remote sensing

MDPI

Article

TROPOMI-Based $PM_{2.5}$ Estimates and Their Evaluation During a High-Pollution Event in Germany

Jana Handschuh [1,*], Frank Baier [1], Víctor Molina García [2], Peter Friedl [1] and Diego Loyola [2]

[1] German Aerospace Center (DLR), German Remote Sensing Data Center, 82234 Weßling, Germany; frank.baier@dlr.de (F.B.)
[2] German Aerospace Center (DLR), Remote Sensing Technology Institute, 82234 Weßling, Germany
* Correspondence: jana.handschuh@dlr.de

Highlights

What are the main findings?

- Surface $PM_{2.5}$ concentrations can be accurately estimated using a Random Forest model solely based on TROPOMI satellite observations as predictor variables, without relying on meteorological or chemical transport model data.
- Additional to the commonly used AOD, TROPOMI parameters such as ALH improve model performance for $PM_{2.5}$ estimation.
- The satellite-only model outperforms CAMS forecasts in reproducing spatial patterns, temporal variability and pollution peaks during high-$PM_{2.5}$ episodes.

What are the implications of the main findings?

- Satellite-only, machine learning-based $PM_{2.5}$ products can provide an independent and complementary source of near-real-time air quality information, subject to the availability of satellite observations.
- The method supports operational Copernicus-based $PM_{2.5}$ monitoring and improves surveillance of extreme pollution events.
- The transferability of the method to Sentinel-4 supports future high-temporal-resolution $PM_{2.5}$ products across Europe, relevant for health impact assessment and air quality policy.

Abstract

Fine particulate matter ($PM_{2.5}$) remains one of the most relevant pollutants affecting air quality and human health worldwide. While satellite-derived aerosol optical depth (AOD) is commonly used to estimate surface $PM_{2.5}$ concentrations, most existing approaches rely heavily on auxiliary meteorological model data. This study presents a novel approach that derives $PM_{2.5}$ for Germany and neighboring countries for the year 2022 based on TROPOMI satellite observations by applying a Random Forest (RF) algorithm. In addition to AOD, various TROPOMI products related to atmospheric composition are included to assess their added value for improving model performance. A comparison with CAMS forecasts is performed to demonstrate that the satellite-based model can more realistically reproduce both spatial patterns and temporal dynamics of $PM_{2.5}$. Furthermore, with a case study for March 2022 the model's ability to capture pollution peaks during high-pollution events, which are particularly relevant for public health assessments, is illustrated. The TROPOMI-based RF model achieves high accuracy despite the absence of meteorological input and successfully captures the spatiotemporal variability of $PM_{2.5}$ concentrations. The results of the study highlight the potential of TROPOMI data for near-real-time $PM_{2.5}$ monitoring and underline its value as an independent, observation-based alternative to chemical transport

Check for updates

Academic Editor: Yuan Wang

Received: 18 December 2025
Revised: 25 January 2026
Accepted: 9 February 2026
Published: 11 February 2026

model forecasts. As part of the DLR project INPULS, the proposed approach provides an important step toward the development of an operational daily satellite-based $PM_{2.5}$ product from the atmospheric Copernicus Sentinel missions and contributes to improving air quality surveillance, both under common and extreme pollution conditions.

Keywords: $PM_{2.5}$; air pollution; TROPOMI; Sentinel-5P; machine learning; random forest; Germany

1. Introduction

Fine particulate matter ($PM_{2.5}$), defined as airborne particles with an aerodynamic diameter smaller than 2.5 μm, is well known to cause a range of serious health problems. In particular, $PM_{2.5}$ pollution increases the morbidity and mortality from respiratory and cardiovascular diseases [1] including the three leading causes of death attributable to $PM_{2.5}$: obstructive pulmonary disease, ischemic heart disease and stroke [2].

In Europe, air quality, and in particular $PM_{2.5}$ pollution, is primarily driven by anthropogenic emissions, most notably fossil fuel combustion, residential heating and agricultural activities [3]. However, natural processes can also substantially degrade air quality. Large-scale events such as wildfires, dust storms and stable meteorological conditions periodically contribute to increased particulate concentrations [4]. It is essential to systematically detect, quantify and monitor the population exposure to both anthropogenic and natural sources of fine particulate matter, in order to assess health risks and to inform effective mitigation and adaption strategies, following the WHO Air Quality Guidelines [1].

Ground-based air quality observation sites are sparse and are predominantly installed in or near major urban areas. This hinders an accurate mapping of the spatio-temporal variability of $PM_{2.5}$ pollution and its actual impacts on human health [5]. To obtain a comprehensive picture of population-wide exposure, it is essential to rely on spatially continuous observations of air quality and $PM_{2.5}$ concentrations. Chemical transport models (CTMs) are frequently employed for this purpose, as they provide temporally consistent and spatially continuous datasets. However, in most cases, the spatial resolution is too coarse for the local scale. Also, the accuracy of such models is highly dependent on their input data, such as emission inventories and meteorological fields, which can lead to high biases compared to observations [6]. Moreover, CTMs are based on complex numerical algorithms resulting in high computational costs. Exceptional events, such as Saharan dust intrusions or wildfires, are typically not well captured in near-real time (NRT) model data when no assimilation of ground-based observations and/or actual emissions is performed. Due to the resolution limits, model-based products tend to spatially smooth out fine-scale structures, thereby underestimating local variability [7].

This is where satellite observations become particularly valuable for air quality monitoring. Satellite sensors enable spatially extensive sampling of the atmosphere and provide consistent observations of atmospheric pollutants. However, coverage can be limited by clouds and retrieval restrictions such as the treatment of surface reflectivity [8]. A major challenge in utilizing satellite data for air quality assessment lies in deriving near-surface concentrations from the column-integrated retrievals. This remains an active area of research, requiring synergistic approaches that combine remote sensing, in situ observations and advanced modeling techniques.

In numerous previous studies, remotely sensed Aerosol Optical Depth (AOD) has proven to be a reliable proxy for estimating near-surface $PM_{2.5}$ concentrations [9] and a variety of approaches have therefore been developed (see, e.g., reviews by [10,11]).

Machine learning (ML) algorithms have recently gained prominence due to their flexibility and superior predictive performance. Unlike linear regression or other simple statistical methods, ML models are capable of capturing complex, nonlinear relationships between AOD and $PM_{2.5}$ and other determinators such as meteorological conditions or surface characteristics [12]. Among the most commonly applied algorithms are Random Forest (RF), Neural Networks (NN) and extreme Gradient Boosting (XGB), which have demonstrated strong predictive capabilities across various spatiotemporal contexts [13]. Unlike NNs, often regarded as "black-box" models, as their internal decision-making process is difficult to trace or explain, ensemble-based methods such as RF and XGB offer better interpretability. With analysis tools, e.g., Importance Ranking, their decision-making process is traceable. Therefore, these methods are more suitable when the goal is not only accurate prediction but also an understanding of the underlying drivers of $PM_{2.5}$ variability.

A wide range of satellite-based AOD products is currently available and used worldwide to estimate $PM_{2.5}$ concentrations. The European Copernicus satellite Sentinel-5P, equipped with the TROPOspheric Monitoring Instrument (TROPOMI), was specifically designed to monitor atmospheric gases and aerosols on a daily basis with comparatively high spatial resolution [14]. Within the DLR project INPULS, innovative algorithms and processors are developed to generate global Level-3 TROPOMI products in NRT. In addition to AOD, several aerosol-related parameters as well as high-quality products for other key air pollutants such as NO_2, O_3 and CO are derived. TROPOMI data have been widely applied in air quality studies across the globe [15–19]. However, most research has focused on nitrogen dioxide (NO_2) and on the direct analysis of column-integrated quantities derived from sensor data [20]. Only a limited number of studies have explored the use of TROPOMI observations for estimating near-surface pollutant concentrations [21–23], particularly $PM_{2.5}$ [24–26]. Our previous work demonstrated the potential of TROPOMI-derived AOD as a predictor for PM2.5 concentrations [27]. Nevertheless, additional parameters available from TROPOMI have not yet been integrated into the machine learning framework for $PM_{2.5}$ estimation. In this study, we aim to evaluate the added value of these parameters, with a particular focus on the Aerosol Layer Height (ALH), as an indicator for vertical aerosol distribution, as well as Aerosol Single Scattering Albedo (ASSA) and Aerosol Index (AI). Both ASSA and AI hold important information on aerosol composition [28,29].

In this study we employ a Random Forest approach to derive near-surface $PM_{2.5}$ concentrations across Germany and parts of the surrounding countries for the study year 2022. Several combinations of predictor variables are tested within the RF framework and evaluated with respect to their trade-off between predictive accuracy and computational efficiency. In addition to the commonly used AOD we investigate the potential added value of various additional TROPOMI parameters for accurate $PM_{2.5}$ estimation. A particular challenge is to assess how well TROPOMI-based RF models are able to reproduce the spatio-temporal $PM_{2.5}$ variability during extreme aerosol events. To this end, we analyze a high-pollution episode over Germany in March 2022 as a case study: aerosol surface concentrations were significantly elevated due to a mixture of processes depending on changing emission sources and weather conditions. Model results are evaluated against ground-based in situ measurements and compared to CAMS forecasts. The overarching objective of this research is to contribute to the development of an operational NRT $PM_{2.5}$ product for Germany—and in the long term for Europe—based on atmospheric Sentinel observations and data-driven modeling techniques.

https://doi.org/10.3390/rs18040562

2. Materials and Methods

2.1. Study Area and Period

The study area comprised Germany and parts of the neighboring countries spanning from 46°N to 56°N and 2°E to 16°E. This region is generally characterized by a diverse mixture of anthropogenic and natural emission sources influencing ambient $PM_{2.5}$ concentrations. Major anthropogenic sources include residential heating, road traffic, industrial processes and energy production with notable emission hotspots in metropolitan and industrialized areas such as the Ruhr area, Berlin or Munich. Agricultural activities, particularly ammonia emissions from livestock farming and fertilizer application, play a significant role in secondary aerosol formation, especially during spring [30]. The seasonal variability of $PM_{2.5}$ concentrations reflects both emission patterns and meteorological conditions. Higher concentrations typically occur in winter due to increased heating demand and stable atmospheric stratification, while lower levels are observed in summer owing to enhanced atmospheric mixing and wet deposition. In addition to local and regional sources, transboundary transport contributes to episodic pollution peaks. Among the most prominent natural events affecting $PM_{2.5}$ over Germany are Saharan dust intrusions and wildfires, which can episodically elevate particulate matter concentrations across large parts of the country. Such events occur several times per year and their intensity and extent depend strongly on synoptic-scale atmospheric circulation patterns.

All analyses in this work were based on data from the year 2022. As a case study, we examined an exceptional high-pollution event that occurred over Europe in March 2022. That episode developed in two phases from March 1 to 10 and from March 12 to 30, the latter being the more severe one [31]. It resulted from a combination of different anthropogenic and natural sources. First, the wildfire activity was anomalously high in March 2022. It was concentrated in Eastern and Southern Europe, mainly associated with drought effects, as well as in Ukraine, where it could also be linked to military activity. Furthermore, from March 13 to 18 an intense dust event took place. Large amounts of mineral dust were transported from the Sahara across the Mediterranean into Western Europe and the study domain [32]. The dust was lifted by storm Celia and channeled towards Europe within a warm conveyor belt. A combination of a low-pressure system south-west of the Iberian Peninsula and a high-pressure system over Central Europe further promoted the dust intrusion. Besides these natural sources, Central Europe was strongly affected by anthropogenic emission sources, including residential heating, agricultural emissions and traffic emissions. An extensive high-pressure system over Central Europe in late March, associated with low wind speeds and a reduced boundary layer height, inhibited the dispersion of the emitted pollutants, causing a strong accumulation [33].

2.2. Data

2.2.1. Ground-Based Observations

Ground-based $PM_{2.5}$ measurements were obtained from the European Environment Agency (EEA) Air Quality Download Service "https://eeadmz1-downloads-webapp.azurewebsites.net/" (accessed on 3 December 2025). A total of 648 monitoring stations located within the study area were included in the analysis (see Figure 1). The stations comprised a mixture of urban, suburban and rural background as well as traffic- and industry-related sites, providing spatially representative coverage across Germany and adjacent regions, although the density of monitoring stations was considerably higher in the urban areas.

