atmosphere

MDPI

Article

A Physics-Informed Framework Linking Satellite AOD and Ambient Particulate Matter: A Pilot Study

Giorgia Proietti Pelliccia [1,2], Erika Brattich [1], Andrea Faggi [1], Silvana Di Sabatino [1] and Tiziano Maestri [1,*]

[1] Department of Physics and Astronomy "Augusto Righi", University of Bologna, 40126 Bologna, Italy; giorgia.proietti6@unibo.it (G.P.P.); erika.brattich@unibo.it (E.B.); andrea.faggi3@unibo.it (A.F.); silvana.disabatino@unibo.it (S.D.S.)

[2] Department of Engineering, University of Basilicata, 85100 Potenza, Italy

* Correspondence: tiziano.maestri@unibo.it

Abstract

Recently, numerous studies have exploited satellite Aerosol Optical Depth (AOD) to estimate near-surface particulate matter (PM) concentrations, with the aim of overcoming the limited spatial and temporal coverage of ground-based air quality monitoring networks. Despite significant progress, the relationship between AOD and PM remains highly uncertain, mainly due to the inadequate representation of local aerosol microphysical properties and of hygroscopic growth effects. In particular, satellite AOD is retrieved at ambient relative humidity, whereas standard PM measurements are performed under dry conditions. This study proposes a physics-informed, semi-empirical approach that overcomes these limitations by directly relating satellite AOD to PM measured at ambient humidity. Co-located measurements, from a Light Optical Aerosol Counter (LOAC) in the urban area of Bologna (Po Valley, Italy) during 2023, are used. This study is designed as a pilot application to evaluate the physical consistency of the proposed framework under well-characterised observational conditions, including spatial co-location, temporal matching to satellite overpasses, and exclusion of precipitation and desert dust events. The LOAC provides particle number size distribution and particle-type classification, which are used to estimate key aerosol properties controlling the AOD–PM theoretical relationship, including the Effective Radius, Extinction Efficiency, and aerosol Mass Density. These quantities, together with Mixing Layer Height, are combined within a theoretical framework linking PM and AOD, allowing for the derivation of a physically based scaling coefficient without relying on empirical hygroscopic growth corrections. The results show that using ambient $PM_{2.5}$ alone already yields a moderate linear correlation with AOD normalized by Mixing Layer Height (Pearson's $R = 0.56$) whereas no meaningful correlation is found when using standard dry $PM_{2.5}$. When aerosol microphysical properties derived from LOAC measurements are incorporated, the correlation substantially improves ($R = 0.76$), with regression slopes close to unity and reduced errors, independently of the season. These results demonstrate that explicitly accounting for aerosol size and optical properties enhances the physical consistency and robustness of satellite-based PM estimates. The proposed framework also provides a pathway to indirectly derive aerosol hygroscopic growth factors by coupling ambient PM estimates from satellite observations with conventional dry PM measurements. This opens new perspectives for characterizing aerosol–humidity interactions from space and for improving air quality monitoring in regions lacking of dense in situ networks.

Keywords: satellite remote sensing; particulate matter; aerosol optical depth; air quality monitoring

Check for updates

Received: 22 April 2026
Revised: 19 June 2026
Accepted: 22 June 2026
Published: 24 June 2026

1. Introduction

The health effects of aerosol pollution are well documented in the literature [1,2], with effects ranging from allergies to chronic respiratory diseases and even carcinogenic effects, as stated by the World Health Organization (WHO) in 2013 [3]. For this reason, the continuous monitoring of aerosol concentrations within urban areas and in background and remote locations is of primary importance for studies focused on exposure and risk assessment practices.

Aerosol monitoring is usually performed with conventional and standardized procedures that employ daily in situ measurements of particulate matter mass concentration (PM, in $\mu g/m^3$) at fixed stations. PM is defined according to particle size: PM_1, $PM_{2.5}$ and PM_{10} refer to the PM of a particle whose diameter is lower, respectively, than 1, 2.5 and 10 µm In the present work, only $PM_{2.5}$ is considered. According to Martin et al. [4], the European average is 3 to 4 PM monitors per million of inhabitants, while in other continents the station density is even more sparse. As a result, most surface areas are underrepresented in terms of the spatial and temporal characterization of PM. The Po Valley, which is the study area for the present work, has a denser monitoring network than the European average; nevertheless, spatial coverage remains insufficient to capture the full heterogeneity of PM concentrations at urban and local scales, supporting the need for satellite-based approaches.

In the few last decades, the scientific community has pursued the idea of using satellite observations to increase the time and space resolutions of the sparse ground-based network [5]. Satellite Aerosol Optical Depth (AOD) can be correlated to the ground-based PM by an approximate linear relationship if the aerosols are confined near the surface, such as in urban areas, where the aerosols are well-mixed below the Mixing Layer Height (MLH) [6,7]. For example, Wang and Christopher [6] found a linear correlation coefficient R equal to 0.7 (out of 1) between MODIS AOD and hourly $PM_{2.5}$; the correlation increased to values higher than 0.9 when lowering the time resolution to monthly averages. Koelemeijer et al. [7] considered several European cities and found a linear correlation coefficient of 0.6 between MODIS (Moderate-Resolution Imaging Spectroradiometer) AOD and hourly $PM_{2.5}$, over rural and background urban stations. Other studies obtained different results, depending on the location, time window and the type of linear model chosen to relate AOD and PM [8]. In general, multi-linear models perform worse when correlated variables are included without interaction terms, i.e., without product terms that capture the combined effect of two or more predictors. Multi-linear models that incorporate a variable selection algorithm, such as stepwise regression or regularization methods, to identify the most informative subset of predictors tend to perform better [9].

More complex methodologies employ data-driven models (statistical or machine learning) or chemistry transport models [10]. In the few last years, several studies trained ML models to estimate $PM_{2.5}$ from satellite AOD over Europe and world-wide (e.g., [11–16]). Each methodology has its strengths and weaknesses. For example, data-driven models require large datasets for the training phase, and identifying the most relevant predictor variables requires careful feature selection. On the other hand, the use of chemical transport models requires great computational power, and their accuracy depends on the quality of input data such as emission sources, land use, meteorology, etc. Lastly, semi-empirical methodologies employ measurements at selected locations to estimate the coefficients of the linear function that relates AOD and PM (e.g., [17]). This type of methodology presents low computational costs and is based on physical relationships, but the input data may be spatially and temporally sparse [18]. Nonetheless, a semi-empirical method can be implemented with some caution and tailored to the case study, depending on the available data. Ferrero et al. [19] considered factors such as the spatial representativeness of the ground-based PM with respect to the satellite footprint and the temporal difference between the satellite overpass time and the daily or hourly PM measurements, as well as the vertical characteristics of the atmosphere. Table A1 summarizes some notable results from multiple previous studies.

In the revised Ambient Air Quality Directive from the European Commission [20], the annual limit for $PM_{2.5}$ has been lowered to 10 μg/m^3, while the 24 h average limit is 25 μg/m^3, to be exceeded less than 18 times during the year. The Po Valley, in Northern Italy, is an area known for frequent exceedance of air pollution limit values, including those for aerosol pollution. This is due to its high urbanization and industrialization density, and it is exacerbated especially in the winter period by its peculiar geography bounded by the Alps and the Apennine mountains limiting atmospheric circulation, as well as by weather patterns characterized by stability conditions that often trap and accumulate pollutants for long periods.

In recent decades, MODIS AOD has been used with empirical, statistical and machine learning methodologies to estimate the PM spatial variability in the area. Koelemeijer et al. [7] and Di Antonio et al. [21] managed to highlight the contrast between the high atmospheric pollution in the Po Valley compared to nearby areas. The study by Di Nicolantonio et al. [22] was one of the first to implement a semi-empirical methodology, and found a correlation of 0.68 and 0.59 between MODIS AOD and daily $PM_{2.5}$, respectively, using data from MODIS Terra and MODIS Aqua, reaching values of 0.7 when using monthly averages. More recent studies [19,23,24] exploited MAIAC (Multi-Angle Implementation of Atmospheric Correction) AOD product [25], retrieved by MODIS observations with a spatial resolution of 1 km × 1 km, to better characterize the spatial variation of PM.

Among recent studies in the area, Ferrero et al. [19] and Arvani et al. [23] considered the use of a hygroscopic growth function within their methodologies. Hygroscopicity refers to the property of some aerosols to grow in size according to their water uptake, changing their optical properties as well as their mass concentration. A corrective function is needed when relating the AOD, which is retrieved at ambient humidity, and the PM, which is measured in controlled environments with RH < 50%. The use of this function should improve the correlation between the two quantities. The main issue is the empirical nature of this function, whose coefficients depend on the local aerosol characteristics [26,27]. Ferrero et al. [19] found that hygroscopic corrections did not significantly improve the AOD–PM correlation for any of the size fractions considered (PM_1, $PM_{2.5}$ and PM_{10}).

In recent years, Optical Particle Counters (OPCs) have been employed as part of air quality monitoring networks because of their portability, higher temporal resolution and limited costs with respect to conventional reference instrumentation utilized for PM measurement [28]. OPCs detect the particle number (PN) concentration in specific size bins and retrieve the PM mass concentration through an internal algorithm. If the OPC is operated at ambient conditions, it estimates the ambient or "wet" PM, which is supposedly better correlated with AOD than the "dry" PM from standard measurements. In fact, the hygroscopic function is not needed when relating "wet" PM with AOD at ambient conditions.

Several previous studies have combined OPC data with satellite AOD. deSouza et al. [29] used OPC size distributions to constrain MISR component-specific AOD in Nairobi, deriving surface PM concentrations by scaling satellite retrievals with ground-based size distribution data. Kim et al. [30] incorporated OPC measurements alongside lidar and sunphotometer data to improve empirical MODIS AOD-$PM_{2.5}$ correlations at a single site in Korea, using aerosol profile information to screen thin cloud and dust cases. Fawole et al. [31] estimated $PM_{2.5}$ from satellite AOD using a Bayesian regression model with planetary boundary layer height and Relative Humidity as inputs, relying on reference ground-based PM measurements rather than OPC-derived ambient PM. However, none of these studies use OPC-derived aerosol microphysical properties (Effective Radius, Extinction Efficiency, and Mass Density) to estimate a theoretical scaling coefficient that directly links ambient $PM_{2.5,wet}$ to satellite AOD, explicitly avoiding empirical hygroscopic corrections. The present study fills this gap.

The present study compares estimates of PM retrieved by a specific Optical Particle Counter, the LOAC (Light Optical Aerosol Counter), operated at ambient conditions, with MAIAC AOD for the year 2023 over the city of Bologna, in the Po Valley. The study also attempts to find a semi-empirical linear function to relate the "wet" $PM_{2.5}$ and the satellite AOD, with the use of auxiliary data. The use of a single, well-instrumented site allows for the minimization of spatial representativeness errors between satellite AOD and ground-based measurements, which are a major source of uncertainty in multi-site AOD–PM studies.

PN measurements from the LOAC are used as an estimate of the aerosol Particle Size Distribution (PSD), allowing for the computation of the Effective Radius of the aerosol mixture. The Effective Radius is an important variable in the correlation between the optical properties and the mass concentration of aerosols and strongly affects the radiative properties derived from satellite. In addition, the LOAC provides estimates of dominant aerosol type (such as salt, mineral, and carbonaceous aerosols) [32] which is used in the semi-empirical methodology together with Effective Radius and MLH.

