PSI, Switzerland
[c] Global Institute for Urban and Regional Sustainability, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China
[d] Section of Mathematics, University of Geneva, rue du Conseil-Général 7-9, 1205 Geneva, Switzerland
[e] Swiss Tropical and Public Health Institute, Kreuzstrasse 2, 4123 Allschwil, Switzerland
[f] University of Basel, Basel, Switzerland
[g] Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Athens 15236, Greece
[h] CIQSO-Center for Research in Sustainable Chemistry, Associate Unit CSIC-University of Huelva "Atmospheric Pollution", Campus El Carmen s/n, 21071 Huelva, Spain
[i] Umweltbundesamt, Germany
[j] Institute of Environmental Assessment and Water Research (IDAEA-CSIC), 08034 Barcelona, Spain
[k] Institute of Chemistry, Eötvös Loránd University, Budapest, Hungary

* Corresponding authors.
E-mail address: abhishek.upadhyay@psi.ch (A. Upadhyay).

https://doi.org/10.1016/j.envint.2026.110143
Received 29 September 2025; Received in revised form 8 February 2026; Accepted 8 February 2026
Available online 11 February 2026

[l] IMT Nord Europe, Institut Mines-Télécom, Université de Lille, Centre for Energy and Environment, Lille 59000, France
[m] MRC Centre for Environment and Health, Environmental Research Group, Imperial College London, London W12 0BZ, United Kingdom
[n] Datalystica Ltd., Villigen 5234, Switzerland
[o] Air Quality Reference Center, HungaroMet, Budapest, Hungary
[p] Aix-Marseille Université, CNRS, LCE, Marseille, France
[q] ISSEP - Institut Scientifique de Service Public, Liege, Belgium
[r] Department of Environment and Planning, CESAM - Centre for Environmental and Marine Studies, University of Aveiro 3810-193 Aveiro, Portugal
[s] University Grenoble Alpes, CNRS, IRD, INP-G, INRAE, IGE (UMR 5001), 38000 Grenoble, France
[t] Empa, Swiss Federal Laboratories for Materials Science and Technology, 8600 Dübendorf, Switzerland
[u] School of Natural Sciences, Physics, University of Galway, Centre for Climate and Air Pollution Studies, Ryan Institute, Galway H91 CF50, Ireland
[v] European Commission, Joint Research Centre, Ispra 21027, Italy
[w] Environmental Protection Agency of Lombardy (ARPA Lombardia), 20124 Milan, Italy
[x] Department of Physics University of Genoa & INFN-Genoa, Via Dodecaneso 33, 16146, Italy
[y] HPRU in Environmental Exposures and Health, Imperial College London, 86 Wood Lane, London W12 0BZ, the United Kingdom of Great Britain and Northern Ireland
[z] Laboratoire de Météorologie Physique, Université Clermont Auvergne-CNRS, Aubière 63170, France
[aa] National Institute for Nuclear Physics - Florence Section, via Sansone 1, 50019 Sesto Fiorentino (FI), Italy
[ab] Air Quality and Climate Department, Estonian Environmental Research Centre (EERC), Marja 4D, Tallinn, Estonia
[ac] Now at Institute of Physics, University of Tartu, Tartu, Estonia
[ad] German Meteorological Service (DWD), Hohenpeissenberg 82383, Germany
[ae] Leibniz Institute for Tropospheric Research (TROPOS), Leipzig 04 318, Germany
[af] Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil-en-Halatte 60550, France
[ag] Atmospheric Composition Research, Finnish Meteorological Institute, Helsinki 00560, Finland
[ah] Aerosol Physics Laboratory, Tampere University, Tampere 33100, Finland
[ai] TNO, Department of Air Quality and Emissions Research, Utrecht, Netherlands
[aj] Institute of Chemical Process Fundamentals, Czech Academy of Sciences, Rozvojová 1, Prague 6 16500, Czech Republic
[ak] AGH University of Krakow, Faculty of Physics and Applied Computer Science, Krakow 30-059, Poland
[al] Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute, Nicosia 2121, Cyprus
[am] Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Gif-sur-Yvette 91190, France
[an] National Institute of Research and Development for Optoelectronics INOE 2000, Magurele 077125, Romania
[ao] SRI Center for Physical Sciences and Technology (FTMC), Vilnius 10257, Lithuania
[ap] NILU, Kjeller 2007, Norway
[aq] Environmental Radioactivity & Aerosol Technology for Atmospheric & Climate impact Lab, INRaSTES., NCSR "Demokritos", Athens 15310, Greece
[ar] Italian National Research Council - Institute of Atmospheric Sciences and Climate (CNR-ISAC), Bologna 40129, Italy
[as] Estonian Environmental Research Centre, Tallinn, Estonia
[at] Insitute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki 00014, Finland
[au] Atmospheric Composition Research, Finnish Meteorological Institute, Helsinki 00560, Finland
[av] ANDRA, DRD/GES Observatoire Pérenne de l'Environnement, 55290 Bure, France
[aw] Andalusian Institute for Earth System Research (IISTA-CEAMA), Av. del Mediterráneo s/n, 18006 Granada, Spain
[ax] National Research Council, Institute of Polar Sciences (CNR-ISP), Bologna 40129, Italy
[ay] School of Earth and Atmospheric Sciences, Queensland University of Technology, Gardens Point, Brisbane, Queensland 4000, Australia
[az] Arpae Emilia-Romagna, Centro Tematico Regionale Qualità dell'Aria, Bologna 40139, Italy
[ba] Department of Environmental Sciences, Aarhus University, Denmark
[bb] College of Environmental Sciences and Engineering, Peking University, Beijing, China

ARTICLE INFO

Keywords:
Organic aerosol
Chemical transport modelling
CAMX
Machine learning
Random forest
Downscaling
Exposure

ABSTRACT

Organic aerosol (OA) is a major component of atmospheric particulate matter (PM), affecting both human health and climate. However, high-resolution estimates of OA exposure needed for exposure analysis remain scarce. Here, we integrate a chemical transport model (CAMx) with a random forest (RF) machine learning approach to bias-correct and downscale daily OA concentrations across Europe. CAMx OA simulations at ~15 km resolution show moderate agreement with observations (r = 0.55). By combining these outputs with high-resolution land-use data and training the RF model on ~48,000 daily OA measurements from 137 sites, prediction accuracy improved (r = 0.65), with ~l5% reduction in root mean square error. The resulting maps provide European daily OA concentrations at ~250 m resolution for alternate years from 2011 to 2019. The model captures key spatial features, including elevated OA in the Po Valley, Southeastern, and Central Europe, as well as intracity variations due to local hotspots. Seasonal analysis reveals higher concentrations in winter, while long-term trends indicate a general decline in OA levels. Exposure estimates show that half of the European population experiences OA levels above 3 μg/m^3, and ~50 million people are exposed to more than 5 μg/m^3, which is the current guideline level recommended by the world health organization for total $PM_{2.5}$. These high-resolution OA maps offer vital critical support for epidemiological research and air quality policy.

1. Introduction

Atmospheric aerosols have adverse effects on both human health and the environment. These aerosols have varying chemical composition, morphology, and size, which determine their impacts on health and climate. Organic aerosol (OA) contributes 20–90% of particulate matter (PM) mass, while its proportion varies by regions and seasons (Jimenez et al., 2009). In Europe, OA accounts for approximately 50% of total aerosol mass (Bressi et al., 2021), with higher contribution in winter compared to summer (Chen et al., 2022). OA can be directly emitted from combustion activities like biomass burning (BB) and fossil fuel combustion or formed in the atmosphere as secondary aerosol through gas-to-particle conversion (Kanakidou et al., 2005). BB is the largest source of OA in Europe in winter, while biogenic emissions represent a major OA source in summer (Jiang et al., 2019). The contribution of vehicular sources is low at the regional scale, yet much higher at urban locations with high traffic density (Chen et al., 2022). Stricter air quality standards that target specific aerosol components are becoming increasingly important from a regulatory perspective. Given the

[1] **Equally contributing/co-first author.**

different toxicity of various aerosol constituents, it is crucial to prioritize regulations on the most harmful components, such as organic compounds.

Epidemiological studies have firmly established the link between air pollution and a range of adverse health effects, including respiratory, cardiovascular and cerebrovascular diseases. PM alone is estimated to cause 4.2 million premature deaths annually (Cohen et al., 2017). Among PM components, Organic carbon (OC) or OA have been specifically associated with adverse health effects. Elevated OC levels have been linked to increased hospital admissions for myocardial infarction and diabetes (Zanobetti and Schwartz, 2009). Secondary organic aerosol (SOA) shows an even stronger association with cardiorespiratory mortality than total $PM_{2.5}$, with mortality rates up to 6.5 times higher attributed to SOA exposure in the U.S. (Pye et al., 2021). In the UK, studies have also linked carbonaceous particles particularly those from the wood burning to non-accidental and respiratory deaths (Piper et al., 2024). While these findings underscore the health risks of OA, most focus on short-term effects or are limited to the U.S. context. There remains a critical need for large-scale, long-term, high-resolution OA data across other regions to better assess its chronic health impacts.

OA levels are typically estimated using ground-based observations, chemical transport models (CTM), and satellite-derived products. Ground observations offer high accuracy but limited spatial coverage, while satellite products often lack sufficient temporal and chemical resolution, limiting their direct use in exposure assessments and epidemiological studies (El Haddad et al., 2024). Satellite data can provide spatially continuous data but retrieving aerosol composition from satellites remains challenging due to algorithmic limitations, though ongoing efforts are improving these methods (Chen et al., 2024; van Donkelaar et al., 2019). CTMs provide continuous, continent-scale OA fields, yet their accuracy is constrained by uncertainties in input data, emissions, and parameterizations, and their coarse resolution is insufficient to capture local primary emissions. Hybrid approaches that integrate CTM outputs with local land-use information and satellite products have successfully improved high-resolution predictions for regulated and emerging pollutants, including PM, NO_2, O_3, and ultrafine particles (Gariazzo et al., 2020; Jianyao et al., 2025; Li et al., 2024; Shen et al., 2022; Upadhyay et al., 2024). However, extending these approaches to predict specific PM components, such as OA, remains limited. Several studies in North America have generated OA maps at regional scales, demonstrating the potential of these hybrid methods (Chen et al., 2020; Weagle et al., 2018). In Europe, recent applications combining CTMs and machine learning have shown promising results for mapping trace metals like copper, a tracer of brake-wear emissions (Upadhyay et al., 2025). Yet, comprehensive, high-resolution maps of OA across Europe using hybrid modeling approaches are still lacking. In a recent study, we applied a machine-learning framework to estimate OA components (HOA, BBOA, and OOA) at dense OC monitoring networks, achieving prediction accuracies with R^2 values of 0.58–0.75 and expanding the source-apportioned OA database by a factor of four (Jouanny et al., 2025). That approach combined CTM-simulated OA components, land-use variables, and on-site OC concentrations, and was trained on source-apportioned OA from aerosol mass spectrometer (AMS) and aerosol chemical speciation monitor (ACSM) observations. While highly valuable for providing source information at sites where only total OC is routinely measured, the predictions are limited to monitoring locations and do not yield gridded OA fields suitable for large-scale exposure assessment.

In the present study, our goal is to develop spatially continuous, high-resolution, daily total OA maps across Europe for alternate years between 2011 and 2019 (2011, 2013, 2015, 2017, and 2019). To achieve this, we integrate CTM predictions for OA and co-emitted pollutants, and land-use features in a machine learning framework trained on an extensive database of ground-based OC and online OA measurements to predict European OA maps. The resulting OA fields provide continent-wide coverage at a resolution and temporal frequency appropriate for epidemiological and exposure analyses from regional to local scales.