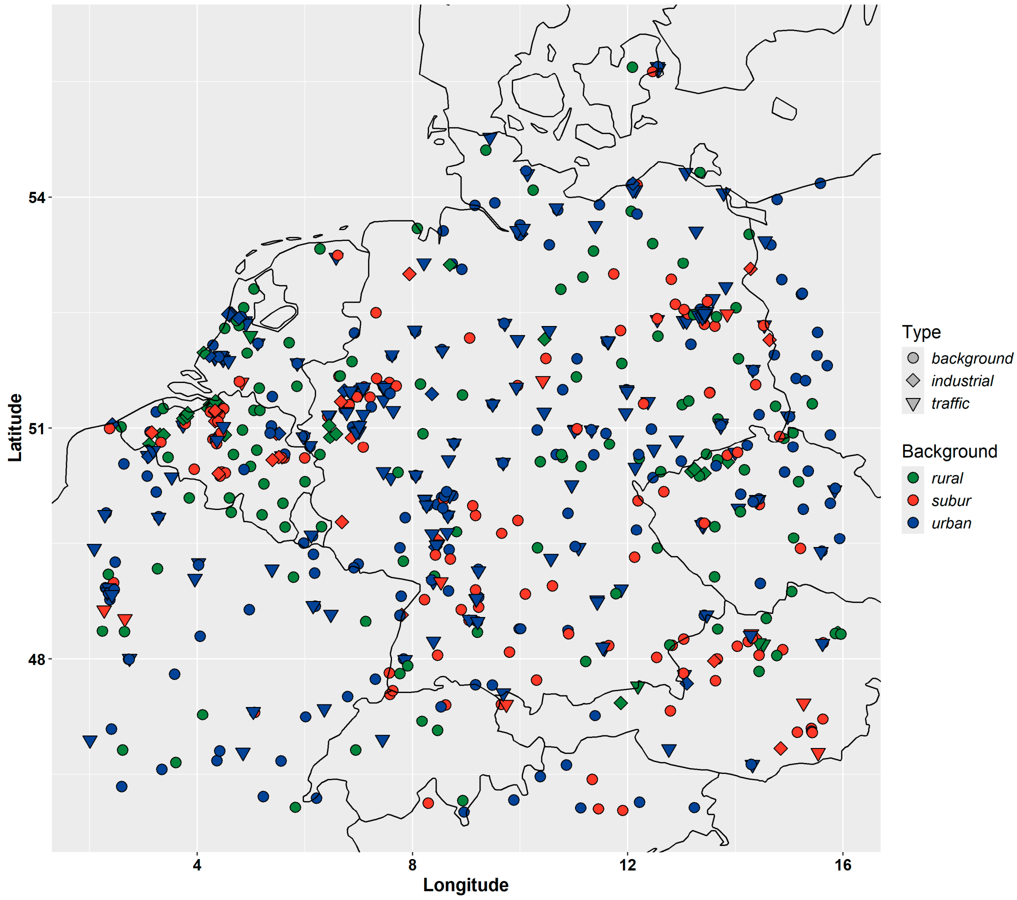

Figure 1. Locations, types and backgrounds of the EEA ground-based $PM_{2.5}$ measurement stations.

As not all stations reported hourly $PM_{2.5}$ concentrations, daily mean values were used for consistency across the dataset. For stations providing hourly data, daily averages were calculated only if at least six hourly measurements were available for a given day. Otherwise, the daily value for that station was classified as a missing observation. This procedure ensured the comparability and temporal consistency of the ground-based dataset used as input for the RF model and for validation of the results.

2.2.2. Satellite Observations

Satellite observations were obtained from the TROPOspheric Monitoring Instrument (TROPOMI; manufacturer Airbus DS-NL, Leiden, Netherlands) aboard the Sentinel-5 Precursor (S5P) polar-orbiting satellite, operated under the European Copernicus Program [14]. TROPOMI is a nadir-viewing imaging spectrometer measuring Top of Atmosphere (TOA) solar radiation reflected by and radiated from the Earth and its atmosphere. It covers spectral bands from the ultraviolet (UV) to the visible (VIS), near-infrared (NIR) and short-wave infrared (SWIR) covering wavelengths from 270 nm to 2385 nm [20]. The sensor has a swath width of 2600 km on the ground and retrieves data at a spatial resolution of ~5.5 km × 3.5 km with daily global coverage. The local equator crossing time is 13:30.

A range of atmospheric variables, including trace gas and aerosol parameters, can be retrieved from TROPOMI measurements. Within the DLR project INPULS, a continuous stream of Level-1B TROPOMI radiance products (band 3 and band 5, NRT) is automatically processed to Level-2 AI, ALH, AOD and ASSA using the TROPOMI aerosol algorithm TropOMAER [29]. Together with ten additional Level-2 NRT trace gas and cloud products, these Level-2 aerosol products are then continuously processed into daily regridded Level-3 NRT products using DLR's Universal Mapper for Atmospheric Spectrometers (UMAS).

While all NRT processing is performed in the cloud using DLR's Processing System for Earth Observation Data (prosEO; "https://github.com/dlr-eoc/prosEO" (accessed on 21 November 2025)), frequent reprocessing campaigns of the entire S5P Level-2 data stack are conducted with more precise offline data and the latest version of UMAS on DLR's High-Performance Data Analytics (HPDA) platform terrabyte. The spatial coverage of individual features varies depending on retrieval sensitivity, cloud contamination and surface reflectance conditions.

For this study, ten different reprocessed daily Level-3 offline products were used (see Table 1). They provide complementary information on aerosol load, optical properties and atmospheric composition, potentially offering valuable insight into the spatial and temporal variability of particulate matter and its precursors. The data were filtered using the quality flags provided within each dataset to exclude low-quality retrievals and ensure consistency. Although a quality flag threshold of 0.5 is commonly recommended, applying this value substantially reduced spatial coverage, particularly for the aerosol-related parameters. Sensitivity tests with lower thresholds showed that values below 0.4 introduced visible artefacts, especially near cloud edges. Therefore, a quality flag threshold of 0.4 was adopted as a compromise between maintaining sufficient spatial coverage and ensuring acceptable quality across the filtered datasets.

Table 1. TROPOMI parameters considered as predictor variables.

Acronym	Name	Unit
AI	Aerosol Index	1
ALH	Aerosol Layer Height	km
AOD550 nm	Aerosol Optical Depth at 550 nm	1
ASSA550 nm	Aerosol Single Scattering Albedo at 550 nm	1
CH_4	Methane	ppm
CO	Carbon Monoxide (total column)	mol/m^2
HCHO	Formaldehyde (total column)	mol/m^2
NO_2	Nitrogen Dioxide (total column)	mol/m^2
O_3	Ozone (total column)	DU
SO_2	Sulfur Dioxide (total column)	DU
UVI	Ultraviolet Index	1

Additional satellite products were considered as predictor variables in this study but not included in the final model setups. These were based on observations of the Moderate Resolution Imaging Spectroradiometer (MODIS; manufactured by Santa Barbara Remote Sensing (SBRS), Santa Barbara, CA, USA) and provided via the NASA data hub LAADS DAAC in approximately 1 km spatial resolution. We used the MODIS products MYD11A1/MYD11A2 for land surface temperature (LST), MYD13A3/MOD13A3 for the normalized difference vegetation index (NDVI) and MCD19Q1 for land cover (LC) information.

All products were resampled onto a common regular $0.01^\circ \times 0.01^\circ$ grid using GDAL to provide a consistent spatial analysis framework throughout this study. This resolution was adopted to maintain consistency with previous model developments and represents a match with the finest native resolution among all considered datasets for the model training process. The oversampling does not increase the effective spatial resolution of the input data, which remains limited by the native resolution of each dataset.

2.2.3. Meteorological Data

Meteorological variables were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) deterministic Atmospheric Model. It provides daily high-resolution 10-day forecasts (SetI—HRES; see [34]). The original spatial resolution of 0.1°

was resampled to a finer 0.01° grid to ensure spatial consistency. To align with the satellite overpass time, data from the 12:00 UTC forecast step were used for analysis. The specific meteorological parameters used in this study are listed in Table 2. Additional parameters such as visibility (Vis), convective available potential energy (CAPE), direct solar radiation at surface (DSR) and dewpoint temperature (TD2m) were tested but excluded from the model training, as they were not beneficial for model performance.

Table 2. Meteorological predictor variables considered for RF model training.

Acronym	Name	Unit
BLD	Boundary Layer Dissipation	J/m^2
BLH	Boundary Layer Height	m
RH	Relative Humidity	%
SP	Surface Pressure	Pa
SSR	Surface Solar Radiation (SW)	J/m^2
STR	Surface Thermal Radiation (LW)	J/m^2
T2m	Temperature in 2 Meters	K
TCWV	Total Column Water Vapor	kg/m^2
TP	Total Precipitation	m
Wind	Wind Speed	m/s
DIR	Wind Direction	Degree

2.2.4. CAMS Forecasts and Reanalysis

For validation and comparison purposes data from the Copernicus Atmosphere Monitoring Service (CAMS) were used, more precisely, the CAMS Regional European Air Quality Forecasts and Reanalysis datasets [35,36]. These are based on an ensemble of 11 state-of-the-art European air quality models, providing hourly data of pollutant concentrations at a horizontal spatial resolution of $0.1° \times 0.1°$. While the forecasts are available in NRT (<24 h), the reanalyses—in which ground-based observations are assimilated and which undergo additional quality control—are released with a temporal delay of at least one year. For the present study, surface-level $PM_{2.5}$ concentrations were extracted from both datasets for the 12:00 UTC timestep, corresponding to the mean overpass time of the satellite observations. The data were resampled to the common 0.01° grid.

2.3. Method

2.3.1. Random Forest

The Random Forest (RF) algorithm is an ensemble learning method that constructs a large number of decision trees and aggregates their predictions to improve model stability and accuracy [37]. Each tree is trained on a bootstrap sample of the training data, while a random subset of predictor variables is used at each node to reduce correlation among trees and enhance model generalization. In this study, the RF implementation provided by the R package ranger() (version 0.17.0) was used [38] due to its computational efficiency and ability to handle large datasets.

Key hyperparameters include the number of trees (ntree), which controls model stability, and the number of variables randomly selected at each split (mtry), which influences the diversity of trees. Ensemble models such as Random Forests are particularly well suited for atmospheric and environmental applications because they can capture nonlinear relationships and variable interactions of complex data [39]. Compared to other machine learning techniques RF models are relatively robust to noise, multicollinearity, outliers and overfitting while providing interpretable measures like feature importance [40].

2.3.2. Model Training

In this study, an RF approach was employed to predict ground-level $PM_{2.5}$ concentrations for the year 2022. The ground-based measurements served as the target variable for model training, while predictor variables included satellite-derived parameters from TROPOMI as well as meteorological variables from ECMWF. To develop an efficient and accurate model—balancing computational cost and predictive performance—different combinations of predictor variables were systematically tested.

A total of 29 potential predictor variables were evaluated during the model development process. An empirical variable selection was performed using the Mean Decrease Accuracy as feature importance measure, in combination with statistical performance indicators including the coefficient of determination, the Pearson correlation coefficient and the root-mean-square error (see Section 2.3.3). Model performance was systematically evaluated for different combinations of predictor variables. Variables that consistently reduced model performance were eliminated in an initial screening step. In a second step, predictors ranked in the lower third of variable importance were individually assessed for their contribution to model performance. Variables that showed no measurable impact on model accuracy were subsequently removed from the final set of predictors.

In total, more than 100 model configurations were tested and compared in terms of predictive accuracy, computational efficiency and spatial coverage. Only the most relevant configurations are presented and discussed here. The evaluated models are summarized in Table 3. All models include the time-related variables day of year (DoY), month and day of week (DOW) as predictor variables. Meteorological parameters were used as a complete set (see Table 2) in most of the model configurations. To assess the added value of satellite-based information, several models were trained with varying subsets of TROPOMI parameters. This allowed for a quantitative evaluation of the contribution of the parameter groups to the overall model performance. The parameter groups included the meteorological predictors providing baseline information on atmospheric dynamics and conditions, aerosol parameters (AOD, AI, ASSA, ALH) which provide information on aerosol optical properties, composition and vertical distribution, trace gas parameters (NO_2, O_3, CO) that act as precursors or share emission sources with $PM_{2.5}$, as well as the UV Index as a radiation-related parameter (UVI).

Table 3. Parameter configuration and sample size for the considered RF models.

Model	Parameters	Sample Size
M1	Meteorology	192,598
M2	Meteorology + AOD	79,984
M3	Meteorology + AOD + AI + ASSA + ALH	29,459
M4	Meteorology + AOD + AI + ASSA	70,216
M5	Meteorology + AOD + AI + ASSA + ALH + NO_2 + O_3 + CO + UVI	23,444
M6	Meteorology + AOD + AI + ASSA + NO_2 + O_3 + CO + UVI	57,673
M7	AOD + AI + ASSA + NO_2 + O_3 + CO + UVI + ALH	23,444
M8	AOD + AI + ASSA + NO_2 + O_3 + CO + UVI	57,673

Certain TROPOMI parameters were also tested but excluded from the analysis due to either insufficient data coverage or negative impacts on model performance. For example, CH_4 exhibited very limited spatial coverage, lowering the spatial and temporal representativity of the training set, while HCHO and SO_2 consistently led to lower model accuracy.

For all model configurations, identical hyperparameter settings were applied to ensure comparability among results. The number of trees was set to 1000 (ntree = 1000) and the number of variables randomly selected at each split was set to five (mtry = 5). Sensitivity tests were conducted to evaluate the influence of the main hyperparameters on model performance. Therefore, value ranges from 3 to 14 were considered for mtry and from 100 to 3000 for ntree. In general, the model performance propagated to saturation with increasing values for ntree, and variations in ntree > 1000 had negligible impact on model performance. The increase in mtry showed a minor effect on predictive accuracy in the value range between 5 and 7. It is generally advisable to increase mtry when a large number of predictor variables is used. However, given the moderate number of predictors in the models considered here, no significant performance improvement was observed. Therefore, mtry was kept constant across all models to ensure consistent and comparable results.