The objective of this work is not to provide a spatially generalizable model, nor to develop an operational PM_{dry} estimation method, but to demonstrate that using ambient-humidity PM with co-located aerosol microphysical properties yields a physically consistent and direct relationship with satellite AOD, without relying on empirical hygroscopic corrections. The conversion from PM_{wet} to regulatory PM_{dry} using a hygroscopic growth factor is identified as a natural next step and left for future work. In contrast to previous multi-site statistical approaches, this study prioritizes the physical consistency of the AOD-$PM_{2.5}$ relationship over spatial coverage. The analysis is therefore conducted at a single, well-instrumented urban site (Bologna), enabling the consistent characterization of aerosols' microphysical properties and their impact on the AOD-PM relationship. Given the limited spatial and temporal extent of the dataset, this work should be regarded as a pilot study aimed at assessing the robustness and interpretability of the proposed methodology. Extending the analysis to multiple sites will be essential to evaluate its transferability across different environments.

2. Materials and Methods

2.1. Study Area

Ground-based observations used in the present work were obtained during an intensive experimental field campaign, conducted in 2023 in the urban area of Bologna, a city located in the southern border of the Po Valley. The classification of the measurement site is urban.

The Po Valley is characterized by a distinct seasonality of pollutant concentrations, including particulate matter. During winter, the frequent stability conditions and low wind velocity tend to favour the trapping and accumulation of pollutants [33]. Particle accumulation gives rise to a concentration gradient at the MLH level, which correlates with a vertical gradient of Relative Humidity [34]. The concentration gradient is also linked to a vertical gradient of aerosol extinction. During winter, aerosol extinction tends to increase in the lowest levels of the atmosphere and concurrently to decrease at altitudes higher than 1 km; during spring and summer, the extinction is evenly distributed in the atmospheric column. As a result, during winter, more than 80% of AOD refers to particles below 1 km [35].

The relative composition of PM also varies during the year according to changes in aerosol emissions. In total, 75% of urban $PM_{2.5}$ in the Po Valley is composed of water-soluble inorganic compounds (nitrates and sulphates) and carbon [36]. Local emission sources in cities comprise different types of fuel due to traffic and residential heating, which leads to different fractions of black carbon and organic compounds during the year [37]. In the city of Bologna, at least 40% of $PM_{2.5}$ is due to traffic emissions, 24% is composed of road dust, and 8% is due to biomass burning; the levels of mineral dust, which are usually negligible, often increase in the summer during episodes of desert dust intrusion [38]. A change in the composition of the aerosol mixture affects the average optical properties retrieved from remote sensing observations, such as satellite AOD. For this reason, when comparing the aerosol mass concentration and the total extinction, it is important to obtain information about the chemical composition or the optical behaviour of the aerosol mixture.

2.2. Methodology

Figure 1 shows the methodological flowchart followed in the present work. Each block refers to one or more Sections presented below, from input data to final statistical verification.

A linear correlation between aerosol extinction at visible wavelengths and $PM_{2.5}$ can be established based on theoretical considerations (e.g., [7]):

$$k_{ext,aer}(z=0) = PM_{2.5}\frac{3\langle Q_{ext,aer}\rangle}{4\rho_{aer}r_{eff}} \tag{1}$$

Equation (1) links the aerosol extinction coefficient $k_{ext,aer}$ at the ground level with the corresponding mass concentration $PM_{2.5}$. All the variables refer to ambient conditions. ρ_{aer} is the aerosol average Mass Density and r_{eff} is the Effective Radius, defined as follows:

$$r_{eff} = \frac{\int_0^\infty \pi r^3 n(r)dr}{\int_0^\infty \pi r^2 n(r)dr} \tag{2}$$

where $n(r)$ is the aerosol PSD, and the average Extinction Efficiency is defined as follows:

$$\langle Q_{ext,aer}\rangle = \frac{\int_0^\infty Q_{aer}(r)\pi r^2 n(r)dr}{\int_0^\infty \pi r^2 n(r)dr}. \tag{3}$$

where $Q_{aer}(r)$ is the single-particle Extinction Efficiency and, for fixed wavelength and aerosol type, is a function of the particle radius.

The AOD is defined as the integral of the aerosol extinction coefficient along the atmospheric column. Assuming that all aerosols in the atmospheric column are well-mixed and placed below the MLH, AOD can be defined as

$$AOD = \int_0^{TOA} k_{ext,aer}(z)dz \cong k_{ext,aer}(0)MLH \tag{4}$$

where TOA stands for Top of Atmosphere. The approximate formula relating $PM_{2.5,wet}$ (at ambient conditions) to AOD (at ambient conditions) is

$$PM_{2.5,wet} \cong \frac{AOD}{MLH}\cdot\frac{4\rho_{aer}r_{eff}}{3\langle Q_{ext,aer}\rangle} \tag{5}$$

The term $A = \frac{4\rho_{aer}r_{eff}}{3\langle Q_{ext,aer}\rangle}$ represents the theoretical linear coefficient between $PM_{2.5,wet}$ and AOD scaled by MLH; that is, $\frac{AOD}{MLH}$. All the variables in A can be estimated from LOAC measurements, so they are consistent with the $PM_{2.5,wet}$ measurements.

Equation (5) depends explicitly on MLH and implicitly on RH. Other meteorological variables such as wind speed, wind direction, solar radiation, and surface pressure do not appear in the theoretical derivation and are not required for the proposed scaling at a single urban site. Their inclusion would be relevant for spatial extrapolation or chemical transport modelling, which are beyond the scope of this study.

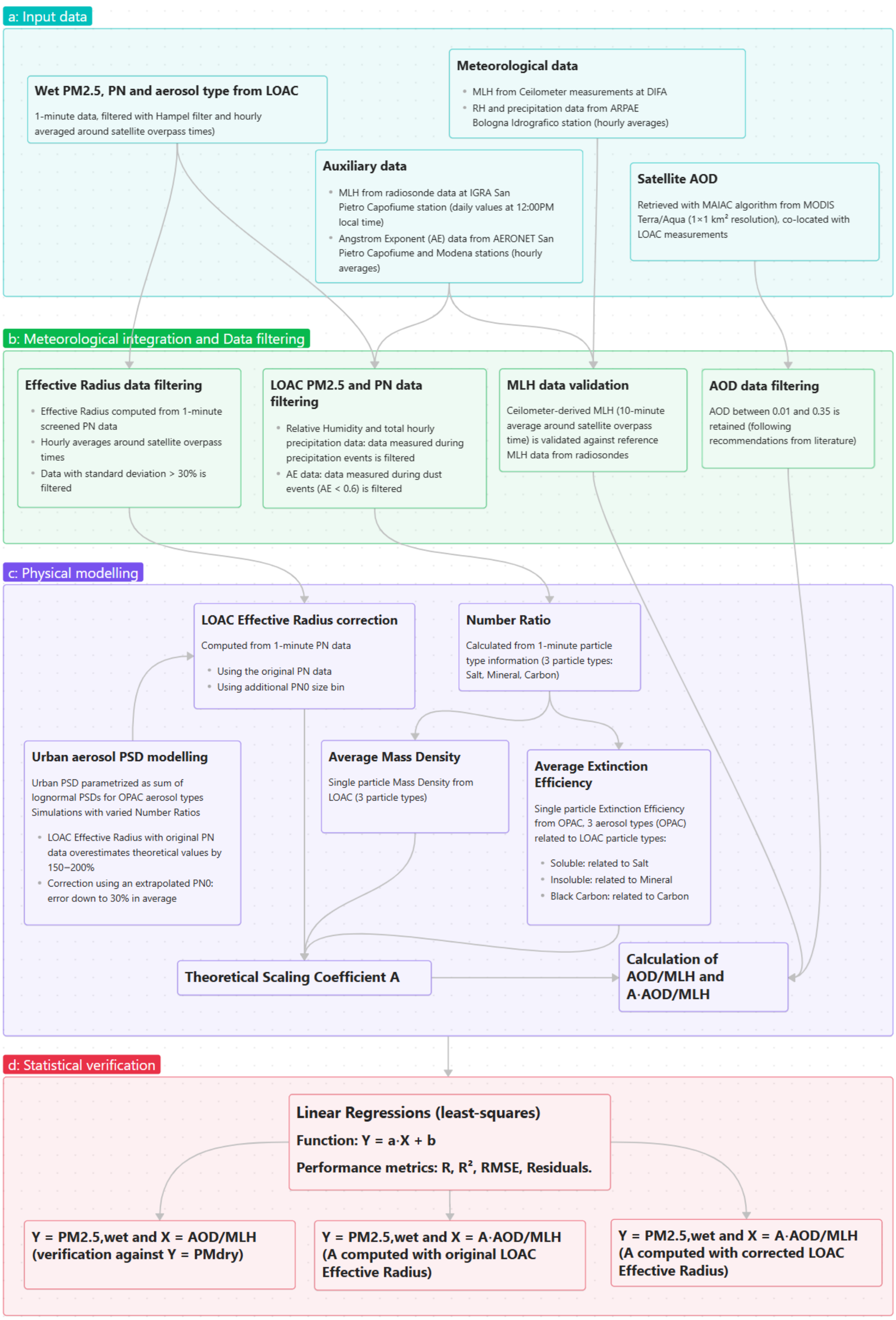


Figure 1. Methodological flowchart illustrating the physics-informed framework. Block (**a**) comprises input data, as explained in Sections 2.3–2.5. Block (**b**) specifies data filtering, as explained in Sections 3.1 and 3.3. Block (**c**) represents the physical modelling and calculation of relevant aerosol physical properties, as explained in Sections 2.2 and 3.2. Block (**d**) shows the statistical verification and results of linear regressions, as explained in Section 3.3.

As mentioned previously, conventional ground-based PM measurements are performed under dry condtitions ($RH < 50\%$) to comply with regulatory standards. The relationship between PM at ambient and at dry conditions is defined as

$$PM_{2.5,dry} = \frac{PM_{2.5,wet}}{f(RH)} \tag{6}$$

where $f(RH)$ is the hygroscopic growth function that depends on the local aerosol composition. As such, $f(RH)$ is usually derived empirically using co-located measurements of aerosol optical properties in ambient and dry conditions [27].

In the present study, the LOAC-derived $PM_{2.5,wet}$ is compared to $\frac{AOD}{MLH}$, consistent with previous studies in the literature. Secondly, the LOAC $PM_{2.5,wet}$ is compared with the $PM_{2.5,wet}$ estimated using Equation (5). Since the relationship is approximated as linear, the comparisons are performed through a linear regression (least-squares fit) with the following function:

$$Y = a \cdot X + b, \tag{7}$$

with $Y = PM_{2.5,wet}$ and $X = \frac{AOD}{MLH}$ in the first case, and $X = \frac{4\rho_{aer} r_{eff}}{3\langle Q_{ext,aer}\rangle} \frac{AOD}{MLH}$ in the second case. The goodness of fit is evaluated using Pearson's correlation coefficient R (95% confidence) and the linear regression performance is measured by the coefficient of determination R^2 and the Root-Mean Square Error (RMSE). By using a linear function, this methodology is able to describe the average linear correlation between PM and AOD, but does not account for possible non-linear and long-term effects.

Theoretically, the value of the slope a in the first case varies throughout the year according to the varying aerosol mixtures and the meteorological conditions that impact the location of interest. In general, the higher the MLH, for example during summer, the lower the linear correlation (e.g., [39]).

In the first case, the coefficient A is considered to be constant. In the second case, the value of A for each point is estimated, and the slope a is close to one if $X = \frac{4\rho_{aer} r_{eff}}{3\langle Q_{ext,aer}\rangle} \frac{AOD}{MLH}$ is a good estimate of $PM_{2.5,wet}$. In both cases, b is the offset between $PM_{2.5,wet}$ and AOD, and, in theory, it does not vary seasonally.