2. Datasets and methodology

2.1. Organic aerosol observation database

We compiled a comprehensive OA observation database to ensure broad spatial and temporal coverage across Europe. Data were sourced from in-house measurements, research collaborators, open-access datasets (e.g., EBAS: https://ebas-data.nilu.no/; last access: December 20, 2025), and federal monitoring agencies. The dataset includes OC measured via offline filter-based thermal-optical analysis, primarily using Sunset Lab instruments and the EUSAAR_2 protocol (Cavalli et al., 2010), as well as OA measured using Aerosol Mass Spectrometers (AMS) and Aerosol Chemical Speciation Monitors (ACSM) (DeCarlo et al., 2006; Fröhlich et al., 2013; Jayne et al., 2000; Ng et al., 2011). We converted OC to OA using a factor of 1.5. Uncertainties associated with this conversion are addressed in Section 2.4.

The final dataset comprises nearly 48,000 daily OA measurements from 137 unique sites across five years (2011, 2013, 2015, 2017, 2019). About 92% of data points are OA estimates converted from OC at 119 sites, while the remaining 8% are direct OA measurements from 18 sites (Fig. 1; site details in Table S1). The dataset spans a wide range of locations, including high- and low-concentration regions as well as urban, suburban, rural, and high-altitude sites making it one of the most comprehensive OA monitoring efforts to date. However, coverage remains limited in certain regions, including Southeastern Europe, remote areas, and high mountain zones. Despite these gaps, the time and spatial resolution of this database represent a key strength of our modeling framework, enabling improved prediction accuracy while also offering a valuable resource for future air quality and health studies.

2.2. Comprehensive air quality model with extensions (CAMx) model

We used the CAMx, version 6.5 (ENVIRON, 2018), to simulate the mass concentrations of major air pollutants across Europe for the years 2011, 2013, 2015, 2017, and 2019. The model domain spans 15°W–35°E longitude and 35°–70°N latitude, with a spatial resolution of 0.125° × 0.25° (latitude × longitude) and 14 vertical layers extending from approximately 20 m to 7000 m above sea level. Gas-phase chemistry was modeled using the Carbon Bond 6 Revision 2 (CB6r2) mechanism (Hildebrandt Ruiz and Yarwood, 2013). The ISORROPIA thermodynamic model was applied to partition inorganic aerosols (sulfate, nitrate, ammonium, sodium, and chloride) between gas and particle phases. OA was modeled using a modified volatility basis set (VBS) module, which distinguishes primary and secondary components from biomass burning, road traffic, other anthropogenic activities, and biogenic sources, estimating primary and secondary biomass burning, hydrocarbon-like, anthropogenic and biogenic OA components. Details on the modified VBS scheme and its parameterization are available in Jiang et al. (2019).

Anthropogenic emissions of major gaseous and particulate pollutants, including POA and precursors of oxygenated OA (OOA), were obtained from the high-resolution European emission inventory CAMS-REG-v4 (Copernicus Atmosphere Monitoring Service–Regional inventory, version 4) developed by TNO (Kuenen et al., 2022). The CAMS-REGv4 version used is referred to as “ref 2”. This version has adjusted particulate matter (PM) emissions for residential combustion to include condensable PM emissions following the approach by Denier van der Gon et al. (2015) and further discussed by Simpson et al. (2022). This inventory includes emissions from public power, industry, transport, residential combustion, agriculture, and other source categories. Non-methane volatile organic compounds (NMVOC) and PM speciation for all emission sources followed the chemical profiles proposed by TNO. Biogenic emissions (isoprene, monoterpenes, and sesquiterpenes), which are critical precursors for SOA, were estimated using the PSI

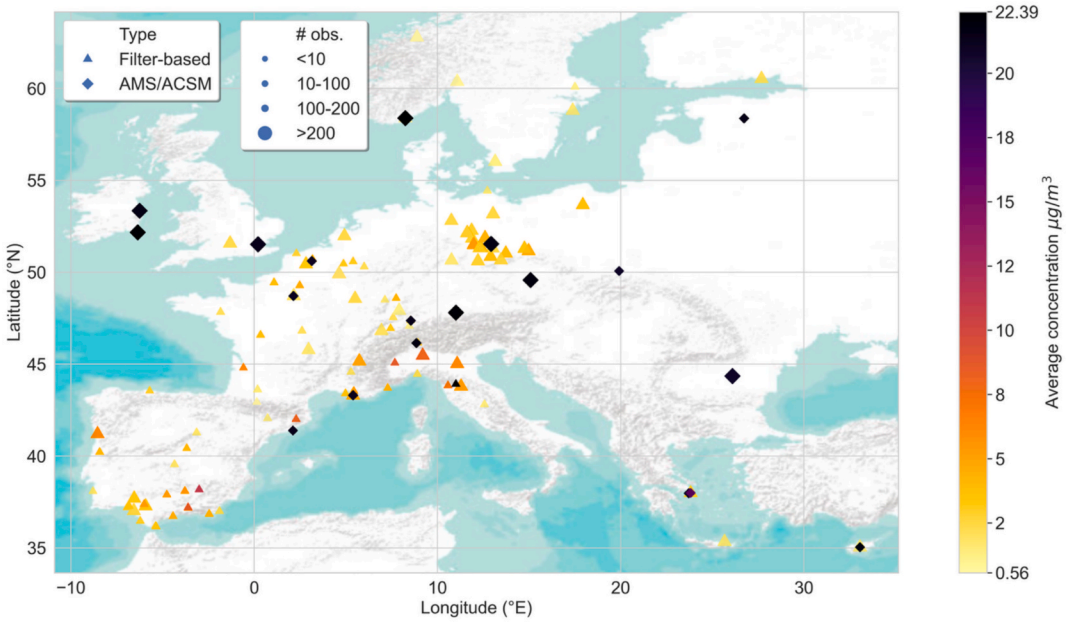

Fig. 1. Locations of organic aerosol (OA) observation sites over Europe. The triangles represent filter-based measurements, and diamonds represent AMS or ACSM (online) measurements. The marker size is proportional to the number of observations available at each location, and the color scale shows the mean OA concentration (μg/m^3) at each location for the whole measurement period. Background map shows land and water areas with terrain. .
Source: © Esri, Source: USGS, Esri, TANA, DeLorne, and NPS

model developed at the Laboratory of Atmospheric Chemistry, Paul Scherrer Institute (Andreani-aksoyoglu and Keller, 1995; Jiang et al., 2019).

Meteorological inputs for the CAMx model were generated using the Weather Research and Forecasting (WRF) model version 3.7.1 (Skamarock et al., 2008) driven by ERA5 reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al., 2011). The WRF setup and parameterization were previously validated for the same modeling domain (Jiang et al., 2019). Chemical initial and boundary conditions were obtained from the global CAM-Chem model (Community Atmosphere Model with Chemistry) and converted to CAMx format using the MOZART2CAMx preprocessor (v3.2.1) (Emmons et al., 2010). Total column ozone data were sourced from NASA's Total Ozone Mapping Spectrometer (TOMS), and photolysis rates were calculated using the Tropospheric Ultraviolet and Visible (TUV) Radiation Model, version 4.8.

2.3. *Machine learning model training*

We used the Random Forest (RF) model (Breiman, 2001), a widely validated tree-based machine learning algorithm, to generate bias-corrected and downscaled OA predictions. We selected RF as the primary modeling approach after testing several algorithms, including gradient boosting (XGBoost) and bagging regressors. RF provided more stable performance across sites, slightly lower variance, and slightly higher predictive accuracy for our dataset. They are also less sensitive to noise, outliers, and heterogeneous feature response relationships, which is advantageous for OA observations collected across different monitoring environments. While gradient boosting is conceptually appealing for bias correction, achieving comparable performance required extensive hyperparameter tuning and substantially increased computational cost. The RF model used our large OA observation database as the target variables and CAMx simulated OA components and other pollutants, meteorological parameters from WRF, and high-resolution land use parameters as predictors. We processed daily OA concentration data to construct a robust target variable, filtering out sites with fewer than 30 time points and observations below 0.1 μg/m^3 (considered below the detection limit). The final dataset for RF modeling included 47,800 daily OA observations from 137 locations, spanning all months, days of the week, and five alternate years between 2011 and 2019, capturing the temporal variability present in the data. The problem is treated as a tabular prediction task; temporal correlations are not explicitly modeled. The list of RF predictors with their description is given in Table S2. The CAMx OA components (HOA, BBOA, PBIS, PAS, PBBS, PRTS) are used individually as predictors in the RF model, together with other covariates. The sum of all components provides total CAMx OA, which is used for validation against observations (Fig. 2a), and not as model input. Land use variables include land use categories (agriculture, industry, transport), road type and length, and population density. Table S2 also provides detailed names and descriptions of all the land use variables used with their source, spatial and temporal resolution, and buffer size, along with the list of other predictors. These land use parameters were obtained from various sources including satellite products and government databases. They were available at resolutions from 25 m to 1000 m for different variables and we re-gridded them to a uniform spatial resolution of ~250 m for modelling purposes. We also used meteorological parameters such as temperature, boundary layer height, wind speed, and atmospheric surface pressure obtained from the WRF model. The CAMx concentration fields set the background for spatial and temporal distributions of OA, while high-resolution land use inputs enabled model downscaling. We treat the problem as a tabular data prediction task, as such we do not explicitly model time correlations.

We optimized the RF model using grid search, systematically testing combinations of hyperparameters to identify those yielding the most accurate predictions. Model evaluation followed a Leave-One-Station-Out Cross-Validation (LOSO-CV) approach, using mean squared error (MSE) as the performance metric. In each iteration, one site was held out as a test set, while 20-fold group cross-validation in the inner loop optimized the RF hyperparameters and assessed performance on unseen locations. Once LOSO-CV predictions were generated, the best hyperparameters were selected using the inner loop with all training data. These were then applied to compute final LOSO predictions for model evaluation. The final model was trained using the optimal hyperparameters and the full training set. To ensure model reliability, we monitored the outer loop's CV error and verified the stability of selected