The model training was based on all available collocations between ground-based $PM_{2.5}$ observations and the corresponding variables from satellite and meteorological data in the year 2022. The number of available collocations (sample size) for each model is given in Table 3. Using the trained RF models, area filling maps of surface $PM_{2.5}$ concentrations were subsequently generated for the study domain. As a side note on performance, the number of predictor variables included in the model showed a strong impact on computational costs during mapping.

To evaluate model performance and spatial consistency a comparison between RF-based $PM_{2.5}$ data and CAMS Regional Air Quality Forecasts was performed. For the map-based analysis, CAMS data were filtered to match the spatial and temporal coverage of the RF estimates, ensuring a consistent basis for comparison. For statistical evaluations, only collocated samples with ground-based observations were considered where both CAMS and RF-based data were available, maintaining comparability across all datasets.

2.3.3. Model Evaluation and Interpretation

Model performance was assessed using a 10-fold cross validation (CV) approach. For this purpose, the data were divided into ten subsets. Then, the model was trained ten times, excluding each of the subsets from the training once to serve as validation set. To assess the model's generalization capability the subset partitioning was done using three different strategies: sample-based, spatial and temporal. In the random approach, samples were randomly assigned to subsets, providing an estimate of overall model performance. In the spatial CV, data were split by stations, allowing us to test the model's ability to predict $PM_{2.5}$ concentrations at independent, unobserved sites. In the temporal approach, the data were divided by days to assess the temporal transferability of the model. The complementary validation strategies provided a comprehensive evaluation of both the spatial and temporal robustness of the RF models.

Performance metrics used in this study included the coefficient of determination (R^2), root-mean-square error (RMSE), mean bias and correlation coefficient (R), which collectively quantified model accuracy, precision and systematic deviations between predicted and observed $PM_{2.5}$ concentrations.

To evaluate variable relevance, permutation-based feature importance scores were extracted from the RF models using the built-in functionality provided by the ranger() R package (version 0.17.0) [38]. These metrics quantify the decrease in model performance when individual predictors are randomly permuted, providing an interpretable measure of their contribution to predictive accuracy. In addition, Partial Dependence Plots (PDPs) were computed to visualize the effect of individual predictors on $PM_{2.5}$ estimates, allowing insight into nonlinear relationships and potential threshold effects. Based on the model prediction data, the R package pdp() (version 0.8.2) was used to produce the PDPs for this study [41].

https://doi.org/10.3390/rs18040562

3. Results

3.1. Model Performances and Interpretation

3.1.1. Statistics

Results of the sample-based, spatial and temporal CV for the eight RF models are summarized in Table 4. All models exhibit a comparable strong performance, with R^2 values from the sample-based CV exceeding 0.8, low RMSE and bias values within a similar range, and high correlations with in situ measurements (R > 0.9). The spatial CV results also indicate comparable, though in some cases slightly lower, predictive performance. This suggests that the models are capable of estimating $PM_{2.5}$ concentrations at previously unobserved locations with satisfactory accuracy. In contrast, the temporal CV results show markedly reduced performance across all models.

Table 4. Validation results for the different RF models using sample-based (random; R), spatial (S) and temporal (T) cross-validation strategies.

Model	**R^2**			**RMSE (μg/m^3)**			**BIAS (μg/m^3)**			**R**		
CV	R	S	T	R	S	T	R	S	T	R	S	T
M1	0.82	0.81	0.58	3.04	3.09	4.65	0.08	0.09	0.06	0.91	0.90	0.76
M2	0.82	0.81	0.62	2.82	2.85	4.09	0.07	0.08	−0.02	0.91	0.90	0.79
M3	0.84	0.84	0.66	3.15	3.17	4.62	0.08	0.08	0.10	0.91	0.92	0.81
M4	0.82	0.82	0.64	2.84	2.88	4.05	0.07	0.06	0.01	0.91	0.90	0.80
M5	0.85	0.85	0.66	3.12	3.17	4.71	0.06	0.08	0.18	0.93	0.92	0.82
M6	0.83	0.82	0.67	2.82	2.86	3.93	0.06	0.05	0.23	0.91	0.91	0.82
M7	0.84	0.83	0.63	3.31	3.35	4.93	0.05	0.06	0.27	0.91	0.91	0.80
M8	0.81	0.80	0.60	3.00	3.02	4.26	0.05	0.03	0.27	0.90	0.90	0.78

As demonstrated by model M1, accurate estimates of $PM_{2.5}$ concentrations can already be obtained using meteorological variables alone. The sample-based CV yielded an R^2 of 0.82. However, the temporal CV results are notably poorer compared to the model which additionally includes AOD as an input variable. This suggests that AOD provides valuable information on the temporal variability of $PM_{2.5}$. On the other hand, the impact of AOD on spatial variability is more limited, as spatial CV results do not differ much between models M1 and M2. When further aerosol-related parameters, including Aerosol Layer Height (ALH), are incorporated, a slight improvement in model performance is observed, with the sample-based R^2 increasing to 0.84. Both the spatial and temporal CV results show improvements compared to the previously discussed models. The inclusion of the other aerosol parameters, namely AI and ASSA, has a more pronounced effect both on the spatial and temporal CV results. Although the impact is less significant than that of ALH, their inclusion still provides an additional benefit, indicating that these variables contribute complementary information for $PM_{2.5}$ prediction.

For models M5-M8 several additional TROPOMI parameters were included as predictors together with the aerosol variables. Model M5, which incorporates the most extensive set of input variables in this study (meteorological parameters plus TROPOMI products), demonstrate the highest overall accuracy with R^2 values of 0.85 for the sample-based and spatial CV and 0.66 for the temporal CV. The associated errors with an RMSE of 3.12 μg/m^3 and a bias of 0.06 μg/m^3 are within the same range as those of the other models. The exclusion of ALH as a predictor again leads to a reduction in overall model performance.

Models M7 and M8 were trained exclusively with TROPOMI-derived parameters and exhibit surprisingly high accuracy. Together with model M3, M7 achieves the second-highest performance among all models, while the spatial and temporal CV results are only slightly lower for the purely TROPOMI-based model. The model that includes all

TROPOMI parameters (M7) performs as good as the model combining meteorological and aerosol variables (M3). This finding suggests that the additional TROPOMI variables capture aspects of atmospheric dynamics and chemistry relevant for $PM_{2.5}$ formation, variability and removal. For instance, the UVI reflects solar radiation intensity driving photochemistry, while O_3 amounts are strongly temperature-dependent. Data on other pollutants such as CO or NO_2 can be linked to secondary $PM_{2.5}$ formed in the atmosphere through chemical reactions and it can also yield information about prevailing meteorological conditions, since their chemistry, spatial and temporal variability is influenced by factors such as temperature, humidity, wind or vertical mixing.

The comparison between models M3/M4 as well as M5/M6 and M7/M8 indicates that ALH has a considerable effect on model performance. Excluding ALH reduces model accuracy by 2–4% when taking the results of all CV strategies into account. ALH provides information on the vertical distribution of aerosols within the atmospheric column. Lower ALH values are typically associated with higher near-surface $PM_{2.5}$ concentrations, making it a valuable predictor for the estimation of $PM_{2.5}$ levels. Unfortunately, retrieving ALH from satellite measurements is quite challenging and is therefore affected by significant data gaps. Including ALH as a predictor reduces the sample size for model training by more than half (see Table 3). Consequently, the mean daily spatial coverage of the resulting gridded $PM_{2.5}$ dataset decreases by about 18% (see Table 5). When ALH is included, not only does the overall sample size get smaller, but the available samples are also much more unevenly distributed over space and time. The strong seasonal variability in sample availability is a general issue across all models (except M1). In particular, the winter months are severely underrepresented with respect to data coverage ranging from 0% to a maximum of 6%. In the case of ALH, large data gaps are also present during summer months, as shown in Table 5, leading to an underrepresentation of summer conditions in the model.

Table 5. Mean daily coverage in % per month and model.

MONTH	M1	M2	M3	M4	M5	M6	M7	M8
ALL	100	38	13	33	10	28	10	28
JAN	100	6	2	3	1	2	1	2
FEB	100	29	16	25	12	18	13	20
MAR	100	59	46	57	39	54	40	56
APR	100	45	17	40	13	32	13	32
MAY	100	48	7	42	6	37	6	37
JUN	100	51	9	46	8	42	8	42
JUL	100	50	5	45	4	40	4	40
AUG	100	58	8	53	6	44	6	45
SEP	100	44	12	37	10	31	10	31
OCT	100	53	26	45	16	28	16	28
NOV	100	16	5	9	3	4	3	5
DEC	100	0	0	0	0	0	0	0

Since RF models are intended for spatial mapping, achieving comprehensive coverage is a key requirement. Although ALH appears to provide valuable information, the resulting $PM_{2.5}$ dataset for the year 2022 are unfortunately highly incomplete. While model accuracy is important, it remains satisfactory even without ALH. Therefore, for the analyses presented in Section 3.2 we used the $PM_{2.5}$ dataset derived from model M8.

3.1.2. Feature Importance and Partial Dependence

Figure 2 presents the permutation-based feature importance of models M5 and M8, illustrating the relative ranking of TROPOMI parameters when meteorological information is in- or excluded. Overall, the time-dependent predictors DoY and month are the most influential variables across all models, with substantially higher mean decrease in accuracy (MDA) values in M8 compared to M5. This observation aligns with the results of the temporal CV and underscores the critical role of variables reflecting the highly heterogenous temporal variability of $PM_{2.5}$. In contrast, static variables such as land cover (LC) and the Normalized Difference Vegetation Index (NDVI) exhibit only minor contributions to the models (see [27]). It can be expected that these variables would play a much more significant role when capturing temporal variability associated with local emissions on shorter time scales, such as harvesting, ploughing or drought periods.

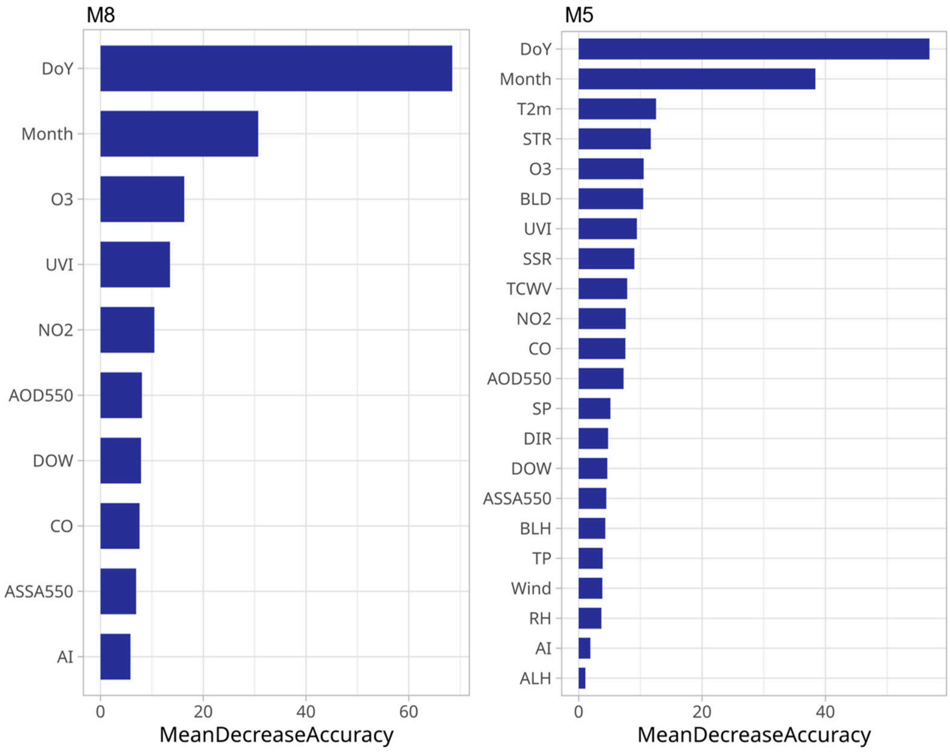

Figure 2. Importance rankings in terms of Mean Decrease in model Accuracy (MDA) for models M5 and M8.

Among meteorological predictors, temperature (T2m), surface longwave radiation (STR) and boundary layer dissipation (BLD) emerge as the most important. The latter reflects the degree of atmospheric mixing within the boundary layer. Greater mixing corresponds to reduced accumulation of $PM_{2.5}$ near the surface. Both temperature and longwave radiation act as indicators of seasonal variation in $PM_{2.5}$ concentrations and may also serve as proxies for conditions favoring the formation of secondary aerosols.

Among the TROPOMI parameters, O_3, UVI and NO_2 rank highest in importance, followed by AOD. Similar to temperature and radiation, UVI may serve as an indicator of seasonal variability in the $PM_{2.5}$ concentrations. While AOD gives direct information on atmospheric particle load, O_3 and NO_2 may serve as indicators of local emissions and secondary aerosol formation.

Interestingly, aerosol-related parameters that directly describe aerosol composition show generally lower importance in the models. This is particularly surprising for the ALH, which has a clear and quantifiable influence on model performance. These findings highlight that feature importance rankings do not necessarily reflect which variables could be neglected without loss of predictive power. Consequently, variable selection based solely on importance rankings, an approach often applied in the literature, proved of limited value in our case.