Lastly, the variable importance assessment of r_{eff} and the other LOAC-derived variables is performed by comparing the correlation of $PM_{2.5,wet}$ with $\frac{AOD}{MLH}$, $\frac{AOD}{MLH} r_{eff}$, $\frac{r_{eff}}{\langle Q_{ext,aer}\rangle} \frac{AOD}{MLH}$, and $\frac{4\rho_{aer} r_{eff}}{3\langle Q_{ext,aer}\rangle} \frac{AOD}{MLH}$.

2.3. OPC Data

OPC measurements were obtained under ambient RH conditions by a LOAC [32] (manufactured by Meteomodem, France) mounted on the rooftop of the Department of Physics and Astronomy (DIFA) of the University of Bologna (lat: 44.499 N, lon: 11.354 E), at about 22 m above the ground (from now on named only DIFA). The surrounding area is a vegetated street inside the historical centre of Bologna. Possible local emission sources are trees and other vegetation (pollen and other biological particles), the low levels of vehicle traffic from Irnerio Street and the surrounding roads (road dust and combustion products), building heating systems during winter (biomass burning), and occasional desert dust events (dust particles). The LOAC dataset covers the months from February to September 2023.

The LOAC retrieves the particle number concentration (number of particles per cm^3, $\#/cm^3$ or simply cm^{-3}) in 19 size bins for particle diameters between 0.2 and 30 µm (Table 1) by the use of laser beams at a 650 nm wavelength. By using two scattering angles (12° and 60°), it also determines the degree of absorption by the detected particle, from transparent to highly absorbing. This estimate is used to classify the particles, divided in five size classes

(Table 2), in four types: droplet, salt, mineral, and carbon. From the PN measurements and the particle type, PM_1, $PM_{2.5}$, and PM_{10} (µg/m^3) are calculated with an internal algorithm which uses a suitable value of particle Mass Density (g/cm^3) according to particle type. In general, all OPC instruments are highly affected by high Relative Humidity, cloudy conditions and precipitation events (e.g., [40]). As such, recent studies recommend to perform measurements with such sensors under low time resolutions (hourly and daily averages) and clear sky conditions [41].

Table 1. Definition of LOAC size bins, respective bin edges, and average diameters.

Size Bin	Bin Edges (µm)	Bin Diameter (µm)
1	0.2–0.3	0.25
2	0.3–0.4	0.35
3	0.4–0.5	0.45
4	0.5–0.6	0.55
5	0.6–0.7	0.65
6	0.70–0.85	0.775
7	0.85–1	0.925
8	1–2	1.5
9	2–3	2.5
10	3–5	4
11	5–7.5	6.25
12	7.5–10	8.75
13	10–12.5	11.25
14	12.5–15	13.75
15	15–17.5	16.25
16	17.5–20	18.75
17	20–22.5	21.25
18	22.5–25	23.75
19	25–30	27.5

Table 2. Size classes of LOAC particle types and respective diameter edges (D).

Size Class	1	2	3	4	5	6
D (µm)	0.2–0.5	0.5–0.9	0.9–5	5–12.5	12.5–20	20–40

The 1 min PN and $PM_{2.5,wet}$ data are screened with a Hampel filter on a 10 min window [42] to remove outliers and then are averaged hourly. The Effective Radius r_{eff} is also computed by discretizing Equation (2), which becomes

$$r_{eff} \cong \frac{\sum_i PN_i R_i^3}{\sum_i PN_i R_i^2} \tag{8}$$

where PN_i is the particle number concentration in the size bin i and R_i is half the nominal diameter of size bin i. This is because the particle number concentration in size bin i is the total value of $n(r)$ within the following size bin: $PN_i \cong \int_{D_i/2}^{D_{i+1}/2} n(r)dr$. The r_{eff} is computed with the original filtered 1 min PN data and then averaged hourly.

The particle type dataset from LOAC is used together with PN data to compute the average aerosol Extinction Efficiency and Mass Density. This is achieved by computing the following:

- The Number Ratio (NR) of each particle type in each size bin;
- The average single-particle Extinction Efficiency of each particle type in each size bin.

Additionally, the proper Mass Density for each particle type is chosen. In Renard et al. 2016 [32], it is advised to use particle type information with a statistical approach. For each hour, the number of occurrences of each particle type x in each size bin i is divided by the total number of observations in each size bin in that hour (maximum of 60); the fraction is then converted into decimal percentage. This percentage defines the hourly Number Ratio $F_{i,x}$ in size bin i for the particle type x, where x can refer to salt (x = 1), mineral (x = 2) or carbon (x = 3).

In Renard et al. 2010 [43], the absorption of different particle types is defined in relation to different values of the imaginary part k of the Refractive Index (RI). The salt and mineral particle types refer to particles with low absorption at 650 nm, while the carbon particle type refers to particles with high absorption at the same wavelength. In the database OPAC (Optical Properties of Aerosols and Clouds) [44], the urban aerosols are defined as a mixture of soluble and insoluble aerosols, characterised by low absorption at 650 nm, and of black carbon (or soot) which has a high absorption at 650 nm. In this study, the assumption made is that each LOAC particle type can be related to each OPAC aerosol type:

- The salt LOAC type related to the "soluble" OPAC type (here considered at RH = 50%);
- The mineral LOAC type related to the "insoluble" OPAC type;
- The carbon LOAC type related to the "soot" OPAC type.

The respective single-particle Extinction Efficiency $Q_{aer}(r)$ values are taken from OPAC. This choice is made taking into account similar values of k between the LOAC and OPAC aerosol types. $Q_{i,x}$ is computed as the average value of $Q_{aer}(r)$ in the size bin i for type x. The average Extinction Efficiency for each LOAC observation is discretized from Equation (3) as

$$\langle Q_{ext,aer} \rangle \cong \frac{\sum_i (Q_{i,1}F_{i,1}PN_i + Q_{i,2}F_{i,2}PN_i + Q_{i,3}F_{i,3}PN_i)R_i^2}{\sum_i PN_i R_i^2} \tag{9}$$

where $F_{i,x}$ denotes the previously defined Number Ratio in the size bin i for particle type x. Similarly, the average Mass Density ρ_{aer} is computed as

$$\rho_{aer} \cong \frac{\sum_i PN_i(m_1F_{i,1} + m_2F_{i,2} + m_3F_{i,3})}{\sum_i PN_i} \tag{10}$$

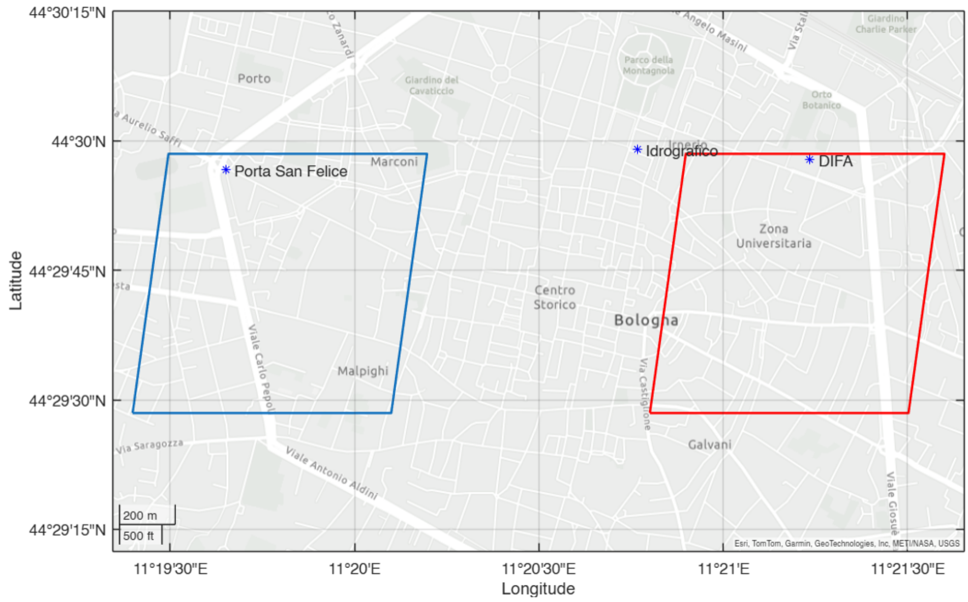
where m_x are the nominal values of Mass Density for each LOAC particle type x (same value for all size bins). The single-particle Mass Density values from the LOAC algorithm are used in this computation, to be consistent with the PM calculation.

2.4. AOD

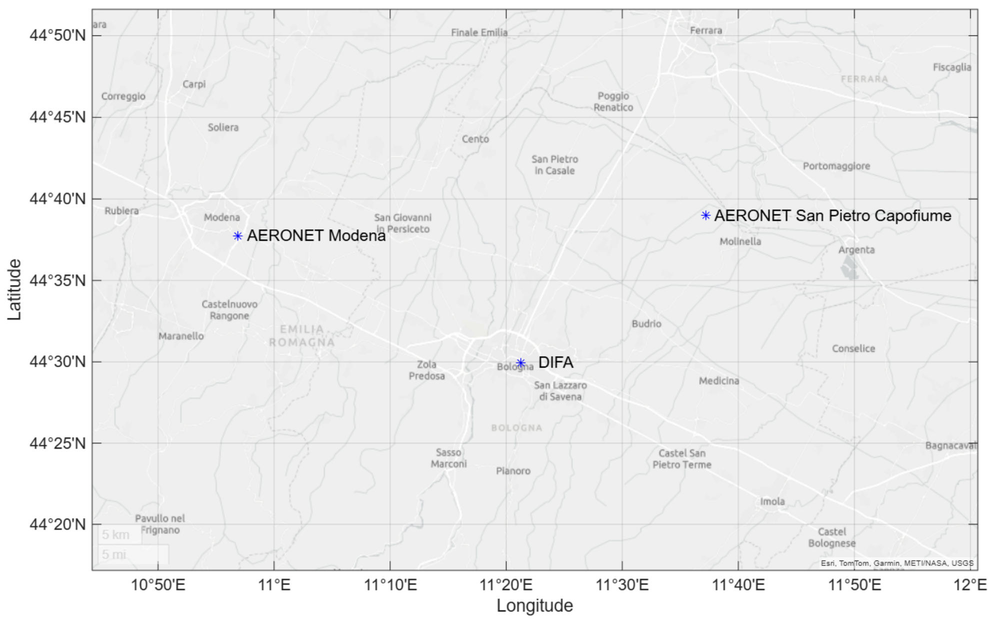
MAIAC AOD is retrieved at 550 nm from MODIS observations at high spatial resolution (1 km × 1 km) on a fixed grid. It has a better performance over bright surfaces compared to other algorithms and has been previously applied over urban areas [45]. In cases of anthropogenic aerosols, an AOD below 0.12 can be overestimated. AOD uncertainty indicated in the product files is related to surface brightness. A value below 0.05 allows the algorithm to perform a standard retrieval [45,46]. A 5% uncertainty value for the single retrieval values can be used. More information on the retrieval can be found at [25]. Intense pollution cases in urban areas within the Po Valley, especially in very high polluted cities such as Milan and Turin, can be sometimes interpreted as clouds and filtered out [23].

The co-location of AOD and ground-based data at DIFA site is performed on the fixed grid used in MAIAC retrievals. The cell containing the DIFA location is selected (Figure 2a). This grid cell refers to the innermost part of Bologna historical centre, characterised by medium-bright surfaces, traffic and low industrial emissions.