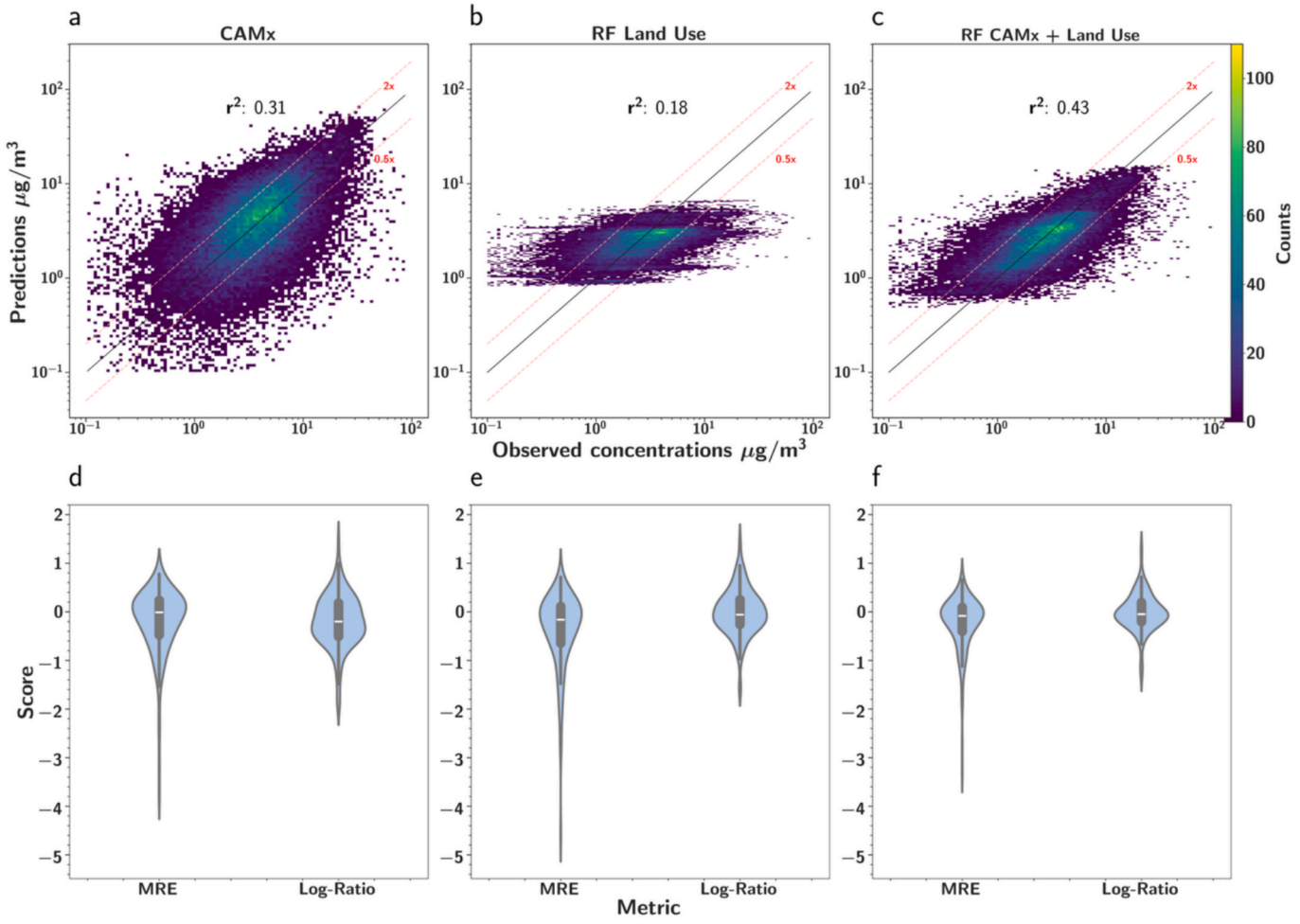

Fig. 2. The range of values of observed and predicted OA was binned to produce a histogram in 2D of the predicted values vs the observed values by the three models (a) CAMx, (b) RF_LU, (c) RF_CAMx_LU, with colors representing the number of total points in each bin. The corresponding station-stratified metrics of the Mean Relative Error (MRE) and log-ratio between observed and predicted values (Log-Ratio) are shown as boxplots in for (d) CAMx, (e) RF_LU and (f) for RF_CAMx_LU.

hyperparameters across iterations.

We trained two versions of the RF model to assess the contribution of land use and CAMx variables to OA predictions. The first model, RF_CAMx_LU, used CAMx outputs, land use, and meteorological data as predictors. The second, RF_LU, excluded CAMx outputs and relied only on land use and meteorological variables. We compared both RF models and the original CAMx predictions against observed OA concentrations. To improve prediction accuracy, we also tested log-transformed versions of the RF models (RF_CAMx_LU log-scale and RF_LU log-scale) to normalize the target variable and perform better for regression tasks whose residuals are homoscedastic in log-scale (Manning and Mullahy, 2001). All RF models were evaluated on the original scale using LOSO cross-validation. Model performance was assessed using station-stratified metrics: Mean Relative Error (MRE), which captures directional bias and is sensitive to concentration magnitude, and the Log-Ratio between observed and predicted values, which is less sensitive to outliers. To interpret the influence of individual predictors, we used permutation importance (Breiman, 2001; Pedregosa et al., 2011) and Shapley Additive Explanations (SHAP) values (Lundberg and Lee, 2017). Permutation importance quantifies the change in model error when feature values are randomly shuffled, while SHAP explains each feature's contribution to specific predictions.

2.4. Uncertainty and limitations

The uncertainties associated with OA modelling in our framework can be broadly categorized into measurement-related and model-related uncertainties.

Measurement uncertainties. We categorized measurement uncertainties into three main sources: (1) offline OC measurements, (2) online OA measurements, and (3) the OA/OC conversion factor used to derive OA from OC.

Offline OC measurements. Uncertainties stem from both sampling and thermal-optical analysis. Sampling-related errors include flow calibration (~3%), positive artefacts from organic vapor adsorption (+10 to +20%) and negative artefacts from OA volatilization (−15 to −5%). Measurement related-errors include calibration accuracy (~3%), detection precision (~2%) and biases associated with the thermal-optical protocol used to separate organic and elemental carbon, with OC measurements using NIOSH and EUSAAR reported to differ by up to 10%. When propagated, these sources yield a total uncertainty of –10% to + 21%, with an average positive bias of ~5% (El Haddad et al., 2011a,b).

Online OA measurements. OA data from ACSM/AMS instruments are affected by detection precision (<1% for 5 μg/m^3 and daily averages), ionization efficiency calibration (~10%), collection efficiency correction (~25%), and the assumed relative ionization efficiency (set to 1.4, but ranges from 1.4 to 1.8). Propagating these errors results in an uncertainty range of –30% to +54%, with an average positive bias of ~12% (El Haddad et al., 2011a,b).

OC to OA conversion. Two factors influence the OC-to-OA conversion: (1) the transmission efficiency of the AMS/ACSM for $PM_{2.5}$,

which is generally high (~80–90%) due to OA being predominantly in the accumulation mode, and (2) the OM/OC ratio, which varies spatially and temporally from 1.5 to 2.0 (variability ~14%), with higher values in rural areas. Considering an average OM/OC ratio of 1.75 and a transmission efficiency of 85%, we calculate an average OA/OC ratio of 1.5 used to convert OC measurements to OA. This ratio is consistent with that determined based on simultaneous online OA and offline OC measurements at few sites. We fitted a model with a conversion factor of 1.8 and compared with the factor 1.5, as detailed in the following section.

Implications. 1. Although most of our OA dataset is derived from converted OC measurements – which dominate the performance of the RF model, values are anchored to AMS/ACSM data through the constant OA/OC ratio of 1.5. If users want to anchor the measurements to offline OA, they should use a factor of 1.75 instead. 2. Using a fixed conversion factor introduces uncertainty to offline OA estimates but also enables a straightforward inference of OC concentrations from OA values. 3. Considering uncertainties from both OC measurements and the OA/OC ratio, we estimate an uncertainty of –19% to +26% for offline-derived OA, compared to –30% to +54% for online OA measurements.

Modelling uncertainties. Errors are primarily associated with the components of the chemical transport model, including errors in meteorological inputs, emission inventories, and the representation of atmospheric chemical and physical processes. CAMx biases are not strictly linear, as they result from multiple interacting sources, including meteorology, emissions, and chemical processes. This non-linearity can increase uncertainty, particularly in regions with sparse observational data where the machine learning model has limited guidance for bias correction. For machine learning when observations fall outside the range covered by the training set, predictions may be biased with unwarranted confidence. Techniques such as conformal prediction (Vovk et al., 2022) may help address these concerns and will be investigated in future work.

3. Results

3.1. Model performance and sensitivity

The observational dataset shows a mean observed OA concentration of 5.0 μg/m^3 across Europe, with a standard deviation of 5.1 μg/m^3, reflecting high regional variability (Fig. 1). Montale (Italy), Kraków (Poland), and Villanueva del Arzobispo (Spain) were the three monitoring locations with the highest mean OA concentrations of 33.6, 18.3, and 16.5 μg/m^3, respectively. We used this observational dataset as the target variable to train the RF model.

We evaluated model performance on a per–left-out-station basis. Table S3 reports the mean performance metrics across all sites, reflecting per-site prediction accuracy. In contrast, Fig. 2 presents the overall metrics calculated from all LOSO predictions combined for each of the models. Both RF based models performed better with log-transformed predictors, and we used those for further analysis and predictions. Fig. 2 compares the performance of CAMx, RF_CAMx_LU log-scale (from now on denoted only as RF_CAMx_LU), and RF_LU on a logarithmic scale through scatter density plots. CAMx achieved a coefficient of determination (R^2) of 0.31 (Pearson's correlation, r = 0.55), capturing the observed value range but exhibiting significant under- and over-predictions at lower concentrations (Fig. 2a). The LOSO predictions for RF_LU (Fig. 2b) had an R^2 = 0.19 (r = 0.43), with values concentrated mainly between 1 and 10, indicating that land-use modeling alone is insufficient to predict regional pollutants like OA. In contrast, RF_CAMx_LU (Fig. 2c) achieved the highest R^2 of 0.43 (r = 0.65), slightly over-predicting lower values but covering the full range of observed OA concentrations. The RMSE decreases from 4.26 μg/m^3 for CAMx to 3.63 μg/m^3 for RF_CAMx_LU, corresponding to a 15% improvement (Table S3). Analysis of MRE and log ratio shows that all models have slightly negative biases. CAMx (Fig. 2d) had a lower median MRE but a wider distribution. In contrast, RF_CAMx_LU (Fig. 2f) had a more constrained MRE distribution. Among the three models, RF_CAMx_LU had the smallest Log-Ratio values, indicating higher accuracy across all observation magnitudes. Additional performance statistics are presented in a per-station stratification in Table S3. The sensitivity analysis of model performance to OA/OC ratios of 1.5 and 1.8 (Table S4) shows only small differences, indicating that the models are weakly sensitive to the choice of OA/OC ratio within this range.

While the land-use model (RF_LU) was capable of high-resolution predictions, it lacks information on pollution sources, chemical reactions, and transport dynamics. By integrating CAMx outputs with land-use variables in RF_CAMx_LU, we leveraged both types of information to produce downscaled predictions of OA levels in Europe. Combining the strengths of both approaches RF_CAMx_LU model showed the highest global R^2 and the lowest median Log-Ratio. Despite slightly higher median MRE than CAMx, it is compensated with less spread of the error bars and outliers for this metric (Fig. 2). We performed Wilcoxon paired tests on the LOSO results, using the null hypothesis that the mean performance metrics of the best-performing model do not differ from those of the other models. The tests confirmed that all previously discussed improvements in overall performance (e.g., R^2, RMSE) are statistically significant. We also applied the Wilcoxon test to model bias (true – predicted), for which no statistically significant differences were found. This indicates that bias alone does not capture the full improvement in predictive skill and highlights the need to consider multiple performance metrics when evaluating model quality.