When meteorological variables are included in the model, predictors such as T2m and STR rank higher in importance than pollutant-related indicators. At the same time, the relative importance of the temporal variables DOY and month decreases slightly, suggesting that these meteorological predictors provide substantial information on the seasonal and temporal variability of $PM_{2.5}$. Among the TROPOMI-based variables, O_3 consistently remains the most influential predictor. Overall, the relative ordering of the TROPOMI parameters remains largely unchanged when meteorological variables are included, although they are shifted to lower ranks due to the presence of certain meteorological predictors. An exception is CO, which ranks slightly higher than AOD in this configuration. However, the importance scores of NO_2, CO and AOD are very close, and these small differences should therefore not be overinterpreted.

In addition to assessing variable importance, it is essential to examine the physical relationships learned by the model, in order to evaluate plausibility of results. Figure 3 displays the PDPs between the estimated $PM_{2.5}$ concentrations and several key predictor variables. Overall, the modeled relationships appear physically reasonable.

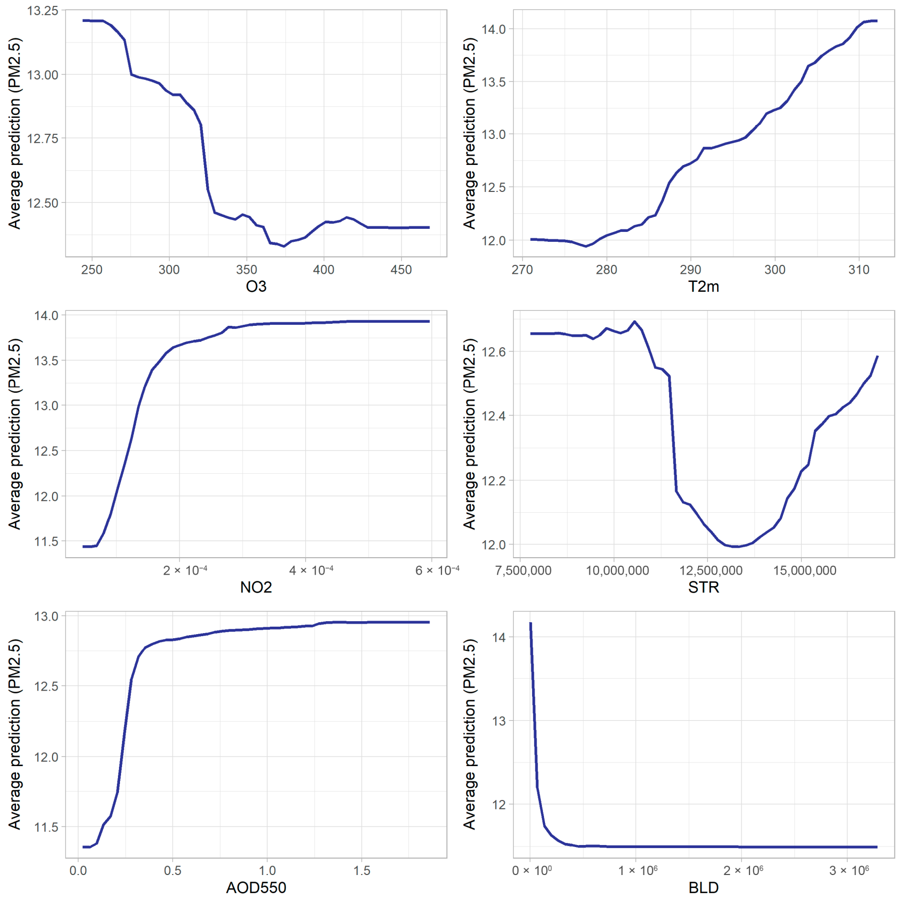

Figure 3. Partial dependence plots for selected predictor variables (O_3, T2m, NO_2, STR, AOD, BLD).

For both NO_2 and AOD, $PM_{2.5}$ concentrations increase with rising NO_2 and AOD values, reflecting similar underlying patterns. These relationships are approximately linear in the lower ranges of NO_2 and AOD but level off beyond a certain threshold, indicating that $PM_{2.5}$ concentrations no longer increase once a specific level of these variables is reached. In contrast, O_3 exhibits an inverse relationship with $PM_{2.5}$. Higher O_3 values correspond to lower $PM_{2.5}$ levels in our model. A threshold effect is also evident here, where $PM_{2.5}$ concentrations stop to decrease over a certain O_3 value.

The model also captures meaningful physical dependencies for the meteorological variables. The relationship between temperature and $PM_{2.5}$ is nearly linear, showing higher $PM_{2.5}$ concentrations at higher temperatures. The dependence between $PM_{2.5}$ and surface longwave radiation is particularly interesting. $PM_{2.5}$ concentrations remain relatively high and constant up to a certain radiation level, after which a sharp decrease occurs, reaching a minimum. After this point, $PM_{2.5}$ increases again with further rising radiation values. This suggests a bidirectional relationship. Higher particle loads may attenuate the downward longwave radiation reaching the surface, while higher radiation levels lead to higher temperatures and may promote aerosol activation. Finally, the relationship between BLD and $PM_{2.5}$ is strong and physically consistent. Only once a certain level of turbulence and vertical mixing is reached does the model predict a reduction in $PM_{2.5}$ concentrations. Under conditions of a shallow or poorly mixed boundary layer, particle accumulation near the surface leads to higher $PM_{2.5}$ levels.

3.2. High Pollution Event (March 2022)

Figure 4 presents the monthly mean values of $PM_{2.5}$ concentrations derived from model M8. Unfortunately, the maps for January and December were rather incomplete due to substantial data gaps in the TROPOMI dataset, which were mainly caused by high cloud cover and poor illumination conditions. Nevertheless, a typical annual pattern of $PM_{2.5}$ concentrations can be observed, with higher concentrations during the cold season and markedly lower values in summer. However, the annual cycle is not very pronounced while the month of March is particularly striking. It exhibits significantly elevated $PM_{2.5}$ concentrations across the entire study area compared to the rest of the year. Almost everywhere, the values approach the legally established limit of 25 $\mu g/m^3$ [42]. In some regions, such as northern France, large parts of Belgium and the Netherlands, Southwestern Germany, as well as Southern Poland and the Czech Republic, this limit is even exceeded in the monthly mean. On average, concentrations in March reach 20 $\mu g/m^3$, which is more than twice as high as during other months of the spring season (see Table 6). The exceptionally high $PM_{2.5}$ concentrations observed in March 2022 can be attributed to a combination of different sources, including residential heating, agricultural emissions, wildfires in Eastern Europe and Sahara dust intrusion [31]. The episode developed in two phases, with the more severe one lasting from March 12 until 30, including an intense dust event from 15 to 18 described in detail by [32]. Large amounts of mineral dust were transported from the Sahara across the Atlantic and Southern Europe into the study area, leading to phenomena such as reddish skies and so-called “blood rain”.

Figure 5 presents time series of $PM_{2.5}$ concentrations for the entire year 2022 (top) and in detail for the month of March (bottom). The values represent averages over all samples that were excluded from training and used for model validation. In addition to the RF-derived $PM_{2.5}$ concentrations (M8), the figure also displays in situ measurements as well as CAMS forecasts and CAMS reanalyses. Besides the relatively high concentrations in January, March clearly stands out within the annual cycle. During that month, concentrations remain above the annual mean value of approximately 10 $\mu g/m^3$ (dashed line in Figure 5) for nearly the entire period, whereas this threshold is exceeded only on a few isolated periods in other months. Overall, the curves of in situ measurements and the RF-derived $PM_{2.5}$ concentrations show very close agreement, with a bias below 0.1 $\mu g/m^3$ and a RMSE of 1.42 $\mu g/m^3$ as indicated in Table 6, demonstrating excellent model performance. The RF model reproduces the temporal evolution of $PM_{2.5}$ concentrations very well, yielding an overall correlation coefficient of 0.9 and accurately capturing the observed peaks. Correlations per month are slightly lower compared to the total set, but the bias and RMSE are constantly low throughout the year.

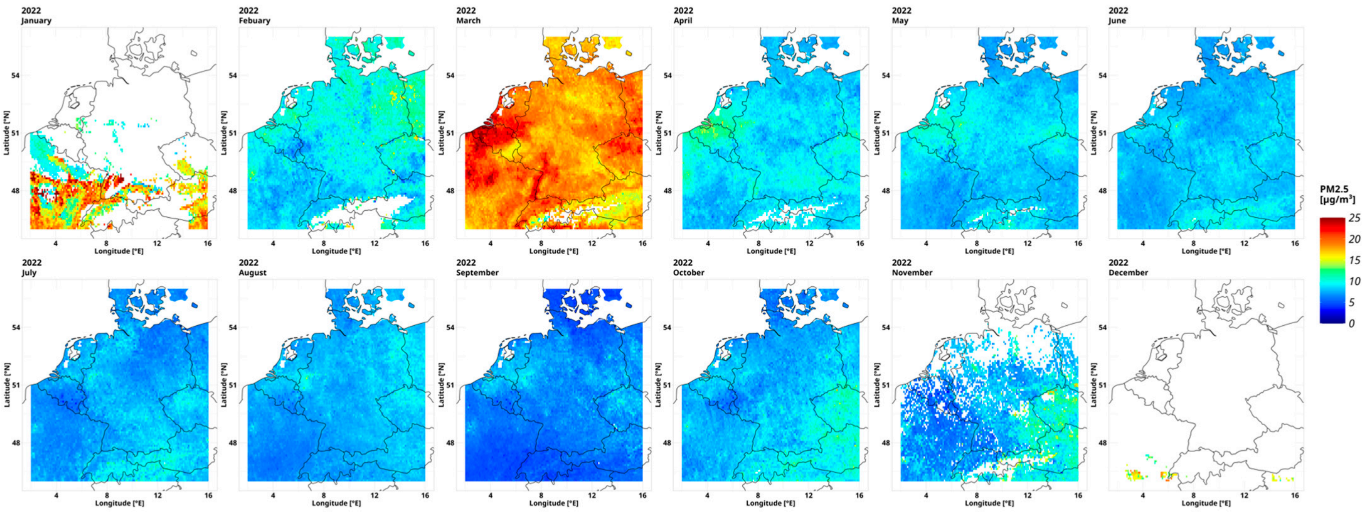

Figure 4. Monthly averages of $PM_{2.5}$ concentrations derived using model M8.

Table 6. Descriptive statistics for the RF model M8, CAMS forecasts and CAMS reanalysis. Average and standard deviation (AVG), bias and RMSE are given in $\mu g/m^3$.

		RF (M8)				CAMS Forecast				CAMS Reanalysis			
MONTH	N	AVG	BIAS	RMSE	R	AVG	BIAS	RMSE	R	AVG	BIAS	RMSE	R
ALL	57,447	10.4 ± 6.0	0.05	1.42	0.90	8.1 ± 5.0	−2.27	4.90	0.77	8.9 ± 5.7	−1.41	3.57	0.88
JAN	295	16.9 ± 6.0	0.06	3.08	0.70	13.4 ± 8.8	−3.46	9.40	0.58	12.4 ± 7.2	−4.39	8.27	0.72
FEB	3389	9.7 ± 3.6	0.07	1.45	0.78	8.1 ± 4.1	−1.53	4.35	0.61	7.8 ± 3.6	−1.78	3.71	0.75
MAR	9770	20.1 ± 7.1	0.06	1.96	0.88	14.1 ± 7.0	−5.97	8.61	0.70	17.0 ± 7.9	−3.13	5.58	0.84
APR	5394	9.6 ± 3.2	0.05	1.33	0.77	8.5 ± 4.2	−1.14	4.03	0.60	8.1 ± 3.8	−1.45	3.52	0.71
MAY	6599	9.0 ± 3.0	0.04	1.23	0.77	7.0 ± 3.0	−1.96	3.87	0.59	8.0 ± 3.3	−0.99	2.86	0.76
JUN	6921	8.3 ± 2.5	0.02	1.19	0.73	6.6 ± 2.6	−1.66	3.44	0.57	7.3 ± 2.9	−0.96	2.65	0.73
JUL	7039	7.5 ± 2.8	0.02	1.27	0.75	5.5 ± 2.3	−1.93	3.81	0.57	6.5 ± 2.7	−0.93	2.91	0.73
AUG	7582	7.9 ± 2.5	0.04	1.16	0.74	6.1 ± 2.4	−1.75	3.26	0.65	7.2 ± 2.9	−0.67	2.46	0.76
SEP	5141	6.8 ± 2.5	0.09	1.18	0.74	6.5 ± 3.4	−0.29	2.83	0.68	6.2 ± 2.8	−0.56	2.57	0.72
OCT	4616	8.4 ± 3.8	0.05	1.32	0.73	7.2 ± 3.8	−1.17	3.54	0.64	7.1 ± 3.3	−1.26	2.94	0.76
NOV	698	8.4 ± 2.8	0.06	1.64	0.64	6.3 ± 3.1	−2.11	4.47	0.54	6.1 ± 2.9	−2.26	4.13	0.67

In contrast, CAMS forecasts exhibit a weaker agreement with station observations in terms of temporal variability. Although CAMS forecasts still reproduce general variations reasonably well, with an overall correlation coefficient of 0.77, $PM_{2.5}$ concentrations are consistently underestimated, and peaks are not well captured. CAMS reanalyses generally show better agreement with station measurements in terms of temporal variability, as indicated by a correlation coefficient of 0.88. However, a systematic underestimation is also observed. Both CAMS forecasts and reanalyses show considerable negative biases compared to station observations with values of 2.3 $\mu g/m^3$ and 1.41 $\mu g/m^3$, respectively. Moreover, both CAMS datasets show significantly higher RMSEs than the RF model with a value of 4.90 $\mu g/m^3$. With respect to the mean values, relative RMSEs are 13% for RF, 60% for CAMS forecasts and 40% for CAMS reanalyses, representing a substantial difference in model accuracy.