The AOD data is acquired for the whole period covered by the LOAC dataset (February–September 2023). AOD retrieval requires cloud-free conditions, limiting data availability. Of the 985 satellite overpasses during the study period, only 26.8% yielded successful retrievals. The retrieval success varied seasonally, ranging from 15–16% in April–May to 41% in August, reflecting regional cloud climatology and precipitation patterns.



(**a**)



(**b**)

Figure 2. (**a**) Location of the ground-based measurements (blue markers), the AOD cell selected for DIFA (red box) and the AOD cell selected for Porta San Felice (blue box). (**b**) Location of the DIFA site relatively to AERONET ground-based stations of Modena and San Pietro Capofiume (blue markers). The IGRA ground-based station of San Pietro Capofiume overlaps with the corresponding AERONET station at this scale. Base map (Bologna, Italy) courtesy of Esri, TomTom, Garmin, GeoTechnologies, Inc., METI/NASA, USGS.

Considering the dataset of co-located AOD retrievals, AOD uncertainty due to surface brightness is lower than 0.045, with a median of 0.016, and AOD values are lower than 0.45, with a median of 0.14. This means that the retrievals used in the following analysis are computed with the standard algorithm, and are likely not affected by overestimation.

2.5. Auxiliary Variables

The Mixing Layer Height is retrieved by a ceilometer Vaisala CL31 (manufactured by Vaisala, Finland) present at DIFA alongside the LOAC. The ceilometer continuously measures the vertical backscatter signal from the ground level to a maximum of 7.5 km and provides estimates of MLH with a temporal resolution of 16 s and a vertical resolution of 10 m. It also provides an estimate of Cloud Height if clouds are detected. The internal algorithm is based on the difference in backscatter signal between the mixing layer, where the signal is higher due to aerosols, and the free atmosphere, where the backscatter signal abruptly drops to very small values in absence of scattering layers (gradient method) [47]. The gradient method estimates the MLH as the height of the greatest negative gradient in the aerosol backscatter signal [48,49].

The ceilometer detects one to three possible MLH values; the highest one usually corresponds to residual layers or incorrectly detected clouds [50,51]. The gradient method requires a robust way to choose the correct MLH value. In the present study, MLH data obtained from radiosondes at the nearest IGRA (Integrated Global Radiosonde Archive) station [52,53] are used as reference. Radiosondes measure the vertical profiles of temperature, humidity, pressure and wind speed with high precision, and provide indirect estimates of MLH. The reference radiosonde measurements are taken at the IGRA station of San Pietro Capofiume (lat: 44.654, lon: 11.623), located 27.4 km northeast of Bologna (Figure 2b), and have a climatology of more than 30 years. The averages of IGRA MLH at 12:00 (local time) between 2021 and 2025 are considered as a reference. The IGRA MLH average for the day is chosen as a variable threshold to select the correct ceilometer MLH values in the same day. The ceilometer dataset, composed of all possible MLH values, is averaged on a 10 min window around the satellite overpass times. In general, the second and third MLH values are higher than the reference IGRA MLH average at the same time and day. The closest ceilometer MLH to the average radiosonde MLH in the same calendar day is chosen as MLH for the present study.

Auxiliary data include quality-controlled data from the local environmental agency (ARPAE) [54]: the 30 min Relative Humidity and hourly Precipitation Rate are measured from the closest meteorological station, Bologna Idrografico (lat: 44.499, lon: 11.346), located 620 m from DIFA; daily "dry" $PM_{2.5}$ from ARPAE Porta San Felice station (lat: 44.499, lon: 11.328), located 2120 m from DIFA. RH data is used in the comparison between $PM_{2.5}$ and AOD to verify that there is no dependence on humidity for both variables, while the "dry" $PM_{2.5}$ is used to compare the results of the linear regressions when using reference data and when using LOAC data. The stations are shown in Figure 2a alongside DIFA and the AOD cells from satellite, with the specific locations marked with blue asterisk symbols.

Lastly, data from AERONET (Aerosol Robotic Network) [55], a global network of ground-based sunphotometers that measure the AOD and related variables [56], is considered. The nearest AERONET stations (Figure 2b) are in the urban area of Modena (lat: 44.629, lon: 10.948), located 35.26 km northwest of Bologna, and in the rural area of Bologna at San Pietro Capofiume (lat: 44.650, lon: 11.620), located 26.94 km northeast of Bologna and 504 m northeast of the IGRA station. AERONET data provides useful information about aerosol columnar properties to complement the data analysis.

2.6. Uncertainties and Data Preprocessing

According to Lyapustin et al. [25], the relative uncertainty of AOD retrieved by the MAIAC algorithm is 0.05, so the uncertainty assigned to AOD is $\Delta AOD = 0.05 \cdot AOD$. The MLH uncertainty is the maximum between the ceilometer measuring uncertainty of 10 m and the standard deviation of the 10 min average. The uncertainty assigned to $X = \frac{AOD}{MLH}$ is calculated with the relative uncertainty rule, as follows:

$$\frac{\Delta X}{X} = \sqrt{0.05^2 + \left(\frac{\Delta MLH}{MLH}\right)^2}. \quad (11)$$

The uncertainty assigned to $A\frac{AOD}{MLH}$ is computed in a similar way.

As the $PM_{2.5}$ and Effective Radius are averaged hourly from the 1 min LOAC data, and the standard deviation is assigned as uncertainty for both variables. The AOD retrieval time is set as the dataset index. The hourly dataset is composed of AOD, MLH, $PM_{2.5,wet}$, r_{eff}, $\langle Q_{ext,aer} \rangle$, ρ_{aer}, A, and RH. All variables except AOD are temporally averaged around the time of satellite overpass.

To ensure consistency and avoid biases introduced by data gaps, only complete rows are retained for the linear regressions and the computation of statistical coefficients. This approach reduces the sample size, but preserves the physical consistency of co-located measurements, which is essential to the presented physics-informed approach. Imputation methods were not employed, as they would introduce artificial correlations between variables and undermine the direct physical relationship between AOD and ground-based PM that this study aims to characterise.

Noise and outliers in LOAC data are addressed using a Hampel filter applied to the LOAC data, as previously mentioned. To minimise the mismatch between the point measurements and the 1×1 km^2 area represented by the MAIAC grid cell, LOAC data is averaged hourly around the time of satellite overpass, while MLH data is averaged in a 10 min window around the time of satellite overpass, as stated above. The location of DIFA site is close to the border of the selected MAIAC grid cell, which could introduce biases related to different emission sources in the adjacent grid cells. However, as stated in Lyapustin et al. [25], AOD values are averaged in a 3×3 grid, and it is safe to assume that this spatial averaging accounts for possible mismatch due to the DIFA site location relative to the grid cell border.

A set of data selection criteria is designed to minimize confounding factors in the AOD-PM relationship. These include the following: (i) co-located measurements; (ii) the temporal matching of ground-based data to satellite overpass times; (iii) exclusion of precipitation events; (iv) the exclusion of desert dust intrusion events (Angstrom Exponent < 0.6 from AERONET); (v) uncertainty filtering (relative uncertainty < 20% to relate $PM_{2.5,wet}$ and $\frac{AOD}{MLH}$; absolute uncertainty to relate $PM_{2.5,wet}$ measured by the LOAC and estimated from AOD and auxiliary data); and (vi) outlier removal using a Hampel filter on LOAC data. These constraints ensure that the remaining data satisfy the theoretical assumptions of the framework (well-mixed boundary layer, clear sky conditions, typical urban aerosol conditions) and that the observed correlations reflect physical relationships rather than artifacts.

3. Results

3.1. Dataset Analysis

The measurement period considered in this work (February–September 2023) was characterized by a particular variability in the meteorological conditions impacting the study area. The month of May 2023 was affected by three exceptional rainfall events (2, 16, and 17 May), with cumulative daily rainfall values higher than 100 mm recorded at the Bologna Idrografico station. Other minor rainfall events occurred throughout the measurement period. Figure 3 shows daily averages of $PM_{2.5,wet}$ from LOAC and RH from Bologna Idrografico for the whole period. Precipitation events with rainfall greater than 20 mm are highlighted in grey. As noted in [41], LOAC measurements are affected by the hygroscopic growth of particles under precipitation events and high RH conditions. This is apparent from the daily values of $PM_{2.5,wet}$ recorded during the study period, which are all below 35 μg/m^3 except for observation records during precipitation events, in which $PM_{2.5,wet}$ reaches values above 50 μg/m^3.

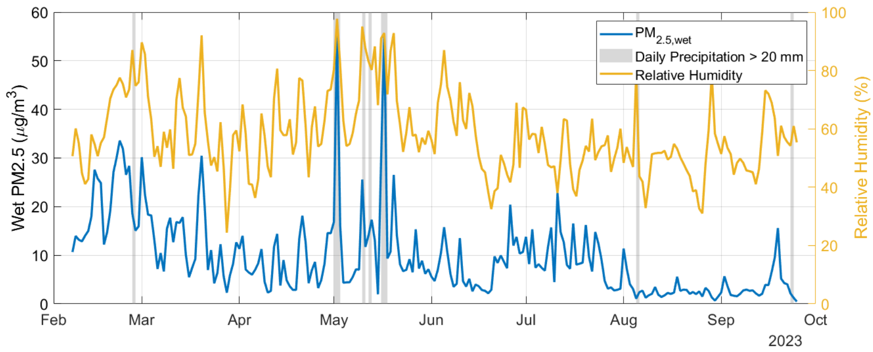
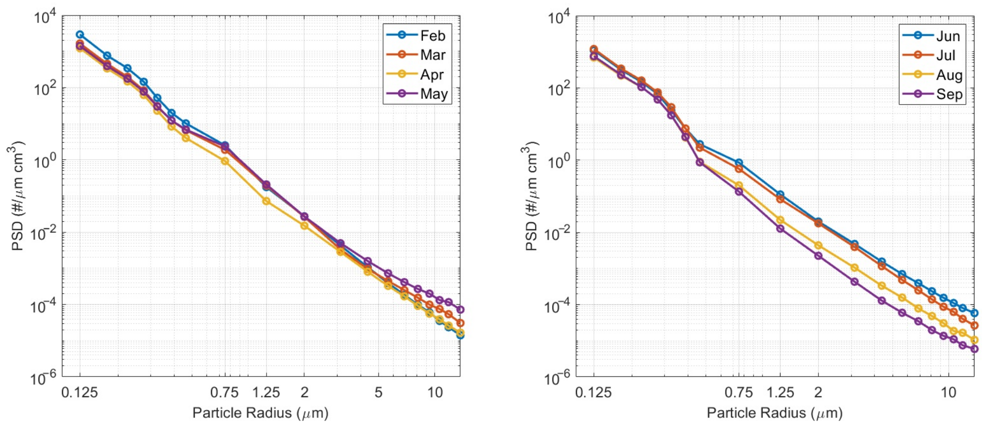


Figure 3. Daily "wet" $PM_{2.5}$ from LOAC (blue line) and daily RH from Bologna Idrografico (yellow line). Grey areas refer to days with total precipitation higher than 20 mm.

Desert dust intrusions were frequent during the months of June and July. The monthly PSD averages in Figure 4 show a corresponding increase in particles with a radius greater than 1 μm, compared to particles with smaller radii during these months. The PSD is computed by dividing the PN by the width of each size bin. The Effective Radius, computed with Equation (8), shows values larger than 2 μm during precipitation events on 2 and 16 May, and higher than 1.5 μm between 19 and 22 June, and between 9 and 11 and 15 and 18 July (Figure 5). Data from AERONET near the stations of Modena and San Pietro Capofiume highlight that, during that time, the Ångström Exponent α was below 0.5 in June and below 0.6 in July. Low values of α are associated with desert dust transport which is frequent in Italy during summer, with dust plumes that also reach the Po Valley and Bologna [57]. Except for precipitation and dust transport events, the Effective Radius is constantly smaller than 0.6 μm during the measurement period.