3.2. Drivers of organic aerosol concentrations

Ambient OA concentrations are shaped by their emissions, precursor sources, physical and chemical transformation, meteorological conditions, and land-use factors. These processes are further emphasized in geographical basins (Salma et al., 2020). The two methods, permutation importance and SHAP, identified biomass burning and other anthropogenic sources as key predictors, with urban fabric, natural green areas, and road networks emerging as the most influential land-use variables (Fig. S1 and S2), it's worth noting that while the results are suggestive, correlations among predictor variables may distort the true importance of these variables (Fig. S3) (Silva and Keller, 2024). We analyzed the relationship between land-use characteristics and high-resolution OA concentrations simulated by the RF_CAMx_LU model using binned boxplots (Fig. 3). OA levels generally decline with increasing natural green cover (Fig. 3a) but rise slightly with larger urbanization (Fig. 3b) and higher density of residential roads (road class 3). Although forests and other green areas contribute to biogenic SOA, our results for the European scale show, these regions tend to have lower OA concentrations than urban areas, with mean OA decreases by ~20% when natural green coverage increases by 50%. The decrease in OA with natural green is steeper in winter season than the summer season, where mean OA decreases by 27% in winter compared to 14% decrease in summer for a 50% increase in green coverage (Fig. S4). In contrast, grid cells with over 30% urban cover or more than 450 m of road class 3 within a 1 km buffer show elevated OA levels. While OA is typically viewed as a regional pollutant, these results highlight a modest but notable local-scale enhancement due to urbanization and road networks. To further evaluate the realism of the downscaled of fields following land-use parameters, we assessed model validation across three urbanization regimes categorized using IMD data. Fig. S6 shows predicted and observed OA comparison as a function of urbanization level and across the full temporal range. The RF_LU_CAMx model captures observed OA concentrations across urbanization levels and temporal scales. These results indicate that the RF_LU_CAMx model resolves sub-grid-scale variability and reproduces key spatial structures across the datasets. The influence of meteorological parameters on OA concentrations over Europe is shown in Fig. S5. OA concentrations increase with temperature, indicating the influence of secondary organic aerosol (SOA) formation (Fig. S5a). However, elevated OA concentrations are also observed at

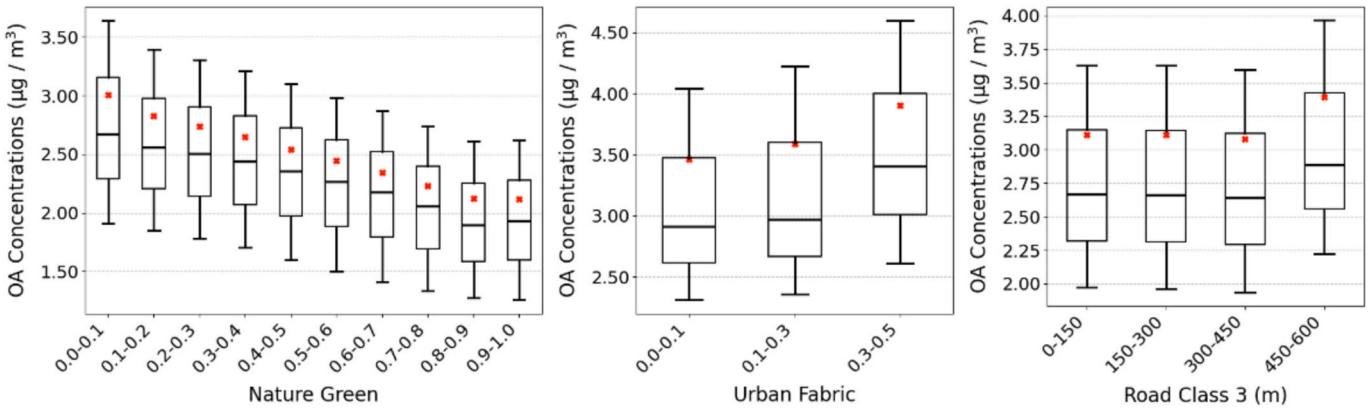

Fig. 3. Box-and-whisker plots showing the relationship between OA concentrations and binned land use parameters: (a) Natural Green, (b) Urban Fabric, and (c) Road Class. Boxes represent the interquartile range (IQR), central lines denote medians, whiskers extend to the 5th and 95th percentiles, and red squares indicate mean OA concentrations for each bin. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lower temperatures, reflecting the impact of biomass burning during winter. Figs. S5(b, c) illustrate that high OA concentrations are associated with low relative humidity (RH) and low planetary boundary layer height (PBLH).

3.3. Spatial distribution of organic aerosol in Europe

We used the RF_CAMx_LU model, trained with the full dataset and optimal hyperparameters, to predict OA concentrations at a fine spatial resolution of 0.005° × 0.0025° (~250 m) across Europe for the years 2011, 2013, 2015, 2017, and 2019. These high-resolution OA predictions over Europe consist of approximately 70 million grids for a single day, totaling 128 billion values for the entire dataset. Fig. 4 illustrates the mean predicted OA concentrations for the selected years across Europe. The results highlight several regions with high concentrations, including the Po Valley, Central Europe (Czechia, southern Poland, Slovakia, and Hungary), and Southeastern Europe (Croatia, Serbia, Bosnia, and Romania). In these areas, yearly mean OA concentrations exceed 10 µg/m^3, which is more than double the revised annual $PM_{2.5}$ guideline level set by the WHO (WHO, 2021). Additionally, highly urbanized areas such as the Benelux region (Belgium, the Netherlands, and Luxembourg), Central Spain, Paris, Warsaw, among others, also showed high OA concentrations (> 4 µg/m^3).

Fig. 5 examines OA distribution across four regions in Europe: the Po Valley, the Iberian Peninsula, Southeastern Europe (Serbia, Hungary, Slovakia, Romania, and Bulgaria), and the Benelux region as predicted by both CAMx and RF_CAMx_LU models. Fig. 5 further illustrates regional differences in OA predictions between RF_CAMx_LU and CAMx. While both models identify major high-concentration zones, CAMx produces broader, less distinct areas of elevated OA, whereas RF_CAMx_LU captures more localized distributions. The refined mapping provided by RF_CAMx_LU highlights the influence of land-use parameters in governing fine-scale spatial variations, with high OA concentrations closely aligning with urbanization and traffic density. For example, in the Benelux region, RF_CAMx_LU shows sharper distinctions between polluted and less polluted areas compared to CAMx, here downscaling enhances urban hotspots while reducing OA in surrounding areas, indicating that CAMx smooths sub-regional variability. Over Eastern Europe, CAMx produces relatively smooth and elevated background OA, whereas RF introduces stronger spatial heterogeneity and generally lower values, particularly in rural areas. Similarly, over Po-Valley RF is enhancing the sub-grid scale heterogeneity. Over the Iberian Peninsula, RF reducing CAMx background concentrations across much of the interior while preserving localized enhancements near urban and coastal regions. These results underscore the importance of high-resolution OA modeling for more accurately characterizing local air quality dynamics and identifying key drivers of OA variability. The difference between OA predicted by RF_CAMx_LU and CAMx over Europe (Fig. S7) indicates that the downscaled and bias-corrected OA generally reduces concentrations in regions where CAMx predicts higher values, while increasing concentrations in areas where CAMx predicts lower values. The zoomed difference maps in Fig. 5 further show that RF_CAMx_LU tends to apply positive corrections along road networks and in smaller urban areas, whereas it applies negative corrections over major cities. In these densely populated urban grids, CAMx's coarse resolution produces uniformly elevated OA concentrations, which are subsequently reduced by the RF_CAMx_LU downscaling.

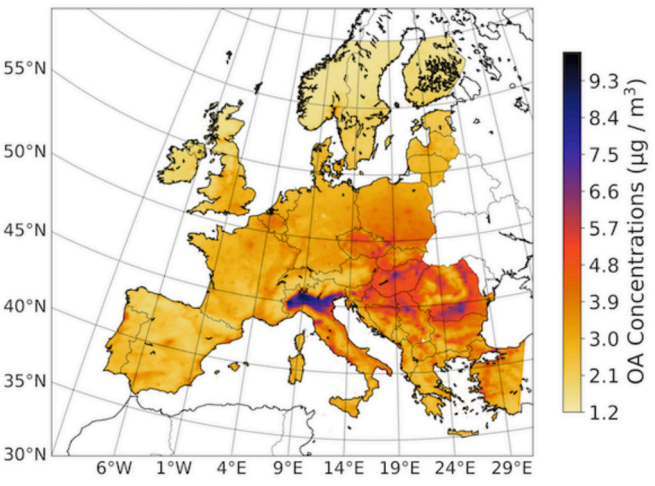

Fig. 4. Average OA concentration (µg/m^3) predicted by RF_CAMx_LU for years 2011, 2013, 2015, 2017, and 2019 at the spatial resolution of 0.0025°×0.005° (~250 m).

3.4. Temporal distribution of OA concentrations

Fig. 6 presents monthly mean OA concentrations across Europe for the simulated years, using RF_CAMx_LU. The maps reveal a distinct seasonal cycle, with higher OA levels in winter and lower concentrations in summer, consistent with previous regional and local studies (Jiang et al., 2019). The seasonal differences are more clearly illustrated by the winter and summer OA means shown in Fig. S8, which highlights a greater regional contrast in concentrations during winter compared to summer. In high-concentration areas like the Po Valley, and Southeastern Europe, OA exceeds 10 µg/m^3 in winter but typically remains below 5 µg/m^3 in summer. This seasonal variation is largely driven by

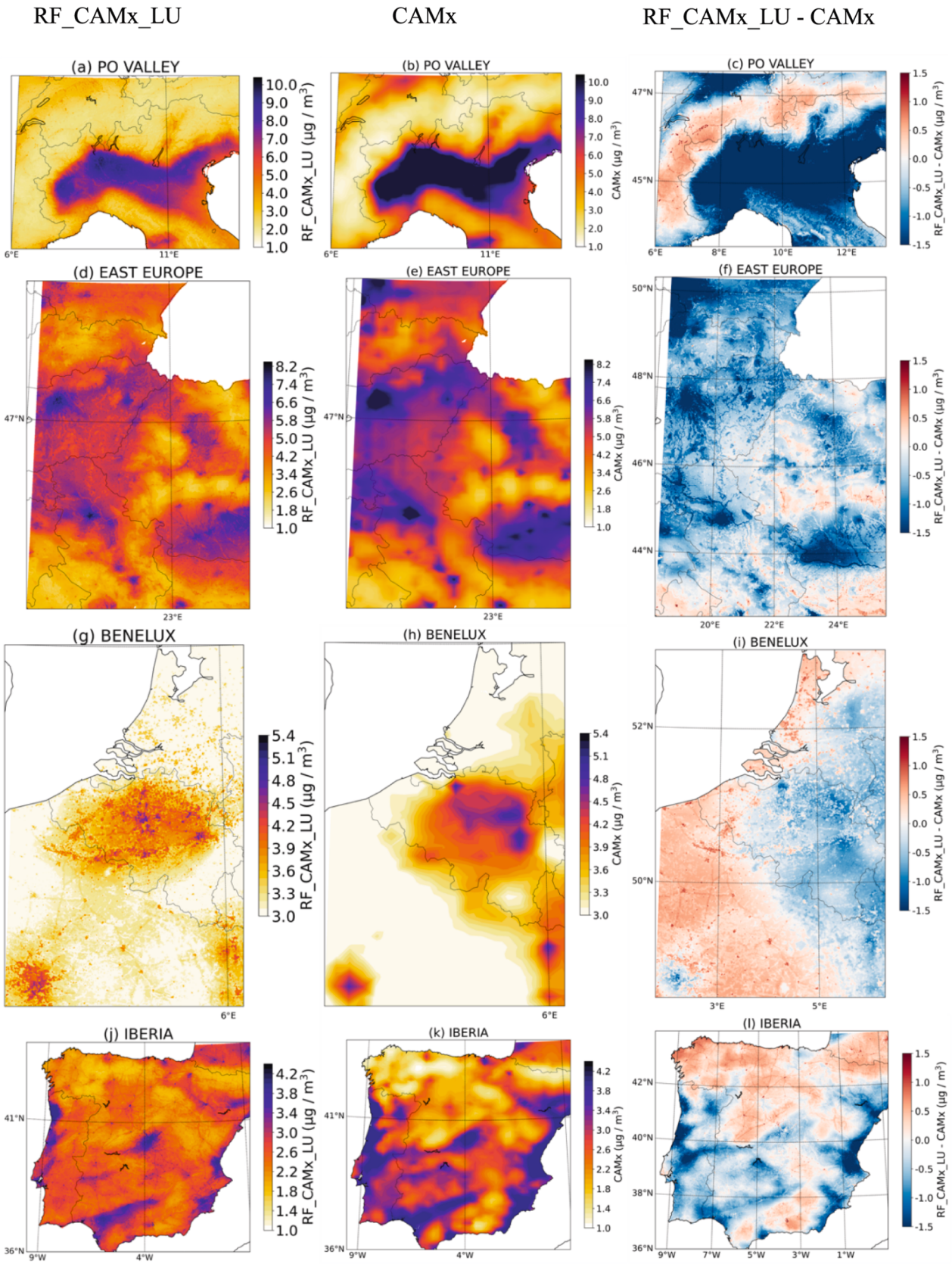

Fig. 5. Yearly mean OA concentrations predicted using RF_CAMx_LU (a, d, g, j), CAMx (b, e, h, k), and their difference (RF_CAMx_LU – CAMx) (c, f, i, l) across four regions: the Po Valley (a–c), Southeastern Europe (Slovakia, Hungary, Serbia, Romania, Bulgaria) (d–f), Benelux (Belgium, the Netherlands, Luxembourg) (g–i), and the Iberian Peninsula (j–l).