At the bottom of Figure 5 a closer look at the temporal evolution of $PM_{2.5}$ levels during March 2022 is provided. As previously mentioned, the concentrations throughout the entire month remain above the annual mean. Particularly during the second half of March concentrations even rise above the monthly mean of around 19 $\mu g/m^3$. Here, it becomes even more evident than in the annual time series that the RF model performs very well in reproducing the temporal variability of $PM_{2.5}$ (R = 0.88), and CAMS reanalyses show a better agreement (R = 0.84) compared to CAMS forecasts (R = 0.7). The mean concentrations from RF and CAMS reanalyses are within the standard deviation of the in situ measurements at all timesteps, although with a pronounced bias in CAMS reanalyses

(-3.13 μg/m^3). CAMS forecasts are not within the standard deviation range and the bias of almost -6 μg/m^3 indicates a pronounced underestimation relative to the in situ measurements. A comparison with other months (see Table 6) reveals that both CAMS forecasts and reanalyses tend to consistently underestimate $PM_{2.5}$ concentrations. Bias values are negative throughout the year, with maxima in March (and January) being twice as large as during "normal" months. Also, the RMSEs are exceptionally high in March with values of 8.61 μg/m^3 and 5.58 μg/m^3 for forecasts and reanalyses, respectively. Correlation coefficients are lowest for CAMS forecasts throughout the entire year 2022, indicating substantial deviations from in situ measurements, particularly at higher concentration levels. In contrast, the RF-derived $PM_{2.5}$ concentrations as well as CAMS reanalyses exhibit much closer agreement in terms of correlation. With respect to RMSE and bias, the RF model performs better than both CAMS forecasts and reanalyses.

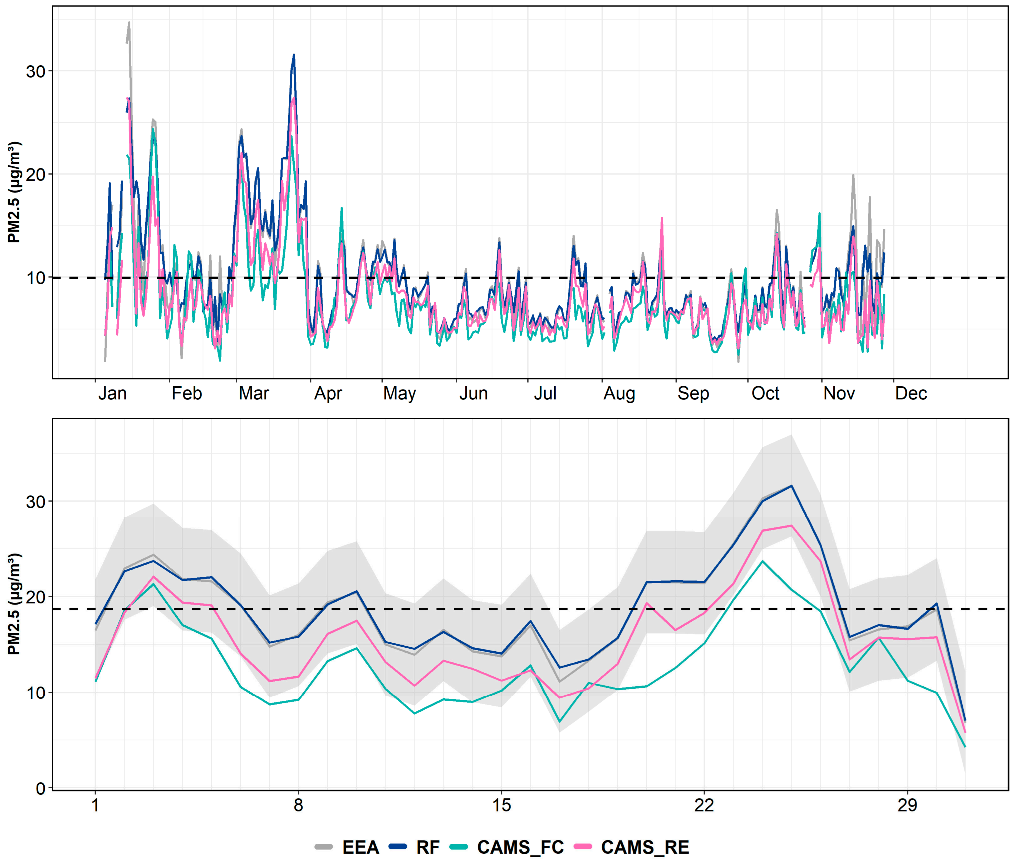

Figure 5. $PM_{2.5}$ time series of daily averages for the complete year 2022 (**top**) and March 2022 (**bottom**) based on RF, in situ measurements, CAMS forecasts and CAMS reanalysis. Values are averaged over all station samples included in the validation dataset. Color codes: observations (gray) with standard deviation (gray shadow); RF (dark blue), CAMS forecasts (turquoise) and CAMS reanalysis (pink). Dashed lines indicate the mean value of observations over the entire time window.

Figure 6 shows boxplots comparing the distributions of $PM_{2.5}$ values from in situ measurements, the RF model and CAMS across different value ranges. In the lowest range, all four datasets show relatively similar distributions, with CAMS slightly underestimating and the RF model slightly overestimating observed $PM_{2.5}$ concentrations. However, as concentration levels increase, the discrepancies among the datasets become more pronounced. The scatter of values, as indicated by the thin lines in Figure 6, increases substantially at higher levels, reflecting larger uncertainties under high-pollution conditions. The largest deviations between the datasets are generally apparent in the highest value range, where values of all datasets no longer overlap with in situ observations. However, the RF-based concentrations remain in closer agreement with the station measurements, although showing a noticeable underestimation relative to the in situ observations in the highest value range, too. In all other value ranges the RF model deviates less from the observations compared to both CAMS forecasts and reanalysis.

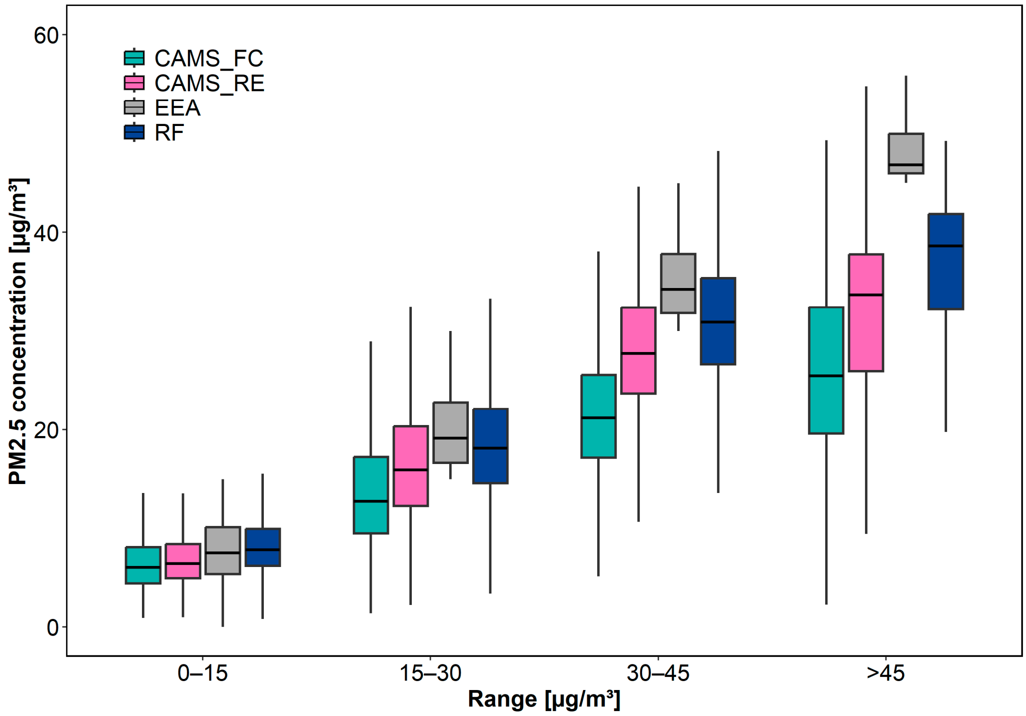

Figure 6. Boxplots for the value distributions of in situ, RF and CAMS datasets for different value ranges.

In the following, we focus on the high-pollution episode end of March 2022 with the aim to analyze how the spatial variability in $PM_{2.5}$ concentrations is represented by the TROPOMI-based RF model compared to CAMS forecasts and reanalysis. In addition, we examine the individual predictor variables included in the RF model regarding their influence on derived $PM_{2.5}$ concentrations.

Figure 7 presents maps of $PM_{2.5}$ concentrations from the RF model as well as CAMS forecasts and reanalysis on a daily basis from the 22 to 26 of March 2022. The RF-based maps show a pronounced increase in $PM_{2.5}$ levels across the study area, reaching a spatially extensive maximum on the 25 of March. Especially high concentrations occur in the northern part of the study area, particularly in Northern France, Belgium, the Netherlands, Northern Germany and Denmark. But in the eastern region covering Poland and Czech Republic, the RF model also indicates substantially increased $PM_{2.5}$ levels. An extensive high-pressure system, associated with dry and stagnant conditions, was present over Europe at that time [33], leading to high air pollution predominantly caused by local emissions. The local emissions mainly originated from agricultural activities in the areas of maximum $PM_{2.5}$ concentrations along the North Sea, but emissions from residential heating also contributed, particularly in Eastern Europe [31].

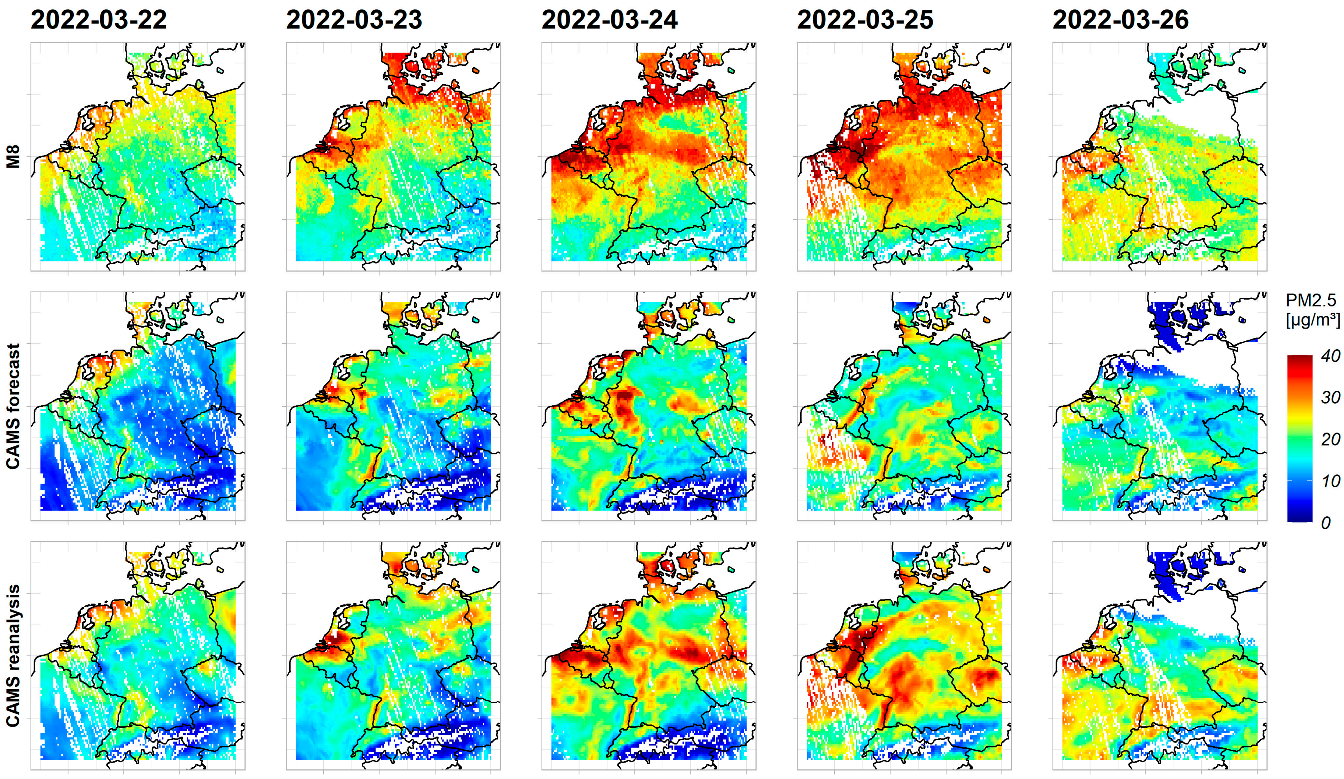

Figure 7. Daily maps of $PM_{2.5}$ concentrations from RF (M8), CAMS forecasts and CAMS reanalysis for the 22 to 26 of March 2022.