(**a**) (**b**)

Figure 4. Monthly PSD averages computed from LOAC PN measurements (logarithmic scales for x and y axis). (**a**) February to May. (**b**) June to September.

Figure 6 shows the monthly particle type occurrences, described by the Number Ratios as defined in Section 2. Only the first three size ranges are shown, consisting of particles with diameters between 0.2 and 5 μm (as specified in Table 2). The salt and mineral types are present in the first two size ranges in varying proportions during the year. In the third size range (0.9–5 μm), the mineral type is the most common, but also the carbon type is consistently present. Particles in the fourth size range (5–12.5 μm) consist of water droplets or unidentified types. Particle types in the two largest size ranges (12.5–40 μm) are mostly not identified, probably due to measured PN concentrations lower than 0.01 #/cm^3.

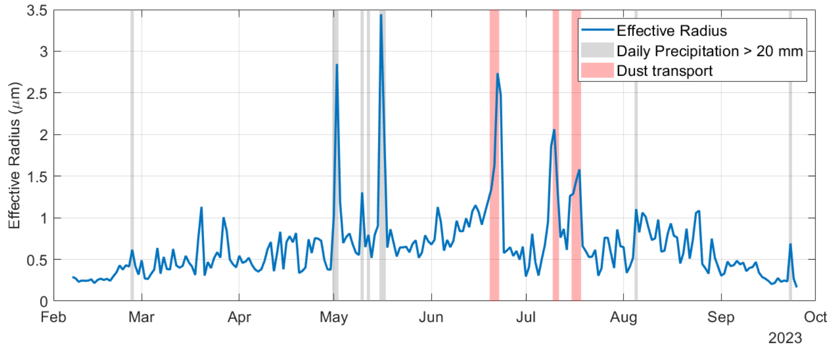


Figure 5. Daily Effective Radius from LOAC (blue line). Grey areas refer to days with total precipitation greater than 20 mm. Red areas refer to dust transport events.

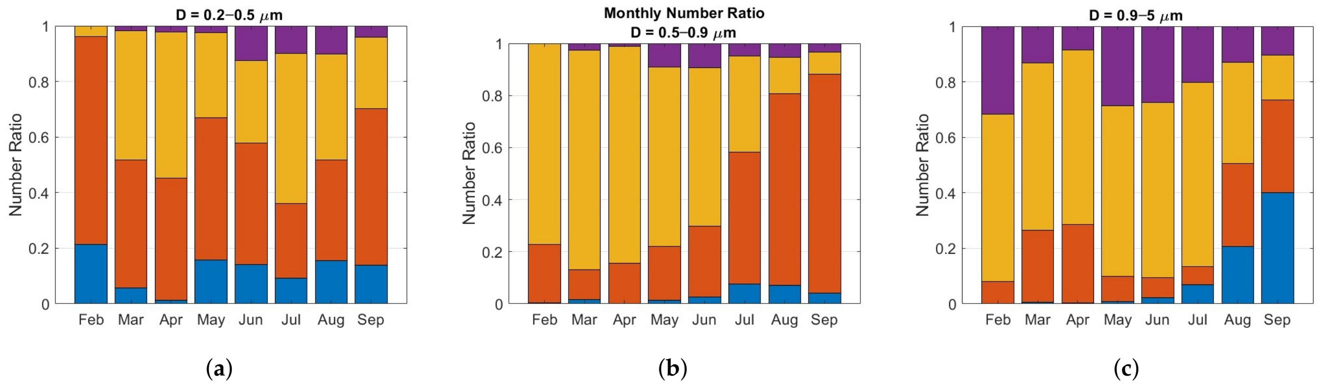


Figure 6. Monthly Number Ratio of the detected particle types in the first three size classes. Blue = droplet, red = salt, yellow = mineral, and purple = carbon. (**a**) Size class 0.2–0.5 μm. (**b**) Size class 0.5–0.9 μm. (**c**) Size class 0.9–5 μm.

3.2. Calculation of Aerosol Properties

Effective Radius. The smallest particle diameter detected by the LOAC is 0.2 μm (Table 1), so that particles with radii less than 0.1 μm are not detected. These particles are almost negligible in the computation of mass concentration, but have a significant contribution in terms of particle number concentration and in the extinction properties. Moreover, the Effective Radius should be calculated considering smaller particles, with radii starting, for example, from 0.05 μm [44]. The same is valid for the computation of $\langle Q_{ext,aer} \rangle$. Before computing the theoretical linear coefficient, the error made by neglecting particles smaller than 0.1 μm in radius should be assessed.

A series of numerical simulations are performed using an original code written in MATLAB (Version 23.2). The aim is to compare a set of "theoretical PSDs" and "LOAC PSDs" with the same aerosol mixtures, and to compare their respective Effective Radii. The code calculates the urban aerosol PSD of as a superposition of three lognormal functions, representing the three aerosol types described in Section 2.3. The resulting PSD, from now on referred to as "theoretical PSD", is the sum of the salt, mineral and carbon PSDs, each multiplied by the respective Number Ratio. The theoretical PSD is computed for particle radii between 0.05 μm and 20 μm. Using this PSD, the code computes the "theoretical Effective Radius" from Equation (2). As a second step, the code uses the theoretical PSD to compute the respective LOAC PN values in each size bin, therefore obtaining the approximated "LOAC PSD", similar to those shown in Figure 4. From the LOAC PSD, the LOAC Effective Radius is computed by using Equation (8).

The comparison between theoretical and LOAC Effective Radii values shows that the LOAC r_{eff} overestimates the theoretical value by a factor between 1.5 and 2, depending on the ratio between the fine and coarse particles used in the simulations. Therefore, the error made on the Effective Radius computed from LOAC PN data without corrections can be as high as 200%. To reduce this error, a simple correction algorithm is applied on the simulations, to quantify the improvement, and then on the actual LOAC data.

Firstly, considering the same set of PSD simulations, an additional size bin, referred to as PN_0, is defined as the PN of particles between 0.1 and 0.2 µm. The value of PN_0 is extrapolated from the LOAC PSD, and specifically from the first four size bins. Secondly, the LOAC Effective Radius is computed again, considering the additional PN_0 size bin. The theoretical PSD is not used in this process, so the algorithm can be applied to real LOAC data. Using the same set of simulations from the previous analysis, the calculations show that the corrected Effective Radius (r'_{eff}) overestimates the theoretical value by only a factor of 0.3 on average. This means that by applying the correction, the error made by computing the Effective Radius from LOAC data changes from approximately 100% to 30% of its theoretical value.

The last step is to apply the same correction to LOAC data. The corrected r'_{eff} values are lower than the original r_{eff} values by a factor between 50% and 90%. The calculated hourly averages of LOAC Effective Radius (both original and corrected) have an uncertainty of 20% on average, rising above 50% during precipitation and desert dust events.

Extinction Efficiency and Mass Density. The average quantities $\langle Q_{ext,aer} \rangle$ and ρ_{aer} are calculated from the hourly Number Ratio, hourly PN, and single-particle $Q_{aer}(r)$ and m_x respectively, applying Equations (9) and (10). Water particles are excluded from the calculation by setting to zero the Number Ratio of the droplet particle type, so the calculated values are not affected by precipitation events, unlike the Effective Radius; desert dust intrusions do not seem to affect them either.

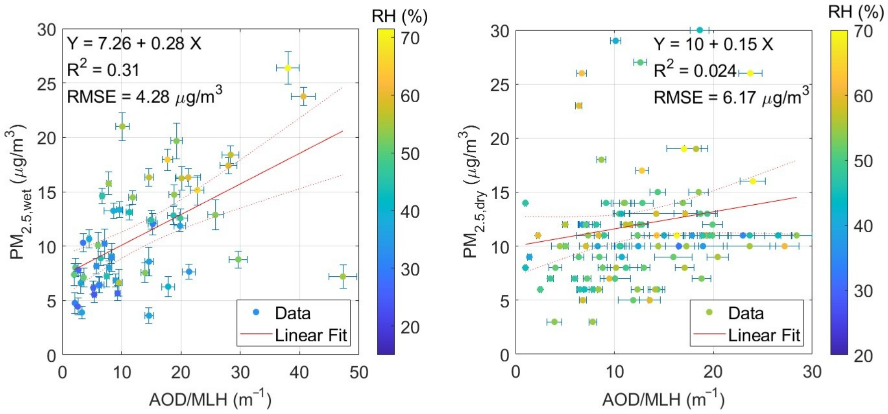
Linear coefficient. Lastly, the coefficient A is computed using all the estimated aerosol properties. A is normalized considering the different units of measurement of $PM_{2.5}$ (µg/m^3), $\frac{AOD}{MLH}$ (m^{-3}), r_{eff} (µm) and ρ_{aer} (g/cm^3). $\langle Q_{ext,aer} \rangle$ is unitless. The variability of A is mostly affected by the variability of the Effective Radius.

3.3. Methodology Application

3.3.1. $PM_{2.5}$ vs. $\frac{AOD}{MLH}$

The linear regression performed uses $Y = PM_{2.5,wet}$ and $X = \frac{AOD}{MLH}$. Data filtering selects $PM_{2.5,wet}$ data with standard deviations less than 3 µg/m^3 and relative uncertainty lower than 20%, AOD data between 0.01 and 0.35, to avoid errors in the retrieval algorithm, and $\frac{AOD}{MLH}$ with a relative uncertainty lower than 20%. The relative uncertainty threshold is set at 20% considering average LOAC data quality under stable conditions and ensuring that quantities represented by different physical units have comparable uncertainties. By using these rules, May and September data are filtered out, as well as several data points during June and July. The final dataset consists of a total of 56 observations available.

Pearson's correlation coefficient R is 0.56 (out of 1), indicating the presence of a linear correlation between the two variables. The linear regression function has $RMSE = 4.28$ µg/m^3 and $R^2 = 0.31$. The linear function is $Y_1 = 7.26 + 0.28 \cdot X$. The residuals are normally distributed around 0, with a median of –0.06. The linear function is shown in Figure 7a with the selected data points. As expected, the relationship between the variables does not show any clear dependence on RH, suggesting that there is no need for a hygroscopic correction factor.



(**a**) (**b**)

Figure 7. Results from the linear regressions. The dotted lines represent the 95% confidence bounds. (**a**) Linear regression between $Y = PM_{2.5,wet}$ and $X = \frac{AOD}{MLH}$. (**b**) Linear regression between $Y' = PM_{2.5,dry}$ and $X_{day} = \frac{AOD}{MLH}$ (daily averages). In the two cases, data points are color-coded with the average hourly and daily RH respectively.

For comparison, the linear regression is repeated using daily "dry" $PM_{2.5}$ from the reference ARPAE station Porta San Felice (SF), in the same temporal window. As $PM_{2.5}$ data is obtained from reference instruments, the uncertainty is assumed to be always below 20%. The MLH is assumed not to change significantly between SF and DIFA; therefore, the ceilometer-derived MLH data is used. The selected AOD refers to the cell whose centre is closest to the ground-based station of Porta San Felice, where $PM_{2.5,dry}$ is measured (blue box in Figure 2a). The AOD values assigned to DIFA and SF are retrieved from the same satellite swath, and as such refer to the same satellite overpass time. For comparison with daily $PM_{2.5,dry}$ data, daily estimates of $\frac{AOD}{MLH}$ are computed as follows. Firstly, the instantaneous $\frac{AOD}{MLH}$ values for each day are computed from the AOD values at satellite overpass time and the respective 10 min MLH averages. Secondly, the resulting $\frac{AOD}{MLH}$ values are averaged daily.