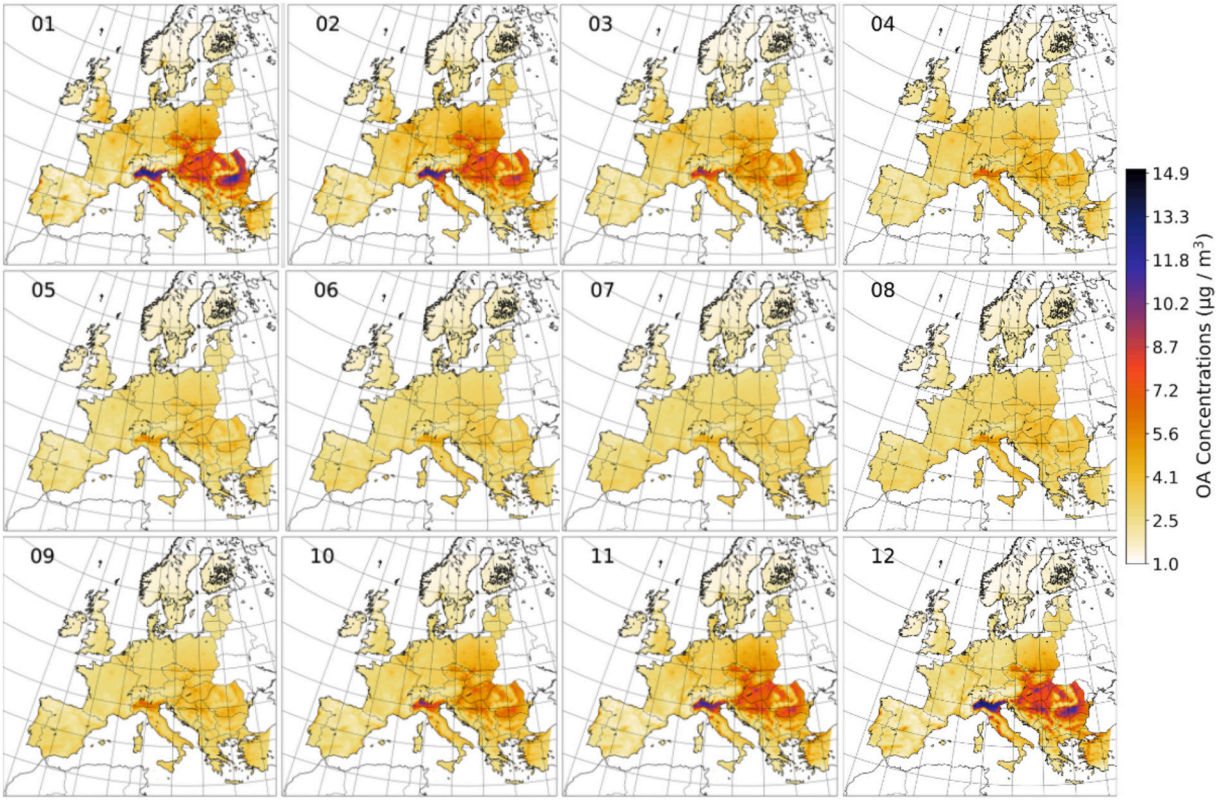

Fig. 6. Mean OA concentrations for each month (labels from 1 to 12 represent months from January to December) averaged over years 2011, 2013, 2015, 2017 and 2019 predicted at ~250 m resolution using the RF_CAMx_LU model.

increased residential biomass burning emissions, and low-dispersion conditions, such as low planetary boundary layer height in winter (Chen et al., 2022; Sandrini et al., 2014). While total OA follows this general trend, individual components may behave differently; For example, biogenic secondary organic aerosol production peaks in summer (Aksoyoglu et al., 2011; Jiang et al., 2019).

Fig. 7 presents temporal trends in OA concentrations across the Po Valley, Southeastern Europe, Benelux, and the Iberian Peninsula. Fig. 7 (a–d) In the Po Valley and Southeastern Europe, OA distributions show a bimodal pattern, indicating two distinct pollution zones. Mean OA concentrations declined from 2011 to 2019, likely reflecting mitigation efforts in residential heating, transport, and fuel combustion. Benelux exhibited a similar but narrower distribution, while the Iberian Peninsula showed a unimodal pattern with occasional peaks up to 12 $\mu g/m^3$. Emission inventories report a 5–10% reduction in OC emissions across Eastern and Western Europe from 2012 to 2023, primarily due to decreased residential emissions (Soulie et al., 2024). In the UK, carbonaceous aerosol levels from residential wood combustion dropped by 2.5–5.5% annually between 2015 and 2021 at three monitoring sites, attributed to the replacement of old fireplaces with more efficient wood stoves (Font et al., 2022).

Monthly variations (Fig. 7e–h, middle row) reveal a winter-dominated bimodal OA distribution in the Po Valley and Southeastern Europe, with lower and more uniform concentrations in summer. Throughout the year, most of the Po Valley and Central Europe experienced OA levels between 2–6 $\mu g/m^3$. However, in winter, concentrations showed a broader range, with significant areas reaching 6–14 $\mu g/m^3$ likely driven by increased local emissions from residential heating and atmospheric stagnation. In contrast, summer concentrations were more evenly distributed, likely due to enhanced atmospheric mixing and regional SOA formation. The winter shift toward higher concentrations in source regions and lower concentrations elsewhere underscores the importance of high-resolution maps in capturing this spatial variability. Benelux and the Iberian Peninsula exhibited similar seasonal patterns, with greater spatial heterogeneity in winter. Notably, Benelux showed wider concentration variability in summer compared to other regions, likely because primary emissions in this region are less linked to residential heating and more to transport-related activities, including ports and airports. Fig. S9 shows monthly OA variation and distribution over whole Europe, here monthly mean markers indicate show winter maxima, early-summer minima, and a stronger secondary peak in late summer. Spread is greatest in winter and late summer, reflecting higher spatial variability during elevated OA periods.

The weekly OA cycle, depicted by a distribution for each day of the week, (Fig. 7i to 7l, bottom row) shows higher concentrations on weekdays, peaking on Wednesdays and Thursdays, with a decline over the weekend. However, this variation is less pronounced than seasonal changes, suggesting that while traffic and industrial activity contribute to OA levels, dominant sources such as biomass burning, and secondary aerosol formation do not follow a strong weekly trend.

The temporal patterns of OA (annual, seasonal, weekly) in the RF_CAMx_LU model largely reflect the underlying CAMx drivers: year-specific emissions control annual trends, meteorology drives seasonal variability, and weekly emission profiles generate weekly cycles. Land-use and population variables vary only at annual resolution and have limited influence on temporal trends. While the temporal dynamics remain similar, the RF model provides bias correction and high-resolution spatial downscaling, which refine the concentration distributions shown in Fig. 7. For comparison, Fig. S10 shows the corresponding distributions from CAMx alone, showing qualitatively similar trends but with different distributions.

3.5. Population exposure to OA

We estimated the European population's exposure to OA using the high-spatial resolution annual mean OA concentrations and gridded

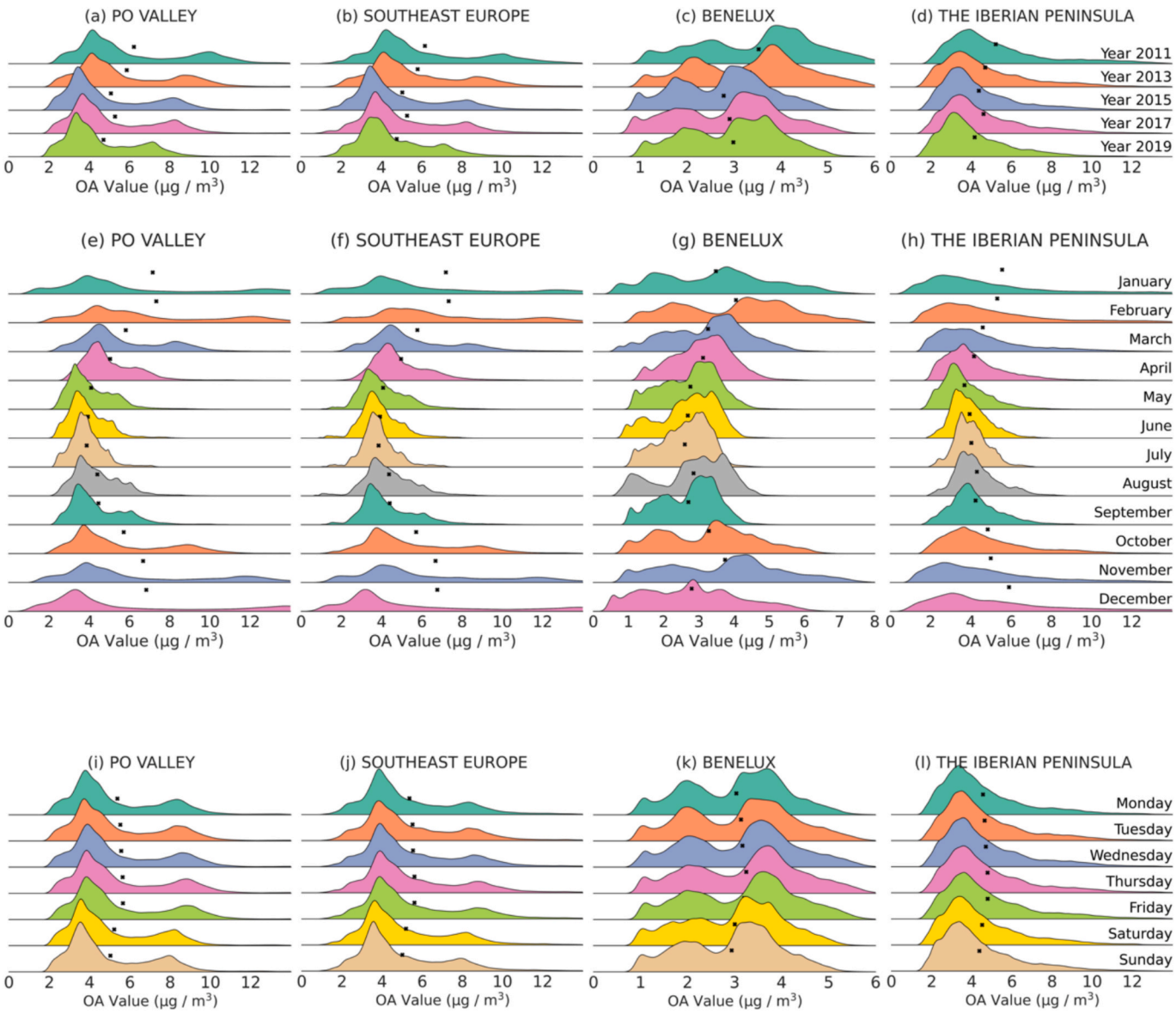

Fig. 7. Distribution of organic aerosol (OA) concentrations using all simulated years for the four regions of interest: (a–d) annual variation in the Po Valley, Southeastern Europe, Benelux, and the Iberian Peninsula, respectively; (e–h) corresponding monthly variations; and (i–l) corresponding weekly (day-of-week) variations for each region. The black dot represents the mean OA concentration.

population data. OA levels were grouped into relevant concentration bins and analyzed alongside population density and total population per bin (Fig. 8). The results show that while ~35% (~190 million) of the European population experienced annual mean OA levels below 3 µg/m^3, 47.65 million people were exposed to concentrations exceeding 5 µg/m^3, which is higher than the revised WHO guideline level for $PM_{2.5}$. Areas with higher population density had elevated OA levels, amplifying health risks. Notably, areas with OA concentrations above 8 µg/m^3 have a median population density exceeding 500 inhabitants/km^2. For reference, the mean population density for the European Union is ~109 inhabitants/km^2 (ranging from ~18 to 1700 inhabitants/km^2 among EU countries). Population exposure derived from the downscaled OA fields is more skewed than that based on CAMx OA (Fig. S11). The high-resolution RF_CAMx_LU fields shift exposure toward lower OA bins and introduce greater variability in population density within the high-OA categories, reflecting the fine-scale urban hotspots that CAMx cannot resolve. Comparing Fig. 8 with Fig. S11 shows that downscaling sharpens spatial contrasts and reduces averaging effects, leading to more realistic population-weighted exposure estimates.