While the CAMS forecasts also indicate an increase in $PM_{2.5}$ concentrations over the northern region, they insufficiently reproduce the actual magnitude and extent of $PM_{2.5}$ pollution showing differences of more than 20 $\mu g/m^3$ in the maximum areas. Only in a few, spatially confined areas do CAMS forecasts reach the maximum $PM_{2.5}$ values obtained from the TROPOMI-based estimates, which, as shown in the previous analysis, are closer to the ground truth. Though CAMS forecasts capture the main pollution hotspots, they substantially underestimate the overall $PM_{2.5}$ burden during that episode. This can be at least partly due to CAMS's known underestimation of residential heating emissions [43].

The CAMS reanalysis exhibit $PM_{2.5}$ patterns that are much more consistent with those derived from the RF model than the corresponding CAMS forecasts. In particular, pollution hotspots are more clearly pronounced in the reanalyses, and regions of elevated $PM_{2.5}$ concentrations appear more spatially extended than in the forecasts. Nevertheless, the $PM_{2.5}$ gradients in CAMS reanalyses remain relatively sharp and the spatially continuous extent of elevated $PM_{2.5}$ concentrations is still underestimated, indicating limitations in representing regionally widespread pollution events.

Figure 8 presents maps of all the predictor variables contributing to $PM_{2.5}$ estimation. The strongest correlations are observed with AOD and CO. Regions with increasing $PM_{2.5}$ concentrations over the analyzed period also show elevated AOD and CO values. This contrasts with findings from Section 3.1. where feature importance rankings showed only moderate relevance of these variables for $PM_{2.5}$ estimation and AOD only improved the temporal correlation. However, results here show an exceptional period with strong $PM_{2.5}$ increase over a well-defined region.

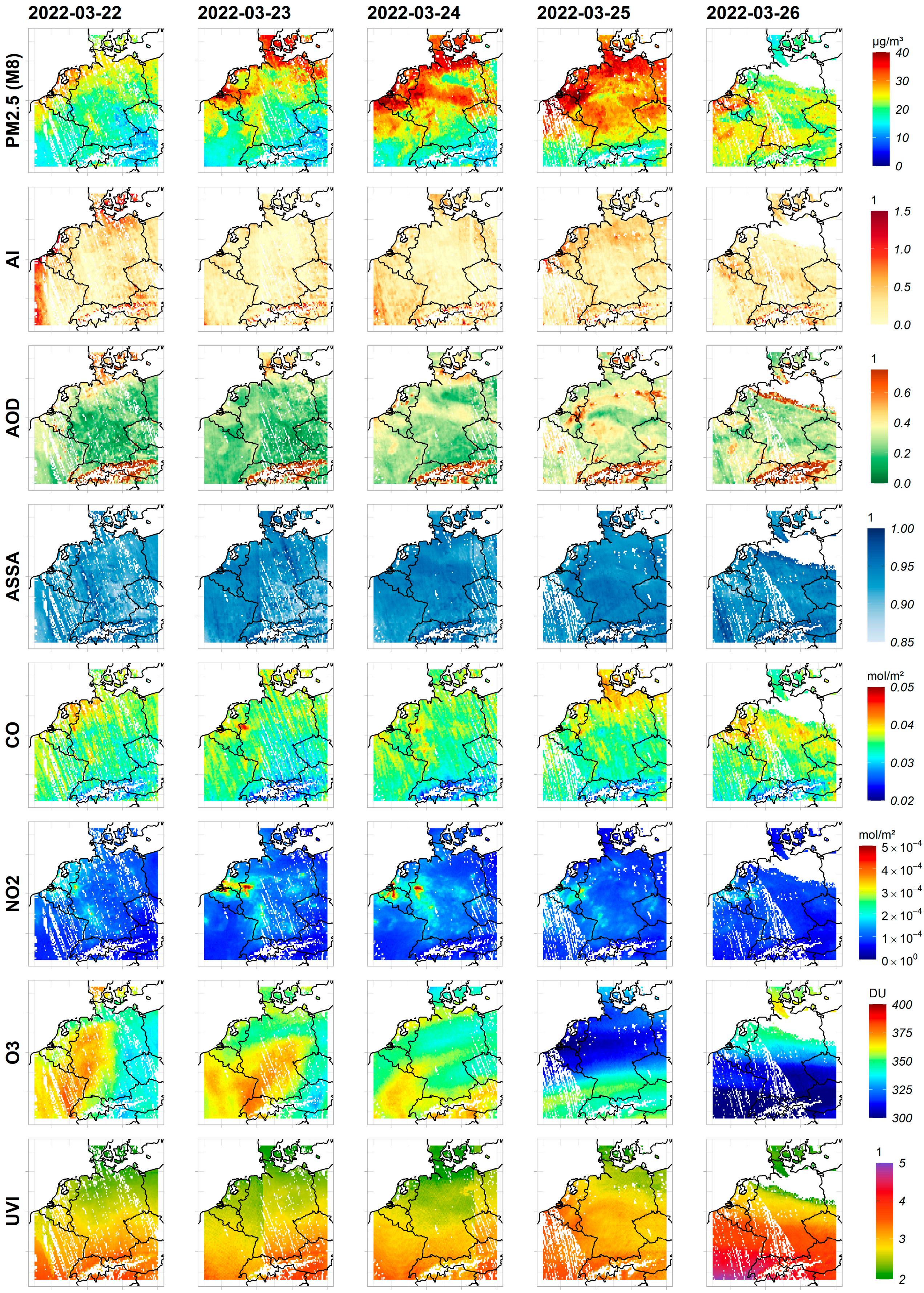

Figure 8. Maps of RF and CAMS results for all input parameters.

The north–south gradient in $PM_{2.5}$ concentrations can be also linked to the UVI, displaying a similar pattern. However, since the UVI is largely latitude-dependent, this relationship is likely coincidental rather than causal. Ozone exhibits an anticorrelation with $PM_{2.5}$ concentrations. Areas with high $PM_{2.5}$ tend to coincide with ozone minima, possibly due to reduced photolysis rates [44]. The ASSA values remain relatively constant throughout the study period, providing no evidence for changes in aerosol composition during this high-pollution episode. Similarly, the AI shows limited variability, though slight positive correlations with estimated $PM_{2.5}$ concentrations are apparent. Higher AI values indicate increased abundance of light-absorbing aerosols in the atmosphere, which may indicate the presence of larger numbers of carbonaceous particles, which can be, for example, associated with residential heating emissions. Local NO_2 hotspots coincide with the estimated $PM_{2.5}$ peaks, but these are spatially confined and do not appear to significantly contribute to the widespread $PM_{2.5}$ enhancements observed across the study region.

4. Discussion

We compared the performance of several RF model configurations trained with different combinations of meteorological and TROPOMI-based predictor variables using multiple CV strategies. Under sample-based CV, all models exhibited similarly high performance, indicating that the choice of input variables has only a minor effect when training and validation samples are randomly mixed. The more informative differences between the model configurations emerged under spatial and temporal CV. While the spatial CV showed overall only minor deviations from the sample-based CV, model performances were consistently lower under temporal CV, indicating that temporal generalization is more challenging than spatial generalization. There is an increased difficulty of accurately predicting $PM_{2.5}$ concentrations for time periods not represented in the training data, which may be due to the greater heterogeneity in time compared to space [11]. Higher temporal variability can arise from local and/or episodic emission events, changing meteorological conditions or the influence of passing intrusions, all of which are difficult to generalize, particularly when the training is based on a limited time span, as only a single year of data was considered in this study. Our results show that augmenting a purely meteorology-driven model with AOD substantially improves temporal generalization, demonstrating the importance of aerosol-related information for capturing the seasonal variability of $PM_{2.5}$. Spatial generalization is further enhanced when additional aerosol parameters (AI, ASSA, ALH) are included. This clearly highlights the added value of incorporating multiple aerosol-related predictors beyond AOD alone.

Previous studies have consistently emphasized the importance of meteorological variables for estimating $PM_{2.5}$ concentrations (e.g., [26]) and could even show that using meteorological data alone can be sufficient for accurate $PM_{2.5}$ prediction [45]. Our results confirm that $PM_{2.5}$ can indeed be reliably estimated using meteorological data alone. However, the findings also demonstrate that models including aerosol and other satellite-derived information outperform those relying exclusively on meteorological inputs. The relevance of aerosol information, particularly AOD, becomes more pronounced when only limited station measurements are available [46], but previous studies have generally demonstrated that incorporating AOD results in enhanced model performance [47]. Our results indicate that incorporating additional aerosol parameters further improves model performance and even show that models based solely on TROPOMI-derived parameters provide comparable high accuracies in estimating ground-level $PM_{2.5}$ concentrations. In many aspects such models offer distinct advantages. In general, it is a key benefit to use only a limited number of predictors, as a smaller set of parameters reduces preprocessing efforts as well as computational costs and improves model interpretability [48]. Moreover, the input data maintain consistency by avoiding the combination of data sources that differ in spatial and temporal resolution, physical basis or uncertainty. While the use of model-based data as input for $PM_{2.5}$ estimation offers the distinct advantage of complete spatial coverage (see Table 5), it must be recognized that such data can exhibit inconsistencies arising from changing assimilation data. For example, during the corona pandemic in 2020 missing aircraft measurements and other unforeseen environmental changes significantly degraded the quality of meteorological forecasts [49]. We could observe a negative effect on the prediction accuracy of RF models when including model-based meteorological data in that year [7]. Hence, though satellite observations carry uncertainties, mainly introduced by retrieval algorithms, they nevertheless provide observational evidence that can exceed model-based data in terms of consistency, reliability and uncertainty. Ultimately, an important advantage is that a computationally efficient RF model based solely on TROPOMI observations enables the operational monitoring of $PM_{2.5}$ in NRT.

Our analyses indicate that the inclusion of certain predictors such as ALH can improve overall model performance but at the same time substantially reduce the number of available training samples. This reduction in sample size affects the representativeness of specific time periods within the training dataset, as satellite data coverage varies strongly across seasons. In particular, winter months tend to be underrepresented, which poses a general challenge for model generalization. This limitation becomes especially critical when only a single year of data is considered for training, as it may impair the model's ability to generalize across time and increase the risk of overfitting [50]. These findings are directly relevant to the comparison of model performance across the different RF configurations, which have considerable differences in sample size.

The impact of limited temporal representativeness is also reflected in the feature importance analysis. Temporal predictors such as DOY and month rank highest, with even higher importance scores in model configurations that do not include meteorological variables. When meteorological predictors are added, the relative importance of DOY and month decreases, while variables such as T2m and STR gain importance, indicating that meteorological variables capture a substantial fraction of seasonal and temporal variability in $PM_{2.5}$. Importantly, the dominance of temporal predictors does not imply that the model fails to resolve spatial heterogeneity. Spatial CV results remain comparable to those obtained from sample-based CV, suggesting that spatial variability is reasonably captured. Moreover, spatial patterns shown in Figure 8 reveal a clear correspondence between aerosol-related predictors, such as AOD, and the derived $PM_{2.5}$ fields, highlighting the role of spatially varying aerosol information in shaping $PM_{2.5}$ distributions. Nevertheless, it is acknowledged that the use of a single-year training period represents a key limitation of this study and that extending the analysis to multiple years would likely reduce the dominance of temporal predictors and further improve temporal generalization.

The feature importance further reveals that AOD ranks below the trace gas indicators NO_2 and O_3, a result that is somewhat unexpected given that AOD provides direct information on the atmospheric particle burden. When no other pollutant-related predictors are included in the model, AOD becomes a substantially more important predictor, as previously demonstrated by us and many other studies [7,27,51]. However, NO_2 serves as an effective proxy for anthropogenic activity, population density and industrial emissions, thereby capturing information on local emission sources that is partly shared with $PM_{2.5}$ [44,52]. In addition, NO_2 and O_3 act as indicators of photochemical processes and secondary aerosol formation pathways, thereby indirectly capturing aspects of aerosol composition [53]. These processes are not directly represented by AOD alone and may therefore explain the higher importance ranking of NO_2 and O_3. A similar relationship was expected for CO, acting as a main precursor in secondary $PM_{2.5}$ formation. However, CO generally exhibits a lower importance ranking, in contrast to findings by [25], who identified CO as the most influential chemical predictor for $PM_{2.5}$ during biomass burning episodes, suggesting that the relevance of individual trace gases is strongly context-dependent. Interestingly, CO ranks slightly higher than AOD when meteorological variables are included. However, the differences in importance scores among NO_2, CO and AOD are so small in this case that they leave no room for meaningful interpretation.

The PDP analysis confirms that the RF models have learned relationships that are not only statistically robust but in general, physically meaningful. Nevertheless, some of the identified dependencies are less intuitive and need further investigation. One notable example is the positive relationship between $PM_{2.5}$ and air temperature. This finding is somewhat unexpected, as higher temperatures—typically occurring during summer—are often associated with lower $PM_{2.5}$ concentrations due to a deeper and better-mixed planetary boundary layer [54]. In addition, lower temperatures are commonly linked to higher relative humidity and lower vapor pressure, conditions that favor the formation of secondary organic aerosols (SOAs) [55]. The observed positive relationship may therefore reflect more dynamic processes, such as enhanced aerosol activation over dry soils or increased vertical transport and uplift under specific meteorological conditions.