In this case, the regression is performed on 102 data points between February and September 2023. The value of R is 0.1558, indicating the absence of linear correlation. As is apparent from Figure 7b, "dry" $PM_{2.5}$ and the corrected AOD seem uncorrelated, and there is no correlation with the respective RH value (daily averaged). The linear function is $Y_1' = 10 + 0.15X_1'$, with $R^2 = 0.0243$ and $RMSE = 6.17$ μg/m^3.

To assess whether the higher correlation of $PM_{2.5,wet}$ with AOD is a physical result rather than an artifact of temporal resolution or filtering, two additional analyses are performed. First, daily averages of $PM_{2.5,wet}$ and $\frac{AOD}{MLH}$ over DIFA are compared directly, yielding a dataset of $n_{Obs} = 38$ observations. The linear regression of daily $\frac{AOD}{MLH}$ with daily $PM_{2.5,wet}$ results in $R = 0.51$, $R^2 = 0.26$, and $RMSE = 4.33$ μg/m^3. Despite the reduced temporal matching quality, the correlation with $PM_{2.5,wet}$ remains substantially higher than the correlation with $PM_{2.5,dry}$, particularly in terms of the RMSE. Second, to address the difference in sample size between the two datasets, the daily PM_{dry} regression is restricted to the same $n = 38$ days used in the daily $PM_{2.5,wet}$ analysis. This yields $R = 0.49$, $R^2 = 0.24$, and $RMSE = 7.71$ μg/m^3.

To verify that this result is not an artifact of the specific day selection, 50 random subsets of $n_{Obs} = 38$ data points are drawn from the full PM_{dry} dataset ($n_{Obs} = 102$). The resulting R values are distributed around $\bar{R} = 0.39$ ($\sigma = 0.12$), $\bar{R}^2 = 0.15$ ($\sigma = 0.12$), and $\overline{RMSE} = 5.5$ μg/m^3 ($\sigma = 1.1$ μg/m^3). The fact that the 38-day subset selected via $PM_{2.5,wet}$ filtering performs better than the average random subset ($R = 0.49$ vs. $\bar{R} = 0.39$) confirms that the day selection is not random: it preferentially retains days with low hourly

PM variability, characterized by clear sky conditions, for which the AOD-PM relationship is physically better constrained. This information is accessible only through the use of hourly $PM_{2.5,wet}$ data. These results confirm that the superior performance of $PM_{2.5,wet}$ is a physical consequence of using ambient-humidity measurements, not an artifact of filtering or temporal resolution.

3.3.2. $PM_{2.5}$ vs. $A\frac{AOD}{MLH}$

The linear regressions between $PM_{2.5,wet}$ measured by LOAC and $PM_{2.5,wet}$ calculated using Equation (5) are performed by considering $X = \frac{4\rho_{aer} r_{eff}}{3\langle Q_{ext,aer}\rangle}\frac{AOD}{MLH}$. As the comparison refers to the same physical quantity, an absolute uncertainty filter should be applied. In the present work, the threshold of 2 $\mu g/m^3$ is considered, as a measure of the LOAC instrumental uncertainty. The filtered subset comprises 49 data points between February and July.

Pearson's correlation coefficient is 0.76, indicating a greater linear agreement between the variables. The linear regression function has $RMSE = 3.11\ \mu g/m^3$ and $R^2 = 0.58$, which indicates a slightly better linear fit compared to results from previous studies such as [22,58]. The linear function, presented in Figure 8a, is $Y = 6.27 + 0.81 \cdot AX$.

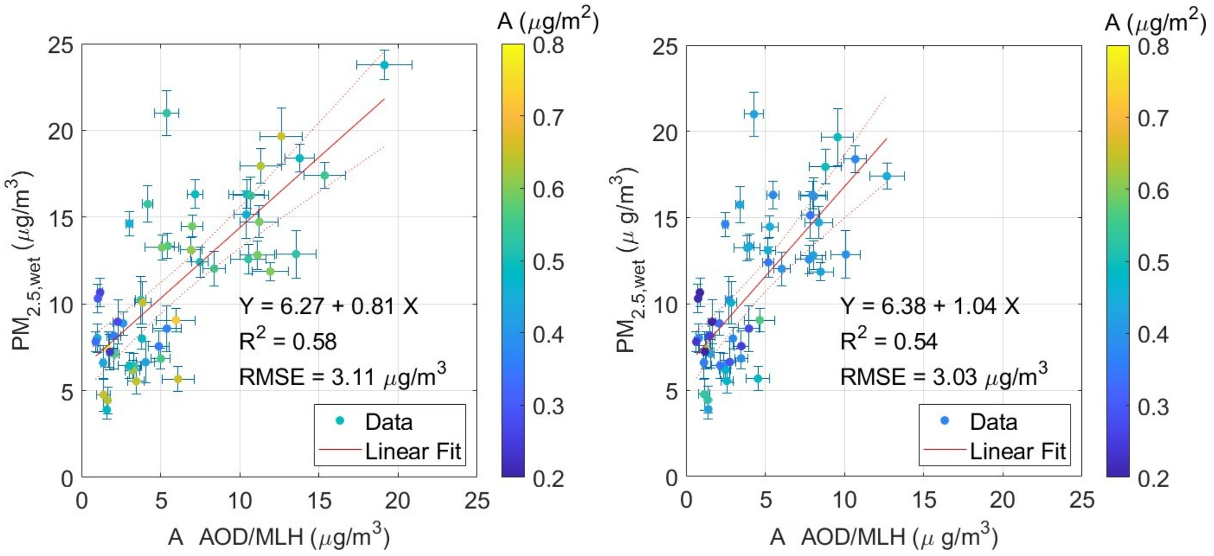


Figure 8. Results from the linear regressions with auxiliary data. The dotted lines represent the 95% confidence bounds. (**a**) Linear regression between $Y = PM_{2.5,wet}$ and $A\frac{AOD}{MLH}$. (**b**) Linear regression between $Y = PM_{2.5,wet}$ and $A_1\frac{AOD}{MLH}$, where A_1 is computed with the new values of r_{eff}. In both cases, the data points are color-coded with the respective values of A and A_1.

The linear regression is repeated by substituting the original LOAC r_{eff} with the corrected r'_{eff}. In this case, the slope of the linear function reaches a value of $a = 1.04$ with $R^2 = 0.54$, $RMSE = 3.03\ \mu g/m^3$ and $R = 0.73$ (shown in Figure 8b).

In both cases, the distribution of the residuals has a normal shape that is slightly negatively skewed, with an average close to 0 and a median of -0.2. The mean value of the residuals is also defined as Mean Bias. A Mean Bias close to zero indicates the absence of a systematic under- or overestimation in the model. The slight negative skewness simply indicates that the distribution of residuals has a slightly longer left tail, but their effects cancel out when considering the average tendencies of the model. For completeness, for both cases, the additional statistical parameters are the following:

- The Mean Absolute Error, defined as the average of the absolute value of residuals, is respectively $MAE = 2.38\ \mu g/m^3$ and $MAE = 2.34\ \mu g/m^3$;
- The relative $RMSE$, defined as the $RMSE$ divided by the average Y, is respectively $RMSE_{rel} = 0.29$ and $RMSE_{rel} = 0.28$.

To verify that the improvement from the previous correlation (R = 0.56) to the current one (R = 0.76) is mainly physical and not an artifact of the sample reduction (from n_{Obs} = 56 to n_{Obs} = 49), the first dataset ($X = \frac{AOD}{MLH}$) is reduced to the same 49 data points used in the current Section. The correlation coefficient R between $PM_{2.5,wet}$ and $X = \frac{AOD}{MLH}$ is 0.58. Thus, there is an increase from 0.56 to 0.58 only due to the reduced sample size, while there is a larger improvement in linear correlation, from 0.58 to 0.76, due to the physical information carried by the aerosol properties contained in the scaling coefficient A.

Table 3 summarizes the linear correlation results obtained in each case considered so far. All reported Pearson's R values are statistically significant at the $p < 0.001$ level (two-tailed test), given the sample sizes (n_{Obs} = 49–56). For n_{Obs} = 0.49 (n_{Obs} = 0.56), the minimum R required for $p < 0.001$ is 0.294 (0.279), well below the resulting R = 0.76 and 0.73 (R = 0.56). Thus, the observed correlations are extremely unlikely to arise by chance. Moreover, for n_{Obs} = 102 the minimum R required for significance is 0.197, above the result of R = 0.16 for the case of daily PM_{dry}, consistent with the conclusion that no meaningful linear relationship between PM_{dry} and $\frac{AOD}{MLH}$ exists prior to selection of daily PM_{dry} data according to PM_{wet} data. In the case of the 38-day subsets, the minimum R for significance is 0.334, so the result of R = 0.49 is statistically significant.

The values of the offset b in all linear functions are compatible with each other within confidence bounds. This means that the offset depends on other factors, such as the difference between the point measurements (LOAC and ceilometer data) and the satellite AOD retrieval (with a footprint of 1 km^2). Theoretically, the slope of the linear function between $PM_{2.5,wet}$ and $X = A\frac{AOD}{MLH}$ should be close to one, as A represents the theoretical slope that relates $PM_{2.5}$ to AOD and should account for all the variability. Cross-validation was not performed due to the limited sample size (56 and 49 data points). This is a limitation of the dataset used in the present study. Future, multi-site studies with larger samples should include cross-validation.

Table 3. The results from all the linear regressions with the function $Y = aX + b$, where $Y = PM_{2.5,wet}$ and X is indicated in the first column. The regressions were performed on 49 data points apart from the first, which was performed on 56 data points. The coefficients a and b are listed with relative uncertainty based on 95% confidence bounds.

X	R	R^2	RMSE (μg/m^3)	a + % Error	b + % Error
$\frac{AOD}{MLH}$	0.56	0.31	4.28	0.28 + 40%	7.26 + 26%
$\frac{4\rho_{aer}r_{eff}}{3\langle Q_{ext,aer}\rangle}\frac{AOD}{MLH}$	0.76	0.58	3.11	0.81 ± 24%	6.27 ± 23%
$\frac{4\rho_{aer}r'_{eff}}{3\langle Q_{ext,aer}\rangle}\frac{AOD}{MLH}$	0.73	0.54	3.03	1.04 ± 27%	6.38 ± 23%

3.3.3. Variable Importance Assessment

The variable importance of r_{eff}, $\langle Q_{ext,aer}\rangle$ and ρ_{aer} is assessed by adding one variable at a time to $X = \frac{AOD}{MLH}$ and comparing Pearson's coefficient values. Each time that one variable is added, only the data with relative uncertainty lower than 20% is considered.

The linear correlation coefficient between $PM_{2.5,wet}$ and $X = \frac{AOD}{MLH}$ is R = 0.5587 as previously stated. Multiplying X by r_{eff}, a value of R = 0.5992 is obtained between $PM_{2.5,wet}$ and $X = \frac{AOD}{MLH}r_{eff}$. Multiply by $\langle Q_{ext,aer}\rangle$, and a value of R = 0.6082 is obtained. Lastly, multiplying by $\frac{4}{3}\rho_{aer}$, a value of R = 0.6103 is obtained between $PM_{2.5,wet}$ and $X = A\frac{AOD}{MLH}$.

The last value is lower than the one obtained in the previous analysis because in this case the considered data points have uncertainties greater than 2 μg/m^3 on X.