4. Conclusions

The integration of CTM with machine learning, using a uniquely extensive OA observation database with wide spatial and temporal coverage, has significantly enhanced prediction accuracy and resolution of daily OA concentration. This approach reveals key OA hotspots and long-term trends, demonstrating the importance of targeted mitigation measures, such as addressing biomass burning and vehicular emissions. OA concentrations in regions such as the Po Valley and Southeastern Europe often exceed twice the WHO annual $PM_{2.5}$ guideline, underscoring the critical need for higher resolution, source-specific OA maps to meet WHO guidelines. Expanding the observation network to Eastern Europe, improving emission inventories, and refining chemical parameterizations will further enhance predictions. This methodology, applicable to other pollutants such as secondary inorganic aerosols, offers a promising pathway for better understanding regional aerosol pollutants. Generated OA maps are suited for epidemiological studies aiming at understanding the health impacts of OA.

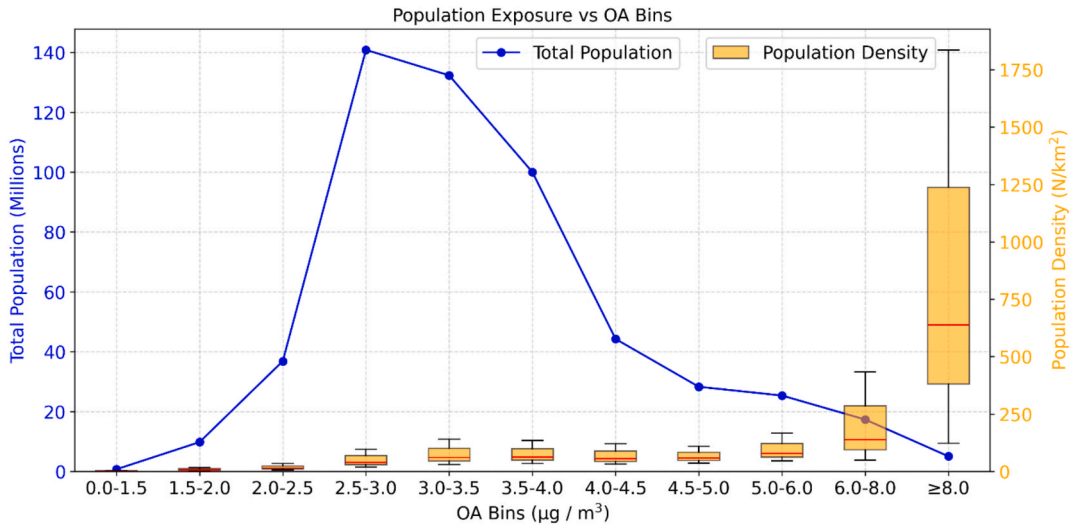

Fig. 8. Total population exposed (blue line) and population density (box-whiskers) for bins of OA concentrations (µg/m^3) over Europe. Orange boxes represent the interquartile range, whiskers represent the 5th–95th percentile, and horizontal red lines denote the median of population density. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

CRediT authorship contribution statement

Daniel Trejo Banos: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **Abhishek Upadhyay:** Methodology, Formal analysis, Data curation, Writing – original draft, Visualization, Software. **Yun Cheng:** Writing – review & editing, Visualization, Methodology, Data curation. **Jianhui Jiang:** Writing – review & editing, Methodology. **Petros Vasilakos:** Writing – review & editing, Formal analysis, Data curation. **Andrea Nava:** Writing – review & editing, Methodology, Formal analysis. **Pavol Ševera:** Writing – review & editing, Methodology. **Benjamin Flueckiger:** Writing – review & editing, Data curation. **Aikaterini Bougiatioti:** Writing – review & editing, Data curation. **Ana Maria Sanchez De La Campa Verdona:** Writing – review & editing, Data curation. **Andrea Schemmel:** Writing – review & editing, Data curation. **Andrés Alastuey:** Writing – review & editing, Data curation. **Anikó Vasanits:** Writing – review & editing, Data curation. **Anna Font:** Writing – review & editing, Formal analysis, Data curation. **Anna Tobler:** Writing – review & editing, Data curation. **Aude Bourin:** Writing – review & editing, Data curation. **Attila Machon:** Writing – review & editing, Data curation. **Benjamin Chazeau:** Writing – review & editing, Data curation. **Benjamin Bergmans:** Writing – review & editing, Data curation. **Célia A. Alves:** Writing – review & editing, Data curation. **Céline Voiron:** Writing – review & editing, Data curation. **Christoph Hueglin:** Writing – review & editing, Data curation. **Chunshui Lin:** Writing – review & editing, Data curation. **Claudio A. Belis:** Writing – review & editing, Data curation. **Cristina Colombi:** Writing – review & editing, Data curation. **Cristina Reche:** Writing – review & editing, Data curation. **Daniel Alejandro Sanchezrodas Navarro:** Writing – review & editing, Data curation. **Dario Massabò:** Writing – review & editing, Data curation. **David C. Green:** Writing – review & editing, Data curation. **Eleonora Cuccia:** Writing – review & editing, Data curation. **Evelyn Freney:** Writing – review & editing, Data curation. **Fabio Giardi:** Writing – review & editing, Data curation. **Francesco Canonaco:** Writing – review & editing, Data curation. **Gaëlle Uzu:** Writing – review & editing, Data curation. **Gang I. Chen:** Writing – review & editing, Data curation. **Hannes Keernik:** Writing – review & editing, Data curation. **Harald Flentje:** Writing – review & editing, Data curation. **Hartmut Herrmann:** Writing – review & editing, Data curation. **Hasna Chebaicheb:** Writing – review & editing, Data curation. **Hilkka Timonen:** Writing – review & editing, Data curation. **Hugo Denier van der Gon:** . **Iasonas Stavroulas:** Writing – review & editing, Data curation. **Imre Salma:** Writing – review & editing, Formal analysis, Data curation. **Jaroslav Schwarz:** Writing – review & editing, Data curation. **Jaroslaw Necki:** Writing – review & editing, Data curation. **Jean Sciare:** Writing – review & editing, Data curation. **Jean-Eudes Petit:** Writing – review & editing, Data curation. **Jean-Luc Jaffrezo:** Writing – review & editing, Formal analysis, Data curation. **Jeni Vasilescu:** Writing – review & editing, Data curation. **Jesús D. De La Rosa:** Writing – review & editing, Data curation. **Julija Pauraite:** Writing – review & editing, Data curation. **Jurgita Ovadnevaite:** Writing – review & editing, Data curation. **Karl Espen Yttri:** Writing – review & editing, Data curation. **Konstantinos Eleftheriadis:** Writing – review & editing, Data curation. **Laurent Poulain:** Writing – review & editing, Data curation. **Livio Belegante:** Writing – review & editing, Data curation. **Lucas Alados-Arboledas:** . **Manousos-Ioannis Manousakas:** Writing – review & editing, Data curation. **Marco Paglione:** Writing – review & editing, Data curation. **Marek Maasikmets:** Writing – review & editing, Data curation. **María Cruz Minguillón:** Writing – review & editing, Data curation. **Maria I. Gini:** Writing – review & editing, Data curation. **Matteo Rinaldi:** Writing – review & editing, Data curation. **Michael Pikridas:** Writing – review & editing, Data curation. **Minna Aurela:** Writing – review & editing, Data curation. **Nicolas Marchand:** Writing – review & editing, Data curation. **Olga Zografou:** Writing – review & editing, Data curation. **Olivier Favez:** Writing – review & editing, Formal analysis, Data curation. **Petr Vodička:** Writing – review & editing, Data curation. **Petra Pokorná:** Writing – review & editing, Data curation. **Radek Lhotka:** Writing – review & editing, Data curation. **Samira Atabakhsh:** Writing – review & editing, Data curation. **Sébastien Conil:** Writing – review & editing, Data curation. **Sonia Castillo:** Writing – review & editing, Data curation. **Stefania Gilardoni:** Writing – review & editing, Data curation. **Stephen M. Platt:** Writing – review & editing, Data curation. **Stuart K. Grange:** Writing – review & editing, Data curation. **Vanes Poluzzi:** Writing – review & editing, Data curation. **Varun Kumar:** Writing – review & editing, Data curation. **Véronique Riffault:** Writing – review & editing, Data curation. **Wenche Aas:** Writing – review & editing, Data curation. **Xavier Querol:** Writing – review & editing, Data curation. **Yulia Sosedova:** Data curation. **Nicole Probst-Hensch:** Writing – review & editing, Formal analysis. **Danielle Vienneau:** Writing – review & editing, Formal analysis. **André S.H. Prévôt:** Writing – review & editing, Formal analysis. **Kees de Hoogh:** . **Kaspar R. Daellenbach:** Writing – review & editing, Funding acquisition, Formal analysis, Data curation. **Ekaterina Krymova:** Writing – original draft, Project administration, Methodology, Formal analysis, Conceptualization. **Imad El**

Haddad: Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge all the support for this work. PSI and SDSC team acknowledge funding from Swiss data science centre (SDSC grant C20-08), the Swiss Federal Office of Environment (FOEN), and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 884104 (PSI-FELLOW-III-3i). Offline data from Swiss sites were sourced from the Swiss National Air Pollution Monitoring Network NABEL, operated by FOEN and Empa. Data from Irish stations were supported by the EU Horizon (grant no. 101081430-PARIS), EPA Ireland and Department of the Environment, Climate and Communications. Data from Czech site were supported by the Ministry of Education, Youth and Sports of the Czech Republic, within the project for support of the national research infrastructure ACTRIS–CZ (LM2023030). Measurements from the Cyprus Atmospheric Observatory were supported financially by the H2020-EMME-CARE (GA 856612) research grants. Measurements from the Athens NOA Air Monitoring Station were supported by the "PANhellenic infrastructure for Atmospheric Composition and climatE change" (MIS 5021516) which is implemented under the Action "Reinforcement of the Research and Innovation Infrastructure", funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund). IDAEA-CSIC acknowledges the funding from Regional Catalan Government (grant no. AGAUR 2021SGR00447). TROPOS measurements were supported by the infrastructure projects ACTRIS (EU FP7; grant no. 262254), ACTRIS-2 (grant no. 654109), and by the German Federal Ministry of Research, Technology and Space (BMFTR) under grant agreements 01LK2002A-G. INOE acknowledge the support of the Core Program within the Romanian National Research Development and Innovation Plan 2022-2027, carried out with the support of MCID, project no. PN 23 05. Samples in France were collected through numerous research and air quality assessment programs, including CARA and MERA (both funded by the Ministry of Environment via LCSQA), DECOMBIO, CAMERA, and QAMECS (all funded by ADEME), ACME and MIAI-Airquality (funded by the University Grenoble Alpes), OPE–Andra (funded by Andra), and multiple initiatives supported by Atmo AuRA, Atmo Sud, Atmo Grand Est, Atmo Hauts de France, Atmo Normandie, IMT Nord Europe, AirBreizh, Atmo Auvergne-Rhône-Alpes, Atmo Grand-Est, Atmo Occitanie, and Lig'Air. We gratefully acknowledge Emmanuel Tison, Nicolas Bonnaire, the many individuals in French AASQA, and staff from several laboratories in France, including IGE and its Air O Sol plateau where a very large fraction of the French samples was analyzed. H. Chebaicheb's PhD was supported by LCSQA (French Ministry of Environment), the Labex CaPPA project (ANR-11-LABX-0005-01), and the CLIMIBIO and ECRIN projects (Hauts-de-France Region & European Regional Development Fund). The ATOLL site, also supported by the Cross-Disciplinary project AREA (R-CDP-24-003-AREA), is part of ACTRIS and contributes to the CARA program of the LCSQA. IMT Nord Europe and INERIS also participated in the COST COLOSSAL Action CA16109. I. Salma acknowledges the support from the Hungarian Research, Development and Innovation Office (contract: Advanced 150835). Ana acknowledgment is given to the Portuguese Foundation for Science and Technology (FCT) for funding UID Centro de Estudos do Ambiente e Mar (CESAM) + LA/P/0094/2020. LCE acknowledges the support of AtmoSud (ToF-ACMS). We acknowledge Dr Anja Tremper and Dr Max Priestman from Imperial College London for their valuable contributions in collecting the ACSM data in London from 2015-2018. Samples in Granada (Spain) were collected through numerous research and air quality assessment programs supported by the Spanish Ministry of Science and Innovation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2026.110143.