Another interesting outcome of the PDP analysis is the apparent anti-correlation between O_3 and $PM_{2.5}$. In reality, the relationship between O_3 and $PM_{2.5}$ is complex and highly context-dependent, varying with season, geographical location, emission sources, meteorological conditions and chemical regime. For instance, ref. [56] reported a positive correlation during summer months, whereas negative correlations were observed in winter and generally under high pollution conditions. Similarly, ref. [57] found a negative relationship between O_3 and $PM_{2.5}$ during winter. The observed anti-correlation may partly be explained by the so-called aerosol–photolysis feedback, whereby $PM_{2.5}$ influences photolysis rates through the absorption and scattering of UV radiation, thereby suppressing photochemical ozone production [44]. Moreover, meteorological factors can modulate the O_3-$PM_{2.5}$ relationship. For example, ref. [58] reported contrasting correlations of $PM_{2.5}$ and O_3 with BLH. $PM_{2.5}$ was negatively correlated with BLH, consistent with the reduced vertical mixing under shallow boundary layers [59]. In contrast, O_3 showed a positive correlation with BLH, as enhanced solar radiation increases the boundary layer and photochemical activity. Overall, these findings highlight that the relationship between O_3 and $PM_{2.5}$ is driven by a combination of chemical, physical and meteorological processes. The PDP results therefore provide valuable insights but should be interpreted in the context of the prevailing environmental conditions, underscoring the need for further targeted investigations.

We compared the RF-based $PM_{2.5}$ estimates with both CAMS forecasts and reanalyses and found that the RF predictions generally showed closer agreement with in situ observations across all value ranges. In particular, CAMS forecasts substantially underestimated $PM_{2.5}$ concentrations and exhibited pronounced deviations from station measurements in their temporal evolution. CAMS reanalyses showed correlation levels comparable to those of the RF-based estimates but still displayed a systematic underestimation, which became increasingly pronounced at higher concentration levels. The RF model achieved markedly better agreement with observations at low and moderate $PM_{2.5}$ concentrations, while also showing a tendency to underestimate very high values. This behavior can partly be attributed to data imbalance, as low and moderate concentrations occurred far more frequently and therefore dominated the training dataset. As a result, fewer samples were available to constrain the model at high concentration levels. In addition, ensemble-based models such as Random Forests are known to smooth extreme values, as predictions are averaged across many decision trees, which further contributes to the underestimation of rare extremes [37].

Despite these limitations, the RF model was still able to capture elevated $PM_{2.5}$ concentrations more accurately than both CAMS forecasts and reanalyses, benefiting from its observational constraints. CAMS products rely strongly on prescribed emission inventories, which limits their ability to reproduce episodic high pollution events [60]. CAMS reanalyses partially mitigate this limitation through the assimilation of ground-based measurements, leading to improved performance in regions with dense station coverage. Accordingly, we observed high agreement between CAMS reanalyses and RF predictions in areas with many monitoring stations. However, CAMS reanalyses tend to exhibit sharp spatial gradients and fail to realistically capture spatially extended elevated $PM_{2.5}$ concentrations, reflecting their reliance on localized data assimilation. In contrast, while ground-based measurements are required for model training, the RF-based approach is independent of station data at the prediction stage, representing a clear advantage for spatially continuous applications. The strong gradients and limited spatial extent of elevated $PM_{2.5}$ concentrations in CAMS products may also result from the assimilation of background stations only, where traffic and industrial stations are not considered, while such information is implicitly included in the RF training. Furthermore, the coarser spatial resolution of CAMS products likely contributes to the underestimation of peak concentrations, as the mismatch between grid-based estimates and point measurements tends to smooth extreme values [61], an effect that becomes more pronounced at coarser resolutions.

Overall, our results indicate that CAMS products are less suitable for analyzing high-pollution episodes or extreme $PM_{2.5}$ events, whereas the RF-based approach demonstrates an improved ability to reproduce spatial extent, magnitude and temporal variability of elevated $PM_{2.5}$ concentrations. While this conclusion is based on a single-year analysis, it is consistent with the findings from one of our previous studies, which reported similar results for a five-year comparison between RF-based $PM_{2.5}$ estimates and CAMS reanalyses [7]. Importantly, since the primary objective of this work is the development of an operational NRT $PM_{2.5}$ product, comparisons with CAMS forecasts are particularly relevant. In this context, our results demonstrate a clear advantage of the proposed RF-based method.

5. Conclusions

This study applied a Random Forest approach to estimate daily surface $PM_{2.5}$ concentrations from TROPOMI measurements over Germany and surrounding countries for the year 2022. By systematically testing different combinations of input variables, we evaluated the influence of various predictors on model performance. Beyond the traditional use of AOD and meteorological variables, several TROPOMI atmospheric composition products were assessed for their added value. The results demonstrate that additional TROPOMI parameters improve model accuracy and, very remarkably, that reliable $PM_{2.5}$ estimates can be obtained using TROPOMI data alone. This challenges the prevailing statement in the literature that meteorological information is indispensable for accurate $PM_{2.5}$ estimation.

Among the tested variables, the Aerosol Layer Height (ALH) showed a particular positive effect on model performance. However, its currently limited spatial coverage restricts its operational use. With the upcoming Sentinel-4 mission providing hourly observations of atmospheric composition over Europe, these limitations are expected to diminish, enabling more continuous $PM_{2.5}$ mapping, potentially including ALH as a main predictor.

To show the potential of our RF model for NRT applications in $PM_{2.5}$ pollution assessment, we compared it to CAMS forecast data, as these are also available in NRT. Our results revealed that the TROPOMI-based RF model more accurately reproduced the spatial and temporal variability of $PM_{2.5}$, especially during high-pollution episodes. Although deviations increased at very high concentration levels, pollution peaks were well captured, and the RF-based maps provided finer spatial detail than CAMS forecasts. This capability is crucial, as the frequency and intensity of high-pollution events are expected to rise under future climate conditions due to enhanced droughts, wildfires and dust events [62].

Overall, this study highlights the potential of satellite-driven machine learning approaches for NRT air quality monitoring. While the current analysis was limited to Germany and surrounding regions, the results demonstrate the feasibility of the proposed approach under a wide range of geographical, meteorological and emission conditions. The transferability of the method to the geostationary mission Sentinel-4 therefore represents a promising perspective, although further validation across Europe, particularly in regions with distinct emission characteristics or complex topography, is required. Subject to such validation and the availability of sufficiently dense ground-based observations, Sentinel-4-based applications could enable daily, high-resolution mapping of $PM_{2.5}$ across Europe. These products have the potential to support exposure assessment, health risk evaluation and the development of effective mitigation and policy strategies.

Author Contributions: The authors contributions are as follows: conceptualization, J.H. and F.B.; methodology, J.H.; software, J.H., V.M.G. and P.F.; validation, J.H.; formal analysis, J.H.; investigation, J.H.; resources, F.B. and D.L.; data curation, J.H., V.M.G. and P.F.; writing—original draft preparation, J.H.; writing—review and editing, F.B., D.L., P.F. and V.M.G.; visualization, J.H.; supervision, F.B. and D.L.; project administration, D.L.; funding acquisition, D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was carried out within the framework of the DLR project INPULS and did not receive any external funding. The corresponding funding codes are MF-8472373-KTR 2472922 and DFD-8472372-KTR 2472924.

Data Availability Statement: The datasets generated in this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors gratefully acknowledge the computational and data resources provided through the joint high-performance data analytics (HPDA) project "terrabyte" of the German Aerospace Center (DLR) and the Leibniz Supercomputing Center (LRZ). We also thank the data suppliers, EU/ESA/DLR for the operational TROPOMI L2 products, DLR/NASA for the TROPOMI aerosol products, EEA for the access to the in situ $PM_{2.5}$ measurements and ECMWF for the meteorological data used in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Health Organization. WHO Global Air Quality Guidelines: Particulate Matter ($PM_{2.5}$ And PM_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. 2021. Available online: https://www.who.int/publications/i/item/9789240034228 (accessed on 5 December 2025).
2. Sang, S.; Chu, C.; Zhang, T.; Chen, H.; Yang, X. The global burden of disease attributable to ambient fine particulate matter in 204 countries and territories, 1990–2019: A systematic analysis of the Global Burden of Disease Study 2019. *Ecotoxicol. Environ. Saf.* **2022**, *238*, 113588. [CrossRef] [PubMed]
3. European Environment Agency (EEA). Air Quality in Europe 2022. 2022. Available online: https://www.eea.europa.eu/en/analysis/publications/air-quality-in-europe-2022 (accessed on 4 November 2025).
4. Sokhi, R.S.; Moussiopoulos, N.; Baklanov, A.; Bartzis, J.; Coll, I.; Finardi, S.; Friedrich, R.; Geels, C.; Grönholm, T.; Halenka, T.; et al. Advances in air quality research–current and emerging challenges. *Atmos. Chem. Phys. Discuss.* **2021**, *22*, 4615–4703. [CrossRef]
5. Manisalidis, I.; Stavropoulou, E.; Stavropoulos, A.; Bezirtzoglou, E. Environmental and health impacts of air pollution: A review. *Front. Public Health* **2020**, *8*, 14. [CrossRef]

6. Friberg, M.D.; Kahn, R.A.; Limbacher, J.A.; Appel, K.W.; Mulholland, J.A. Constraining chemical transport $PM_{2.5}$ modeling outputs using surface monitor measurements and satellite retrievals: Application over the San Joaquin Valley. *Atmos. Chem. Phys.* **2018**, *18*, 12891–12913. [CrossRef]
7. Handschuh, J.; Erbertseder, T.; Baier, F. On the added value of satellite AOD for the investigation of ground-level $PM_{2.5}$ variability. *Atmos. Environ.* **2024**, *331*, 120601. [CrossRef]
8. Chen, Z.Y.; Méndez, R.; Petetin, H.; Lacima, A.; García-Pando, C.P.; Ballester, J. A pan-European, high-resolution, daily total, fine-mode and coarse-mode aerosol optical depth dataset based on quantile machine learning. *Earth Syst. Sci. Data Discuss.* **2023**, *2023*, 1–31. [CrossRef]
9. Sorek-Hamer, M.; Chatfield, R.; Liu, Y. Strategies for using satellite-based products in modeling $PM_{2.5}$ and short-term pollution episodes. *Environ. Int.* **2020**, *144*, 106057. [CrossRef]
10. Zhang, Y.; Li, Z.; Bai, K.; Wei, Y.; Xie, Y.; Zhang, Y.; Ou, Y.; Cohen, J.; Zhang, Y.; Peng, Z.; et al. Satellite remote sensing of atmospheric particulate matter mass concentration: Advances, challenges, and perspectives. *Fundam. Res.* **2021**, *1*, 240–258. [CrossRef]
11. Zhu, S.; Tang, J.; Zhou, X.; Li, P.; Liu, Z.; Zhang, C.; Zou, Z.; Li, T.; Peng, C. Research progress, challenges, and prospects of $PM_{2.5}$ concentration estimation using satellite data. *Environ. Rev.* **2023**, *31*, 605–631. [CrossRef]
12. Ma, Z.; Dey, S.; Christopher, S.; Liu, R.; Bi, J.; Balyan, P.; Liu, Y. A review of statistical methods used for developing large-scale and long-term $PM_{2.5}$ models from satellite data. *Remote Sens. Environ.* **2022**, *269*, 112827. [CrossRef]
13. Gianquintieri, L.; Oxoli, D.; Caiani, E.G.; Brovelli, M.A. State-of-art in modelling particulate matter (PM) concentration: A scoping review of aims and methods. *Environ. Dev. Sustain.* **2024**, *27*, 25889–25911. [CrossRef]
14. Veefkind, J.P.; Aben, I.; McMullan, K.; Förster, H.; De Vries, J.; Otter, G.; Claas, J.; Eskes, H.J.; De Haan, J.F.; Kleipool, Q.; et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens. Environ.* **2012**, *120*, 70–83. [CrossRef]
15. Akcin, H.; Uyar, N.; Yıldırım, Y.; Ozdogan, N. Assessing air pollutant emissions from undercity mines using different satallite tropospheric monitoring sensor data. *Int. J. Environ. Sci. Technol.* **2025**, *22*, 14849–14866. [CrossRef]
16. Davybida, L.I. Air quality impacts of war detected from the Sentinel-5P satellite over Ukraine. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2022; Volume 1254, p. 012112. [CrossRef]
17. El-Zeiny, A.; Nagy, A.; Kaamoush, M. Impact of maritime and industrial activities on air quality along the Egyptian Mediterranean coast using data observation techniques. *Euro-Mediterr. J. Environ. Integr.* **2025**, *10*, 2883–2896. [CrossRef]
18. Rudke, A.P.; Martins, J.A.; Hallak, R.; Martins, L.D.; De Almeida, D.S.; Beal, A.; Freitas, E.D.; Andrade, M.F.; Koutrakis, P.; Albuquerque, T.T.A. Evaluating TROPOMI and MODIS performance to capture the dynamic of air pollution in São Paulo state: A case study during the COVID-19 outbreak. *Remote Sens. Environ.* **2023**, *289*, 113514. [CrossRef]
19. Vîrghileanu, M.; Săvulescu, I.; Mihai, B.A.; Nistor, C.; Dobre, R. Nitrogen Dioxide (NO_2) Pollution monitoring with Sentinel-5P satellite imagery over Europe during the coronavirus pandemic outbreak. *Remote Sens.* **2020**, *12*, 3575. [CrossRef]
20. Reshi, A.R.; Pichuka, S.; Tripathi, A. Applications of Sentinel-5P TROPOMI satellite sensor: A review. *IEEE Sens. J.* **2024**, *24*, 20312–20321. [CrossRef]
21. Balamurugan, V.; Chen, J.; Wenzel, A.; Keutsch, F.N. Spatiotemporal modeling of air pollutant concentrations in Germany using machine learning. *Atmos. Chem. Phys.* **2023**, *23*, 10267–10285. [CrossRef]
22. Chan, K.L.; Khorsandi, E.; Liu, S.; Baier, F.; Valks, P. Estimation of surface NO_2 concentrations over Germany from TROPOMI satellite observations using a machine learning method. *Remote Sens.* **2021**, *13*, 969. [CrossRef]
23. Naseer, E.; Basit, A.; Bhatti, M.K.; Siddique, M.A. Machine learning for area-wide monitoring of surface level concentration of NO_2 using remote sensing data. In Proceedings of the International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE), Lahore, Pakistan, 2–4 December 2022; IEEE: Piscataway, NJ, USA; pp. 1–6. [CrossRef]
24. Han, S.; Kundhikanjana, W.; Towashiraporn, P.; Stratoulias, D. Interpolation-based fusion of Sentinel-5P, SRTM, and regulatory-grade ground stations data for producing spatially continuous maps of $PM_{2.5}$ concentrations nationwide over Thailand. *Atmosphere* **2022**, *13*, 161. [CrossRef]
25. Son, R.; Stratoulias, D.; Kim, H.C.; Yoon, J.H. Estimation of surface $PM_{2.5}$ concentrations from atmospheric gas species retrieved from TROPOMI using deep learning: Impacts of fire on air pollution over Thailand. *Atmos. Pollut. Res.* **2023**, *14*, 101875. [CrossRef]
26. Wang, Y.; Yuan, Q.; Li, T.; Tan, S.; Zhang, L. Full-coverage spatiotemporal mapping of ambient $PM_{2.5}$ and PM_{10} over China from Sentinel-5P and assimilated datasets: Considering the precursors and chemical compositions. *Sci. Total Environ.* **2021**, *793*, 148535. [CrossRef]
27. Handschuh, J.; Erbertseder, T.; Baier, F. Systematic evaluation of four satellite AOD datasets for estimating $PM_{2.5}$ using a random forest approach. *Remote Sens.* **2023**, *15*, 2064. [CrossRef]
28. Falah, S.; Kizel, F.; Banerjee, T.; Broday, D.M. Accounting for the aerosol type and additional satellite-borne aerosol products improves the prediction of $PM2_{.5}$ concentrations. *Environ. Pollut.* **2023**, *320*, 121119. [CrossRef]