4. Discussion

Pearson's correlation coefficient R = 0.56 shows that the "wet" hourly $PM_{2.5}$ is linearly correlated to AOD even without using information about aerosol properties, while MLH is always considered. For comparison, the correlation between AOD and "dry" $PM_{2.5}$ from reference ground-based measurements yields a coefficient R = 0.16, indicating the absence of a meaningful linear correlation. The lower value of R for "dry" $PM_{2.5}$ is partly attributable to its coarser temporal resolution compared to LOAC-derived "wet" $PM_{2.5}$, and underlines the need for reference data with better temporal resolution.

The analysed data show that higher values of both $PM_{2.5,wet}$ and AOD are generally associated with higher values of RH, contrary to "dry" $PM_{2.5}$, which is measured at fixed RH values lower than 50%. The use of PM observations taken at ambient humidity allows for a direct correlation with satellite AOD without hygroscopic corrections, representing an element of novelty with respect to previous works on the same topic while reaching comparable results [7,19,22,23]. While a direct comparison of correlation coefficients with "dry" $PM_{2.5}$-based studies is complicated by differences in measurement protocols and RH treatment, the relative improvement observed here (R increasing from 0.56 to 0.76 when aerosol microphysical properties are incorporated) is consistent with the broader finding that accounting for aerosol physical properties enhances the AOD-PM relationship. The low value of $R = 0.16$ obtained for "dry" $PM_{2.5}$ in this study should not be interpreted as a methodological flaw, but rather reflects the combination of factors that distinguish the present setup from previous works reporting $R \geq 0.6$: those studies used regional or multi-city datasets with large sample sizes and hourly PM measurements, while the presented "dry" $PM_{2.5}$ comparison relies on daily averages matched to a single urban site with a limited number of observations ($n_{Obs} = 102$) and a higher-resolution AOD product (MAIAC 1 km vs. MODIS 10 km). Under these more stringent conditions, the temporal mismatch between daily "dry" $PM_{2.5}$ and instantaneous $\frac{AOD}{MLH}$ is the primary driver of the low correlation.

This interpretation is supported by the additional analysis presented in Section 3.3.1: when both datasets are reduced to the same $n = 38$ daily pairs, "wet" $PM_{2.5}$ still outperforms "dry" $PM_{2.5}$ in terms of RMSE (4.33 vs. 7.71 μg/m^3), and the "dry" $PM_{2.5}$ results on those days are better than the average over random subsets of the same size ($\bar{R} = 0.39$, $\sigma = 0.12$), confirming that the day selection reflects physical constraints rather than a statistical artifact.

Furthermore, by comparing satellite-derived "wet" $PM_{2.5}$ with co-located reference "dry" $PM_{2.5}$ measurements, the local hygroscopic growth factor f(RH) can be inferred without assuming a fixed empirical parameterisation, potentially improving PM estimates in humid environments relative to approaches based on globally averaged RH corrections. A hygroscopic correction remains necessary to convert the "wet" $PM_{2.5}$ estimated from AOD into regulatory "dry" $PM_{2.5}$; this step is identified as future work.

The correlation between $PM_{2.5,wet}$ and $\frac{AOD}{MLH}$ shows no significant dependence on the time of day within the sampling window (9 AM–2 PM). Analysis of Terra (morning) vs. Aqua (afternoon) overpasses revealed comparable scatter in both cases, despite the expected MLH evolution during this period. This suggests that normalisation by MLH adequately accounts for diurnal boundary layer variations within the satellite overpass timeframe.

A Pearson's correlation coefficient of R = 0.76 is obtained when AOD is multiplied by the theoretical coefficient A, indicating a substantially greater linear agreement between the variables. The coefficient of determination $R^2 = 0.58$ is comparable to previous studies that used a linear fit for the correlation [22,58]. The linear relationship holds across the months represented in the dataset without requiring season-specific tuning, although this result cannot be extrapolated to the full annual cycle given that the dataset is weighted toward late winter and spring (February–May, approximately 75.5% of valid pairs); summer months (June–July) contribute only 24.5%, and August and September yielded no valid data points after filtering. The slope of the linear function is $a = 0.81$, below the theoretical value of unity, reflecting the high relative uncertainty of A, which in turn derives from the high relative uncertainty of r_{eff}. By repeating the regression with the corrected r'_{eff}, obtained by extrapolating particle number concentrations below the LOAC detection limit, the slope reaches $a = 1.04$ while maintaining a comparable goodness of fit ($R^2 = 0.54$, $R = 0.73$).

The assessment of variable importance shows that the uncertainty assigned to r_{eff} and X greatly impacts the linear correlation. By considering X data with relative uncertainties lower than 20%, Pearson's coefficient increases by 5% between $X = \frac{AOD}{MLH}$ and $X = A\frac{AOD}{MLH}$: in particular, adding r_{eff} increases R by 4%, while adding $\langle Q_{ext,aer} \rangle$ and ρ_{aer} contributes a further 1%. For comparison, applying an absolute uncertainty threshold of $2\,\mu g\,m^{-3}$ on both Y and X increases R by a further 15%, for a total improvement of 20% from 0.56 to 0.76.

The dominant contribution of r_{eff} is physically consistent with its large dynamic range at this site: as shown in Figure 5, r_{eff} varies from below 0.6 μm during typical conditions to above 1.5 μm during desert dust intrusions, a range substantially larger than the seasonal variability of Extinction Efficiency and Mass Density. This confirms that particle size is the primary source of variability in the AOD-PM relationship at this urban site, and that even an observational estimate of r_{eff} from an OPC instrument substantially improves the physical consistency of satellite-based $PM_{2.5}$ estimates.

5. Conclusions

$PM_{2.5}$ concentrations measured at ambient humidity by an LOAC optical particle counter were analysed alongside co-located satellite AOD retrieved at high spatial resolution. The use of ambient-humidity $PM_{2.5}$ in place of the standard dry measurement allowed a direct comparison with satellite AOD without relying on empirical hygroscopic correction functions, yielding a moderate linear correlation (Pearson's $R = 0.56$) when AOD was normalised by the Mixing Layer Height. No significant correlation was found when reference dry $PM_{2.5}$ measurements were used instead, highlighting the importance of temporal resolution and humidity consistency between the two quantities being compared.

The exploitation of aerosol microphysical and optical properties derived from LOAC measurements (Effective Radius, Extinction Efficiency, and Mass Density) enabled the derivation of a theoretical scaling coefficient linking $PM_{2.5,wet}$ and AOD at ambient conditions. These properties directly control how aerosol mass concentration translates into column-integrated extinction, providing a physical basis for the observed improvement in correlation (Pearson's $R = 0.76$) upon inclusion of the physics-informed coefficient A. The regression slope approaches the theoretical value of unity ($a = 0.81$ with the original r_{eff}, $a = 1.04$ with the corrected r'_{eff}), indicating that the theoretical relationship (Equation (5) accurately predicts both the magnitude and the variability of $PM_{2.5,wet}$ from satellite AOD. The intercept remains consistent across all regression cases within confidence bounds, suggesting that the residual offset reflects the scale mismatch between point measurements and the satellite footprint rather than a systematic physical bias. The linear relationship holds across the analysed period, but the dataset coverage is insufficient to draw conclusions about autumn and winter conditions; multi-year observations are required for a robust seasonal characterization. Multi-year, continuous observations would be required to robustly characterize the seasonal and inter-annual variability in the AOD-PM relationship within this framework; the present results should therefore be regarded as a proof-of-concept rather than a definitive seasonal characterization.

The lower detection limit of the LOAC (0.2 µm diameter) causes the overestimation of the Effective Radius, as the contribution of smaller particles is neglected. In the present work, this is addressed by extrapolating particle number concentrations to sizes below the detection limit using the first four size bins. The corrected Effective Radius, which is on average between 10% and 50% lower than the uncorrected value, improves the agreement between measured and satellite-estimated $PM_{2.5}$, bringing the regression slope closer to the theoretical value of one. This result highlights the central role of the Effective Radius in physically linking aerosol optical and mass properties, and points to the importance of accounting for the instrument detection limit when deriving aerosol microphysical properties from OPC data.

Although the analysed period does not cover a full annual cycle and the analysis is limited to a single site, this work was designed as a pilot study to evaluate the physical consistency of the proposed framework under well-controlled conditions. The use of a single, well-instrumented site allowed for the detailed and self-consistent characterization of aerosols' microphysical properties, which is essential for testing the underlying assumptions of the methodology. Wind-driven transport, solar radiation, and atmospheric pressure were not explicitly accounted for; at a single urban site with co-located measurements, these factors are considered secondary relative to aerosol microphysical properties and MLH, but may become relevant in future multi-site applications. Future work will extend the analysis to multiple locations and longer time periods to assess the transferability and robustness of the approach across different aerosol regimes and environmental conditions.

For regulatory applications, ambient $PM_{2.5}$ estimated from satellite observations can be converted to dry $PM_{2.5}$ using co-located dry PM measurements and the relation $PM_{dry} = PM_{wet} / f(RH)$ (Equation (6)). The proposed framework therefore provides a basis for indirectly deriving local aerosol hygroscopic growth factors by coupling satellite-based ambient $PM_{2.5}$ estimates with conventional dry PM observations, opening new opportunities for characterising aerosol–humidity interactions from space and for improving air quality assessments in regions with limited ground-based monitoring.

Author Contributions: Conceptualization, G.P.P. and T.M.; formal analysis, G.P.P.; data curation, G.P.P. and A.F. and S.D.S.; writing—original draft preparation, G.P.P.; writing—review and editing, T.M. and E.B.; supervision, T.M. and E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the EU—Next Generation EU, Mission 4, Component 2, Investment 3.3 (M.D. 117/2023) CUP J33C22002610002.

Data Availability Statement: The MAIAC dataset, MCD19A2 MODIS/Terra+Aqua Aerosol Optical Thickness Daily L2G Global 1km SIN Grid, was acquired from the Level-1 & Atmosphere Archive and Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the GoddardSpace Flight Center in Greenbelt, MD, USA, (https://ladsweb.modaps.eosdis.nasa.gov/, accessed on 30 June 2025) and is accessible at https://doi.org/10.5067/MODIS/MCD19A2.006 [25]. The AERONET data is available at https://aeronet.gsfc.nasa.gov/, accessed on 30 June 2025 [55]. $PM_{2.5}$ and meteorological data from ARPAE are accessible at https://dati.arpae.it/, accessed on 30 June 2025 [54]. Ceilometer data and LOAC data available on request from the authors.

Acknowledgments: We graciously thank our colleagues Francesco Barbano and Luigi Brogno for the availability and curation of the Ceilometer data. The study was partly developed within the RETURN "Multi-risk science for resilient communities under a changing climate" (National Recovery and Resilience Plan—NRPP, Mission 4, Component 2, Investment 1.3—D.D. n.341 15/3/2022, PE0000005).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Table A1. Comparative table of previous studies, indicating the study region and year of the dataset used in the study; AOD and PM datasets; correlation models; and results.