Data availability

High-resolution annual mean organic aerosol (OA) concentrations over Europe, predicted using integrated CAMx and random forest modelling, are available on Zenodo at the following https://doi.org/10.5281/zenodo.18516697, for the years 2011, 2013, 2015, 2017, and 2019. The observed OA concentration used to compile the OA database cannot be shared directly; however, it may be obtained upon request from the respective data owners or the corresponding authors.

References

Aksoyoglu, S., Keller, J., Barmpadimos, I., Oderbolz, D., Lanz, V.A., Prévôt, A.S.H., Baltensperger, U., 2011. Aerosol modelling in Europe with a focus on Switzerland during summer and winter episodes. Atmos. Chem. Phys. 11, 7355–7373. https://doi.org/10.5194/acp-11-7355-2011.

Andreani-aksoyoglu, S., Keller, J., 1995. Estimates of monoterpene and isoprene emissions from the forests in Switzerland. J. Atmos. Chem. 20, 71–87. https://doi.org/10.1007/BF01099919.

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

Bressi, M., Cavalli, F., Putaud, J.P., Fröhlich, R., Petit, J.-E., Aas, W., Äijälä, M., Alastuey, A., Allan, J.D., Aurela, M., Berico, M., Bougiatioti, A., Bukowiecki, N., Canonaco, F., Crenn, V., Dusanter, S., Ehn, M., Elsasser, M., Flentje, H., Graf, P., Green, D.C., Heikkinen, L., Hermann, H., Holzinger, R., Hueglin, C., Keernik, H., Kiendler-Scharr, A., Kubelová, L., Lunder, C., Maasikmets, M., Makeš, O., Malaguti, A., Mihalopoulos, N., Nicolas, J.B., O'Dowd, C., Ovadnevaite, J., Petralia, E., Poulain, L., Priestman, M., Riffault, V., Ripoll, A., Schlag, P., Schwarz, J., Sciare, J., Slowik, J., Sosedova, Y., Stavroulas, I., Teinemaa, E., Via, M., Vodička, P., Williams, P.I., Wiedensohler, A., Young, D.E., Zhang, S., Favez, O., Minguillón, M.C., Prevot, A.S.H., 2021. A European aerosol phenomenology - 7: High-time resolution chemical characteristics of submicron particulate matter across Europe. Atmos. Environ.: X 10, 100108. https://doi.org/10.1016/j.aeaoa.2021.100108.

Cavalli, F., Viana, M., Yttri, K.E., Genberg, J., Putaud, J.-P., 2010. Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol. Atmos. Meas. Tech. 3, 79–89. https://doi.org/10.5194/amt-3-79-2010.

Chen, C., Litvinov, P., Dubovik, O., Bindreiter, L., Matar, C., Fuertes, D., Lopatin, A., Lapyonok, T., Lanzinger, V., Hangler, A., Aspetsberger, M., de Graaf, M., Tilstra, L. G., Stammes, P., Dandocsi, A., Gasbarra, D., Fluck, E., Zehner, C., Retscher, C., 2024. Extended aerosol and surface characterization from S5P/TROPOMI with GRASP algorithm. Part II: Global validation and Intercomparison. Remote Sens. Environ. 313, 114374. https://doi.org/10.1016/j.rse.2024.114374.

Chen, G., Canonaco, F., Tobler, A., Aas, W., Alastuey, A., Allan, J., Atabakhsh, S., Aurela, M., Baltensperger, U., Bougiatioti, A., De Brito, J.F., Ceburnis, D., Chazeau, B., Chebaicheb, H., Daellenbach, K.R., Ehn, M., El Haddad, I., Eleftheriadis, K., Favez, O., Flentje, H., Font, A., Fossum, K., Freney, E., Gini, M., Green, D.C., Heikkinen, L., Herrmann, H., Kalogridis, A.-C., Keernik, H., Lhotka, R., Lin, C., Lunder, C., Maasikmets, M., Manousakas, M.I., Marchand, N., Marin, C., Marmureanu, L., Mihalopoulos, N., Močnik, G., Nęcki, J., O'Dowd, C., Ovadnevaite, J., Peter, T., Petit, J.-E., Pikridas, M., Matthew Platt, S., Pokorná, P., Poulain, L., Priestman, M., Riffault, V., Rinaldi, M., Różański, K., Schwarz, J., Sciare, J., Simon, L., Skiba, A., Slowik, J.G., Sosedova, Y., Stavroulas, I., Styszko, K., Teinemaa, E., Timonen, H., Tremper, A., Vasilescu, J., Via, M., Vodička, P., Wiedensohler, A., Zografou, O., Cruz Minguillón, M., Prévôt, A.S.H., 2022. European aerosol phenomenology – 8: Harmonised source apportionment of organic aerosol using 22 Year-long ACSM/AMS datasets. Environ. Int. 166, 107325. https://doi.org/10.1016/j.envint.2022.107325.

Chen, H., Zhang, Z., van Donkelaar, A., Bai, L., Martin, R.V., Lavigne, E., Kwong, J.C., Burnett, R.T., 2020. Understanding the joint impacts of fine particulate matter concentration and composition on the incidence and mortality of cardiovascular disease: a component-adjusted approach. Environ. Sci. Technol. 54, 4388–4399. https://doi.org/10.1021/acs.est.9b06861.