29. Torres, O.; Jethva, H.; Ahn, C.; Jaross, G.; Loyola, D.G. TROPOMI aerosol products: Evaluation and observations of synoptic scale carbonaceous aerosol plumes during 2018–2020. *Atmos. Meas. Tech.* **2020**, *13*, 6789–6806. [CrossRef]
30. Timmermans, R.; van Pinxteren, D.; Kranenburg, R.; Hendriks, C.; Fomba, K.W.; Herrmann, H.; Schaap, M. Evaluation of modelled LOTOS-EUROS with observational based PM_{10} source attribution. *Atmos. Environ. X* **2022**, *14*, 100173. [CrossRef]
31. Hamer, P.; Fjæraa, A.M.; Tarrasón, L.; Soares, J.; Meleux, F.; Colette, A.; Ung, A.; Raux, B.; Kuenen, J. Copernicus Atmosphere Monitoring Service Interim Annual Assessment Report on European Air Quality 2022. Available online: https://atmosphere.copernicus.eu/sites/default/files/publications/CAMS271_2021SCx_D1.1.1._202306_2022_Interim_Assessment_Report_v3.pdf (accessed on 6 November 2025).
32. Schulz, M.; Tsyro, S.; Mortier, A.; Valdebenito, A.; Benedictow, A. Saharan Dust Episode of 13–18 March 2022—Episode Analysis: CAMS2_71 Report N°01 in 2022, CAMS71_D3.2.1-2022-2-A_11-18MarchEpisode_v1.0 (CAMS). Available online: https://policy.atmosphere.copernicus.eu/reports/CAMS2-71_D3.2.1-2022-1_PM10_episode_13-18March2022_v1.pdf (accessed on 6 November 2025).
33. Tsyro, S.; Schulz, M.; Mortier, A.; Valdebenito, A.; Benedictow, A.; Timmerman, R.; Kranenburg, R. High PM_{10} Levels: Episode of 20–27 March 2022—CAMS2_71 Episode Analysis Report N°02 in 2022, CAMS71_D3.2.1-2022-2-B_20-27MarchEpisode. Available online: https://policy.atmosphere.copernicus.eu/reports/CAMS2-71_PM10_episode_20-27March2022_final.pdf (accessed on 6 November 2025).
34. ECMWF Forecast User Guide: HRES—High Resolution Forecasts. Available online: https://confluence.ecmwf.int/display/FUG/Forecast+User+Guide (accessed on 23 October 2025).
35. CAMS Regional: European Air Quality Analysis and Forecast Data Documentation. Available online: https://confluence.ecmwf.int/display/CKB/CAMS+Regional%3A+European+air+quality+analysis+and+forecast+data+documentation (accessed on 23 October 2025).
36. CAMS Regional: European Air Quality Reanalyses Data Documentation. Available online: https://confluence.ecmwf.int/display/CKB/CAMS+Regional%3A+European+air+quality+reanalyses+data+documentation (accessed on 21 January 2026).
37. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef] [PubMed]
38. Wright, M.N.; Ziegler, A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **2017**, *77*, 1–17. [CrossRef]
39. Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M.B.; Choirat, C.; Koutrakis, P.; Lyapustin, A.; et al. An ensemble-based model of $PM_{2.5}$ concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **2019**, *130*, 104909. [CrossRef]
40. Gao, X.; Ruan, Z.; Liu, J.; Chen, Q.; Yuan, Y. Analysis of atmospheric pollutants and meteorological factors on $PM_{2.5}$ concentration and temporal variations in Harbin. *Atmosphere* **2022**, *13*, 1426. [CrossRef]
41. Greenwell, B.M. pdp: An R package for constructing partial dependence plots. *R J.* **2017**, *9*, 421–436. [CrossRef]
42. Proposal for a Revision of the Ambient Air Quality Directives. In European Parliamentary Research Service, 2022, vol. 347, Issue October. Available online: https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/734679/EPRS_BRI(2022)734679_EN.pdf (accessed on 6 November 2025).
43. Tarrasón, L.; Hamer, P.D.; Meleux, F.; Colette, A.; Rouïl, L. Interim Annual Assessment Report for 2021, Copernicus Atmosphere Monitoring Service (CAMS) Report. Available online: https://policy.atmosphere.copernicus.eu/reports/CAMS271_IAR_2021_Final.pdf (accessed on 6 November 2025).
44. Qu, Y.; Wang, T.; Yuan, C.; Wu, H.; Gao, L.; Huang, C.; Li, Y.; Li, M.; Xie, M. The underlying mechanisms of $PM_{2.5}$ and O_3 synergistic pollution in East China: Photochemical and heterogeneous interactions. *Sci. Total Environ.* **2023**, *873*, 162434. [CrossRef]
45. Wang, Z.; Chen, P.; Wang, R.; An, Z.; Qiu, L. Estimation of $PM_{2.5}$ concentrations with high spatiotemporal resolution in Beijing using the ERA5 dataset and machine learning models. *Adv. Space Res.* **2023**, *71*, 3150–3165. [CrossRef]
46. Tian, Z.; Wei, J.; Li, Z. How important is satellite-retrieved aerosol optical depth in deriving surface $PM_{2.5}$ using machine learning? *Remote Sens.* **2023**, *15*, 3780. [CrossRef]
47. Bai, K.; Li, K.; Sun, Y.; Wu, L.; Zhang, Y.; Chang, N.B.; Li, Z. Global synthesis of two decades of research on improving $PM_{2.5}$ estimation models from remote sensing and data science perspectives. *Earth-Sci. Rev.* **2023**, *241*, 104461. [CrossRef]
48. Padimala, S.S.K.; Matli, C.S. A comprehensive review delineates advancements in retrieving particulate matter utilising satellite aerosol optical depth: Parameter consideration, data processing, models development and future perspectives. *Atmos. Res.* **2024**, *308*, 107514. [CrossRef]
49. Chen, Y. COVID-19 pandemic imperils weather forecast. *Geophys. Res. Lett.* **2020**, *47*, e2020GL088613. [CrossRef]
50. Ibrahim, S.; Landa, M.; Pešek, O.; Brodský, L.; Halounová, L. Machine learning-based approach using open data to estimate $PM_{2.5}$ over Europe. *Remote Sens.* **2022**, *14*, 3392. [CrossRef]
51. Amiri, Z.; Zare Shahne, M. Modeling $PM_{2.5}$ concentration in Tehran using satellite-based Aerosol optical depth (AOD) and machine learning: Assessing input contributions and prediction accuracy. *Remote Sens. Appl. Soc. Environ.* **2025**, *38*, 101549. [CrossRef]

52. Bridhikitti, A. Applications of Remote Sensing for Air Pollution Monitoring in Thailand: An Early Warning for Public Health. In *Earth Data Analytics for Planetary Health*; Springer Nature: Singapore, 2023; pp. 3–31. [CrossRef]
53. Guo, B.; Wang, Z.; Pei, L.; Zhu, X.; Chen, Q.; Wu, H.; Zhang, W.; Zhang, D. Reconstructing MODIS aerosol optical depth and exploring dynamic and influential factors of AOD via random forest at the global scale. *Atmos. Environ.* **2023**, *315*, 120159. [CrossRef]
54. Kalmus, P.; Ao, C.O.; Wang, K.N.; Manzi, M.P.; Teixeira, J. A high-resolution planetary boundary layer height seasonal climatology from GNSS radio occultations. *Remote Sens. Environ.* **2022**, *276*, 113037. [CrossRef]
55. Thomsen, D.; Iversen, E.M.; Skønager, J.T.; Luo, Y.; Li, L.; Roldin, P.; Priestley, M.; Pedersen, H.B.; Hallquist, M.; Ehn, M. The effect of temperature and relative humidity on secondary organic aerosol formation from ozonolysis of Δ3-carene. *Environ. Sci. Atmos.* **2024**, *4*, 88–103. [CrossRef]
56. Li, Y.; Wang, T.; Wang, Q.; Qu, Y.; Wu, H.; Xie, M.; Li, M.; Li, S.; Zhuang, B. Spatiotemporal Variations of $PM_{2.5}$ and O_3 Relationship during 2014–2021 in Eastern China. *Aerosol Air Qual. Res.* **2023**, *23*, 230060. [CrossRef]
57. Ding, A.J.; Fu, C.B.; Yang, X.Q.; Sun, J.N.; Zheng, L.F.; Xie, Y.N.; Herrmann, E.; Nie, W.; Petäjä, T.; Kerminen, V.-M.; et al. Ozone and fine particle in the western Yangtze River Delta: An overview of 1 yr data at the SORPES station. *Atmos. Chem. Phys.* **2013**, *13*, 5813–5830. [CrossRef]
58. Zhou, Q.; Cheng, L.; Zhang, Y.; Wang, Z.; Yang, S. Relationships between Springtime $PM_{2.5}$, PM_{10}, and O_3 Pollution and the Boundary Layer Structure in Beijing, China. *Sustainability* **2022**, *14*, 9041. [CrossRef]
59. Li, J.; Chen, H.; Li, Z.; Wang, P.; Cribb, M.; Fan, X. Low-level temperature inversions and their effect on aerosol condensation nuclei concentrations under different large-scale synoptic circulations. *Adv. Atmos. Sci.* **2015**, *32*, 898–908. [CrossRef]
60. Akritidis, D.; Katragkou, E.; Georgoulias, A.K.; Zanis, P.; Kartsios, S.; Flemming, J.; Inness, A.; Douros, J.; Eskes, H. A complex aerosol transport event over Europe during the 2017 Storm Ophelia in CAMS forecast systems: Analysis and evaluation. *Atmos. Chem. Phys.* **2020**, *20*, 13557–13578. [CrossRef]
61. Xiao, Q.; Chang, H.H.; Geng, G.; Liu, Y. An ensemble machine-learning model to predict historical $PM_{2.5}$ concentrations in China from satellite data. *Environ. Sci. Technol.* **2018**, *52*, 13260–13269. [CrossRef] [PubMed]
62. Coelho, S.; Rafael, S.; Lopes, D.; Miranda, A.I.; Ferreira, J. How changing climate may influence air pollution control strategies for 2030? *Sci. Total Environ.* **2021**, *758*, 143911. [CrossRef]