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an

analysis of data from the Global Burden of Diseases Study 2015. Lancet 389, 1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6.
DeCarlo, P.F., Kimmel, J.R., Trimborn, A., Northway, M.J., Jayne, J.T., Aiken, A.C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K.S., Worsnop, D.R., Jimenez, J.L., 2006. Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. Anal. Chem. 78, 8281–8289. https://doi.org/10.1021/ac061249n.
Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorolog. Soc. 137, 553–597. https://doi.org/10.1002/qj.828.
El Haddad, I., Marchand, N., Temime-Roussel, B., Wortham, H., Piot, C., Besombes, J.-L., Baduel, C., Voisin, D., Armengaud, A., Jaffrezo, J.-L., 2011. Insights into the secondary fraction of the organic aerosol in a Mediterranean urban area: Marseille. Atmos. Chem. Phys. 11, 2059–2079. https://doi.org/10.5194/acp-11-2059-2011.
El Haddad, I., Marchand, N., Wortham, H., Piot, C., Besombes, J.-L., Cozic, J., Chauvel, C., Armengaud, A., Robin, D., Jaffrezo, J.-L., 2011. Primary sources of PM $_{2.5}$ organic aerosol in an industrial Mediterranean city, Marseille. Atmos. Chem. Phys. 11, 2039–2058. https://doi.org/10.5194/acp-11-2039-2011.
El Haddad, I., Vienneau, D., Daellenbach, K.R., Modini, R., Slowik, J.G., Upadhyay, A., Vasilakos, P.N., Bell, D., de Hoogh, K., Prevot, A.S.H., 2024. Opinion: how will advances in aerosol science inform our understanding of the health impacts of outdoor particulate pollution? Atmos. Chem. Phys. 24, 11981–12011. https://doi.org/10.5194/acp-24-11981-2024.
Emmons, L.K., Walters, S., Hess, P.G., Lamarque, J.-F., Pfister, G.G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S.L., Kloster, S., 2010. Description and evaluation of the model for ozone and related chemical Tracers, version 4 (MOZART-4). Geosci. Model Dev. 3, 43–67. https://doi.org/10.5194/gmd-3-43-2010.
ENVIRON, 2018. User's guide on Comprehensive Air Quality Model with Extensions (CAMx), version 6.50. Report.
Font, A., Ciupek, K., Butterfield, D., Fuller, G.W., 2022. Long-term trends in particulate matter from wood burning in the United Kingdom: dependence on weather and social factors. Environ. Pollut. 314. https://doi.org/10.1016/j.envpol.2022.120105.
Fröhlich, R., Cubison, M.J., Slowik, J.G., Bukowiecki, N., Prévôt, A.S.H., Baltensperger, U., Schneider, J., Kimmel, J.R., Gonin, M., Rohner, U., Worsnop, D.R., Jayne, J.T., 2013. The ToF-ACSM: a portable aerosol chemical speciation monitor with TOFMS detection. Atmos. Meas. Tech. 6, 3225–3241. https://doi.org/10.5194/amt-6-3225-2013.
Gariazzo, C., Carlino, G., Silibello, C., Renzi, M., Finardi, S., Pepe, N., Radice, P., Forastiere, F., Michelozzi, P., Viegi, G., Stafoggia, M., 2020. A multi-city air pollution population exposure study: combined use of chemical-transport and random-Forest models with dynamic population data. Sci. Total Environ. 724, 138102. https://doi.org/10.1016/j.scitotenv.2020.138102.
Hildebrandt Ruiz, L., Yarwood, G., 2013. Interactions between organic aerosol and NOy: Influence on oxidant production. Final report for AQRP project 12.
Jayne, J.T., Leard, D.C., Zhang, X., Davidovits, P., Smith, K.A., Kolb, C.E., Worsnop, D.R., 2000. Development of an aerosol mass spectrometer for size and composition analysis of submicron particles. Aerosol Sci. Tech. 33, 49–70. https://doi.org/10.1080/027868200410840.
Jiang, J., Aksoyoglu, S., Ciarelli, G., Oikonomakis, E., El-Haddad, I., Canonaco, F., O'Dowd, C., Ovadnevaite, J., Minguillón, M.C., Baltensperger, U., Prévôt, A.S.H., 2019. Effects of two different biogenic emission models on modelled ozone and aerosol concentrations in Europe. Atmos. Chem. Phys. 19, 3747–3768. https://doi.org/10.5194/acp-19-3747-2019.
Jiang, J., Aksoyoglu, S., El-Haddad, I., Ciarelli, G., van der Gon, H.A.C., Canonaco, F., Gilardoni, S., Paglione, M., Minguillón, M.C., Favez, O., Zhang, Y., Marchand, N., Hao, L., Virtanen, A., Florou, K., O'Dowd, C., Ovadnevaite, J., Baltensperger, U., Prévôt, A.S.H., 2019. Sources of organic aerosols in Europe: a modeling study using CAMx with modified volatility basis set scheme. Atmos. Chem. Phys. 19, 15247–15270. https://doi.org/10.5194/acp-19-15247-2019.
Jianyao, Y., Yuan, H., Su, G., Wang, J., Weng, W., Zhang, X., 2025. Machine learning-enhanced high-resolution exposure assessment of ultrafine particles. Nat. Commun. 16, 1209. https://doi.org/10.1038/s41467-025-56581-8.
Jimenez, J.L., Canagaratna, M.R., Donahue, N.M., Prevot, A.S.H., Zhang, Q., Kroll, J.H., DeCarlo, P.F., Allan, J.D., Coe, H., Ng, N.L., Aiken, A.C., Docherty, K.S., Ulbrich, I. M., Grieshop, A.P., Robinson, A.L., Duplissy, J., Smith, J.D., Wilson, K.R., Lanz, V.A., Hueglin, C., Sun, Y.L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J.M., Collins, D.R., Cubison, M.J., E., Dunlea, J., Huffman, J.A., Onasch, T.B., Alfarra, M.R., Williams, P.I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J.Y., Zhang, Y.M., Dzepina, K., Kimmel, J.R., Sueper, D., Jayne, J. T., Herndon, S.C., Trimborn, A.M., Williams, L.R., Wood, E.C., Middlebrook, A.M., Kolb, C.E., Baltensperger, U., Worsnop, D.R., 2009. Evolution of organic aerosols in the atmosphere. Science (1979) 326, 1525–1529. Doi: 10.1126/science.1180353.
Jouanny, A., Upadhyay, A., Jiang, J., Vasilakos, P., Via, M., Cheng, Y., Flueckiger, B., Uzu, G., Jaffrezo, J.-L., Voiron, C., Favez, O., Chebaicheb, H., Bourin, A., Font, A., Riffault, V., Freney, E., Marchand, N., Chazeau, B., Conil, S., Petit, J.-E., de la Rosa, J.D., de la Campa, A.S., Navarro, D.-S.-R., Castillo, S., Alastuey, A., Querol, X., Reche, C., Minguillón, M.C., Maasikmets, M., Keernik, H., Giardi, F., Colombi, C., Cuccia, E., Gilardoni, S., Rinaldi, M., Paglione, M., Poluzzi, V., Massabò, D., Belis, C., Grange, S., Hueglin, C., Canonaco, F., Tobler, A., Timonen, H.J., Aurela, M., Ehn, M., Stavroulas, I., Bougiatioti, A., Eleftheriadis, K., Gini, M.I., Zografou, O., Manousakas, M.-I., Chen, G.I., Green, D.C., Pokorná, P., Vodička, P., Lhotka, R., Schwarz, J., Schemmel, A., Atabakhsh, S., Herrmann, H., Poulain, L., Flentje, H., Heikkinen, L., Kumar, V., Denier van der Gon, H.A., Aas, W., Platt, S.M., Yttri, K.E., Salma, I., Vasanits, A., Bergmans, B., Sosedova, Y., Necki, J., Ovadnevaite, J., Lin, C., Pauraite, J., Pikridas, M., Sciare, J., Vasilescu, J., Belegante, L., Alves, C., Slowik, J. G., Probst-Hensch, N., Vienneau, D., Prévôt, A.S.H., Medbouhi, A.A., Banos, D.T., de Hoogh, K., Daellenbach, K.R., Krymova, E., El Haddad, I., 2025. Machine-learning-driven reconstruction of organic aerosol sources across dense monitoring networks in Europe. Environ. Sci. Technol. Lett. 12, 1523–1531. https://doi.org/10.1021/acs.estlett.5c00771.
Kanakidou, M., Seinfeld, J.H., Pandis, S.N., Barnes, I., Dentener, F.J., Facchini, M.C., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C.J., Swietlicki, E., Putaud, J.P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G.K., Winterhalter, R., Myhre, C.E.L., Tsigaridis, K., Vignati, E., Stephanou, E.G., Wilson, J., 2005. Organic aerosol and global climate modelling: a review. Atmos. Chem. Phys. 5, 1053–1123. https://doi.org/10.5194/acp-5-1053-2005.
Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., van der Gon, H., 2022. CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling. Earth Syst. Sci. Data 14, 491–515. https://doi.org/10.5194/essd-14-491-2022.
Li, Z., Ho, K.-F., Lee, H.F., Yim, S.H.L., 2024. Development of an integrated model framework for multi-air-pollutant exposure assessments in high-density cities. Atmos. Chem. Phys. 24, 649–661. https://doi.org/10.5194/acp-24-649-2024.
Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 2017-Decem, 4766–4775.
Manning, W.G., Mullahy, J., 2001. Estimating log models: to transform or not to transform? J. Health Econ. 20, 461–494. https://doi.org/10.1016/S0167-6296(01)00086-8.
Ng, N.L., Herndon, S.C., Trimborn, A., Canagaratna, M.R., Croteau, P.L., Onasch, T.B., Sueper, D., Worsnop, D.R., Zhang, Q., Sun, Y.L., Jayne, J.T., 2011. An aerosol chemical speciation monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol. Aerosol Sci. Tech. 45, 780–794. https://doi.org/10.1080/02786826.2011.560211.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.
Piper, R., Tremper, A., Katsouyanni, K., Fuller, G.W., Green, D., Font, A., Walton, H., Rivas, I., Evangelopoulos, D., 2024. Associations between short-term exposure to airborne carbonaceous particles and mortality: a time-series study in London during 2010–2019. Environ. Pollut. 360, 124720. https://doi.org/10.1016/j.envpol.2024.124720.
Pye, H.O.T., Ward-Caviness, C.K., Murphy, B.N., Appel, K.W., Seltzer, K.M., 2021. Secondary organic aerosol association with cardiorespiratory disease mortality in the United States. Nat. Commun. 12, 7215. https://doi.org/10.1038/s41467-021-27484-1.
Salma, I., Vasanits-Zsigrai, A., Machon, A., Varga, T., Major, I., Gergely, V., Molnár, M., 2020. Fossil fuel combustion, biomass burning and biogenic sources of fine carbonaceous aerosol in the Carpathian Basin. Atmos. Chem. Phys. 20, 4295–4312. https://doi.org/10.5194/acp-20-4295-2020.
Sandrini, S., Fuzzi, S., Piazzalunga, A., Prati, P., Bonasoni, P., Cavalli, F., Bove, M.C., Calvello, M., Cappelletti, D., Colombi, C., Contini, D., de Gennaro, G., Di Gilio, A., Fermo, P., Ferrero, L., Gianelle, V., Giugliano, M., Ielpo, P., Lonati, G., Marinoni, A., Massabò, D., Molteni, U., Moroni, B., Pavese, G., Perrino, C., Perrone, M.G., Perrone, M.R., Putaud, J.-P., Sargolini, T., Vecchi, R., Gilardoni, S., 2014. Spatial and seasonal variability of carbonaceous aerosol across Italy. Atmos. Environ. 99, 587–598. https://doi.org/10.1016/j.atmosenv.2014.10.032.
Shen, Y., Hoogh, K. De, Schmitz, O., Clinton, N., Tuxen-bettman, K., Christensen, J.H., Frohn, L.M., Geels, C., Karssenberg, D., Vermeulen, R., Hoek, G., 2022. Europe-wide air pollution modeling from 2000 to 2019 using geographically weighted regression 168. Doi: 10.1016/j.envint.2022.107485.
Silva, S.J., Keller, C.A., 2024. Limitations of XAI methods for process-level understanding in the atmospheric sciences. Artificial Intelli. Earth Syst. 3, e230045. https://doi.org/10.1175/AIES-D-23-0045.1.
Simpson, D., Kuenen, J., Fagerli, H., Heinesen, D., Benedictow, A., Denier van der Gon, H., Visschedijk, Antoon Klimont, Z., Aas, W., Lin, Y., Yttri, K.E., Paunu, V.-V., 2022. Revising PM2.5 emissions from residential combustion, 2005–2019.
Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W., Powers, J.G., 2008. Advanced Research WRF Version 3. University Corporation for Atmospheric Research. Doi: Doi: 10.5065/D68S4MVH.
Soulie, A., Granier, C., Darras, S., Zilbermann, N., Doumbia, T., Guevara, M., Jalkanen, J.-P., Keita, S., Liousse, C., Crippa, M., Guizzardi, D., Hoesly, R., Smith, S. J., 2024. Global anthropogenic emissions (CAMS-GLOB-ANT) for the Copernicus Atmosphere monitoring Service simulations of air quality forecasts and reanalyses. Earth Syst. Sci. Data 16, 2261–2279. https://doi.org/10.5194/essd-16-2261-2024.
Upadhyay, A., Jiang, J., Cheng, Y., Vasilakos, P., Chen, Y., Banos, D.T., Flückiger, B., Manousakas, M.I., Prévôt, A.S.H., Modini, R.L., de la Campa, A.S., Schemmel, A., Alastuey, A., Bergmans, B., Alves, C.A., Hueglin, C., Colombi, C., Reche, C., Sánchez-Rodas, D., Massabò, D., Diapouli, E., Mazzei, F., Lucarelli, F., Uzu, G., Salma, I., Jaffrezo, J.-L., de la Rosa, J.D., Reusser, J.E., Eleftheriadis, K., Alleman, L.Y., Scerri, M., Severi, M., Favez, O., Prati, P., Traversi, R., Vecchi, R., Becagli, S., Nava, S., Castillo, S., Darfeuil, S., Grange, S.K., Querol, X., Kertész, Z., Ciarelli, G., Probst-Hensch, N., Vienneau, D., Kuenen, J., Van Der Gon, H.D., Daellenbach, K.R., Krymova, E., de Hoogh, K., El-Haddad, I., 2025. High-resolution modelling of particulate matter chemical composition over Europe: brake wear pollution. Environ. Int. 202, 109615. https://doi.org/10.1016/j.envint.2025.109615.

Upadhyay, A., Sharma, P., Chowdhury, S., 2024. Machine Learning Applications in Air Quality Management and Policies, in: Artificial Intelligence for Air Quality Monitoring and Prediction. CRC Press, p. 18.

van der Gon, H.A.C., Bergström, R., Fountoukis, C., Johansson, C., Pandis, S.N., Simpson, D., Visschedijk, A.J.H., 2015. Particulate emissions from residential wood combustion in Europe – revised estimates and an evaluation. Atmos. Chem. Phys. 15, 6503–6519. https://doi.org/10.5194/acp-15-6503-2015.

van Donkelaar, A., Martin, R.V., Li, C., Burnett, R.T., 2019. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ. Sci. Technol. 53, 2595–2611. https://doi.org/10.1021/acs.est.8b06392.

Vovk, V., Gammerman, A., Shafer, G., 2022. Algorithmic learning in a random world. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-031-06649-8.

Weagle, C.L., Snider, G., Li, C., van Donkelaar, A., Philip, S., Bissonnette, P., Burke, J., Jackson, J., Latimer, R., Stone, E., Abboud, I., Akoshile, C., Anh, N.X., Brook, J.R., Cohen, A., Dong, J., Gibson, M.D., Griffith, D., He, K.B., Holben, B.N., Kahn, R., Keller, C.A., Kim, J.S., Lagrosas, N., Lestari, P., Khian, Y.L., Liu, Y., Marais, E.A., Martins, J.V., Misra, A., Muliane, U., Pratiwi, R., Quel, E.J., Salam, A., Segev, L., Tripathi, S.N., Wang, C., Zhang, Q., Brauer, M., Rudich, Y., Martin, R.V., 2018. Global sources of fine particulate matter: interpretation of PM2.5 chemical composition observed by SPARTAN using a global chemical transport model. Environ. Sci. Technol. 52, 11670–11681. https://doi.org/10.1021/acs.est.8b01658.

WHO, 2021. WHO global air quality guidelines.

Zanobetti, A., Schwartz, J., 2009. The effect of fine and coarse particulate air pollution on mortality: a national analysis. Environ. Health Perspect. 117, 898–903. https://doi.org/10.1289/ehp.0800108.