Study	Region (Year)	Data	Model	Results
Wang and Christopher (2003) [6]	U.S.A. (2002)	MODIS AOD, Hourly $PM_{2.5}$	Linear model	R = 0.7 (daily averages), R = 0.9 (monthly averages)
Koelemeijer et al. (2006) [7]	Europe (2003)	MODIS AOD, Daily and hourly $PM_{2.5}$	Linear model with MLH correction	R = 0.6 (hourly averages)
Di Nicolantonio et al. (2009) [22]	Po Valley (2004, 2007, 2008)	MODIS AOD, Daily $PM_{2.5}$	Linear model with MLH and RH corrections	R^2 = 0.68 (MODIS Terra), R^2 = 0.59 (MODIS Aqua), R^2 = 0.72 (monthly averages)
Ferrero et al. (2019) [19]	Po Valley (2006)	MAIAC AOD, Daily PM_1, $PM_{2.5}$, PM_{10}	(a) Linear model with MLH and RH corrections, (b) Multi-linear model with MLH correction (Dataset divided by season)	(a) R^2 = 0.51, (b) R^2 = 0.9 (spring), R^2 = 0.7 (other seasons)
Arvani et al. (2016) [23]	Po Valley (2010–2012)	MAIAC AOD, Daily PM_{10}	Linear model with MLH and RH correction (Dataset divided by season)	R^2 = 0.54, R^2 = 0.95 (AOD divided in bins)
Wang et al. (2010) [58]	China (2008)	MODIS AOD, Daily $PM_{2.5}$ and PM_{10}	Linear model with MLH and RH corrections	R^2 = 0.47 (daily averages), R^2 = 0.66 (monthly and regional averages)

References

1. Sharma, S.; Chandra, M.; Kota, S.H. Health effects associated with $PM_{2.5}$: A systematic review. *Curr. Pollut. Rep.* **2020**, *6*, 345–367. [CrossRef]
2. Thangavel, P.; Park, D.; Lee, Y.C. Recent insights into particulate matter ($PM_{2.5}$)-mediated toxicity in humans: An overview. *Int. J. Environ. Res. Public Health* **2022**, *19*, 7511. [CrossRef] [PubMed]
3. World Health Organization. *World Health Statistics 2013*; World Health Organization: Geneva, Switzerland, 2013.
4. Martin, R.V.; Brauer, M.; van Donkelaar, A.; Shaddick, G.; Narain, U.; Dey, S. No one knows which city has the highest concentration of fine particulate matter. *Atmos. Environ. X* **2019**, *3*, 100040. [CrossRef]
5. Hoff, R.M.; Christopher, S.A. Remote Sensing of Particulate Pollution from Space: Have We Reached the Promised Land? *J. Air Waste Manag. Assoc.* **2009**, *59*, 645–675. [CrossRef]
6. Wang, J.; Christopher, S.A. Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: Implications for air quality studies. *Geophys. Res. Lett.* **2003**, *30*, ASC 4 . [CrossRef]
7. Koelemeijer, R.; Homan, C.; Matthijsen, J. Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. *Atmos. Environ.* **2006**, *40*, 5304–5315. [CrossRef]

8. Chu, Y.; Liu, Y.; Li, X.; Liu, Z.; Lu, H.; Lu, Y.; Mao, Z.; Chen, X.; Li, N.; Ren, M.; et al. A review on predicting ground $PM_{2.5}$ concentration using satellite aerosol optical depth. *Atmosphere* **2016**, *7*, 129. [CrossRef]
9. Chen, J.; De Hoogh, K.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Bauwelinck, M.; Van Donkelaar, A.; Hvidtfeldt, U.A.; Katsouyanni, K.; et al. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Environ. Int.* **2019**, *130*, 104934. [CrossRef] [PubMed]
10. 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]
11. Chen, Z.Y.; Turrubiates, R.F.M.; Petetin, H.; Lacima, A.; Pérez García-Pando, C.; Ballester, J. Estimation of pan-European, daily total, fine-mode and coarse-mode Aerosol Optical Depth at 0.1° resolution to facilitate air quality assessments. *Sci. Total Environ.* **2024**, *918*, 170593. [CrossRef] [PubMed]
12. Gu, J.; Wang, Y.; Ma, J.; Lu, Y.; Wang, S.; Li, X. An Estimation Method for $PM_{2.5}$ Based on Aerosol Optical Depth Obtained from Remote Sensing Image Processing and Meteorological Factors. *Remote Sens.* **2022**, *14*, 1617. [CrossRef]
13. Lin, H.; Li, S.; Niu, J.; Yang, J.; Wang, Q.; Li, W.; Liu, S. Estimation of Ultrahigh Resolution $PM_{2.5}$ in Urban Areas by Using 30 m Landsat-8 and Sentinel-2 AOD Retrievals. *Remote Sens.* **2025**, *17*, 2609. [CrossRef]
14. 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]
15. Fan, Z.; Zhan, Q.; Yang, C.; Liu, H.; Bilal, M. Estimating $PM_{2.5}$ Concentrations Using Spatially Local Xgboost Based on Full-Covered SARA AOD at the Urban Scale. *Remote Sens.* **2020**, *12*, 3368. [CrossRef]
16. 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]
17. Handschuh, J.; Erbertseder, T.; Schaap, M.; Baier, F. Estimating $PM_{2.5}$ surface concentrations from AOD: A combination of SLSTR and MODIS. *Remote Sens. Appl. Soc. Environ.* **2022**, *26*, 100716. [CrossRef]
18. Li, Y.; Yuan, S.; Fan, S.; Song, Y.; Wang, Z.; Yu, Z.; Yu, Q.; Liu, Y. Satellite remote sensing for estimating $PM_{2.5}$ and its components. *Curr. Pollut. Rep.* **2021**, *7*, 72–87. [CrossRef]
19. Ferrero, L.; Riccio, A.; Ferrini, B.; D'Angelo, L.; Rovelli, G.; Casati, M.; Angelini, F.; Barnaba, F.; Gobbi, G.; Cataldi, M.; et al. Satellite AOD conversion into ground PM_{10}, $PM_{2.5}$ and PM_1 over the Po valley (Milan, Italy) exploiting information on aerosol vertical profiles, chemistry, hygroscopicity and meteorology. *Atmos. Pollut. Res.* **2019**, *10*, 1895–1912. [CrossRef]
20. Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on Ambient Air Quality and Cleaner Air for Europe (Recast) . Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202402881 (accessed on 7 August 2025).
21. Di Antonio, L.; Di Biagio, C.; Foret, G.; Formenti, P.; Siour, G.; Doussin, J.F.; Beekmann, M. Aerosol optical depth climatology from the high-resolution MAIAC product over Europe: Differences between major European cities and their surrounding environments. *Atmos. Chem. Phys.* **2023**, *23*, 12455–12475. [CrossRef]
22. Di Nicolantonio, W.; Cacciari, A.; Tomasi, C. Particulate Matter at Surface: Northern Italy Monitoring Based on Satellite Remote Sensing, Meteorological Fields, and in-situ Samplings. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *2*, 284–292. [CrossRef]
23. Arvani, B.; Pierce, R.B.; Lyapustin, A.I.; Wang, Y.; Ghermandi, G.; Teggi, S. Seasonal monitoring and estimation of regional aerosol distribution over Po valley, northern Italy, using a high-resolution MAIAC product. *Atmos. Environ.* **2016**, *141*, 106–121. [CrossRef]
24. Stafoggia, M.; Schwartz, J.; Badaloni, C.; Bellander, T.; Alessandrini, E.; Cattani, G.; de' Donato, F.; Gaeta, A.; Leone, G.; Lyapustin, A.; et al. Estimation of daily PM_{10} concentrations in Italy (2006–2012) using finely resolved satellite data, land use variables and meteorology. *Environ. Int.* **2017**, *99*, 234–244. [CrossRef] [PubMed]
25. Lyapustin, A.; Wang, Y.; Korkin, S.; Huang, D. MODIS Collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* **2018**, *11*, 5741–5765. [CrossRef]
26. Zieger, P.; Fierz-Schmidhauser, R.; Weingartner, E.; Baltensperger, U. Effects of relative humidity on aerosol light scattering: Results from different European sites. *Atmos. Chem. Phys.* **2013**, *13*, 10609–10631. [CrossRef]
27. Hänel, G. An attempt to interpret the humidity dependencies of the aerosol extinction and scattering coefficients. *Atmos. Environ. (1967)* **1981**, *15*, 403–406. [CrossRef]
28. Karagulian, F.; Barbiere, M.; Kotsev, A.; Spinelle, L.; Gerboles, M.; Lagler, F.; Redon, N.; Crunaire, S.; Borowiak, A. Review of the Performance of Low-Cost Sensors for Air Quality Monitoring. *Atmosphere* **2019**, *10*, 506. [CrossRef]
29. deSouza, P.; Kahn, R.A.; Limbacher, J.A.; Marais, E.A.; Duarte, F.; Ratti, C. Combining low-cost, surface-based aerosol monitors with size-resolved satellite data for air quality applications. *Atmos. Meas. Tech.* **2020**, *13*, 5319–5334. [CrossRef]
30. Kim, K.; Lee, D.; Lee, K.; Lee, K.; Noh, Y. Estimation of surface-level $PM_{2.5}$ concentration based on MODIS aerosol optical depth over Jeju, Korea. *Korean J. Remote Sens.* **2016**, *32*, 413–421. [CrossRef]
31. Fawole, O.G.; Ogunjo, S.T.; Olabode, A.; Alabi, W.; Sa'id, R.S. Estimation of surface $PM_{2.5}$ concentration from satellite aerosol optical depth using a constrained observation-based model. *Climate* **2026**, *14*, 1. [CrossRef]

32. Renard, J.B.; Dulac, F.; Berthet, G.; Lurton, T.; Vignelles, D.; Jégou, F.; Tonnelier, T.; Jeannot, M.; Couté, B.; Akiki, R.; et al. LOAC: A small aerosol optical counter/sizer for ground-based and balloon measurements of the size distribution and nature of atmospheric particles—Part 1: Principle of measurements and instrument evaluation. *Atmos. Meas. Tech.* **2016**, *9*, 1721–1742. [CrossRef]
33. Carbone, C.; Decesari, S.; Mircea, M.; Giulianelli, L.; Finessi, E.; Rinaldi, M.; Fuzzi, S.; Marinoni, A.; Duchi, R.; Perrino, C.; et al. Size-resolved aerosol chemical composition over the Italian Peninsula during typical summer and winter conditions. *Atmos. Environ.* **2010**, *44*, 5269–5278. [CrossRef]
34. Ferrero, L.; Riccio, A.; Perrone, M.; Sangiorgi, G.; Ferrini, B.; Bolzacchini, E. Mixing height determination by tethered balloon-based particle soundings and modelling simulations. *Atmos. Res.* **2011**, *102*, 145–156. [CrossRef]
35. Barnaba, F.; Putaud, J.P.; Gruening, C.; dell'Acqua, A.; Dos Santos, S. Annual cycle in co-located in situ, total-column, and height-resolved aerosol observations in the Po Valley (Italy): Implications for ground-level particulate matter mass concentration estimation from remote sensing. *J. Geophys. Res. Atmos.* **2010**, *115*, D19209 . [CrossRef]
36. Ricciardelli, I.; Bacco, D.; Rinaldi, M.; Bonafè, G.; Scotto, F.; Trentini, A.; Bertacci, G.; Ugolini, P.; Zigola, C.; Rovere, F.; et al. A three-year investigation of daily $PM_{2.5}$ main chemical components in four sites: The routine measurement program of the Supersito Project (Po Valley, Italy). *Atmos. Environ.* **2017**, *152*, 418–430. [CrossRef]
37. Scotto, F.; Bacco, D.; Lasagni, S.; Trentini, A.; Poluzzi, V.; Vecchi, R. A multi-year source apportionment of $PM_{2.5}$ at multiple sites in the southern Po Valley (Italy). *Atmos. Pollut. Res.* **2021**, *12*, 101192. [CrossRef]
38. Tositti, L.; Brattich, E.; Masiol, M.; Baldacci, D.; Ceccato, D.; Parmeggiani, S.; Stracquadanio, M.; Zappoli, S. Source apportionment of particulate matter in a large city of southeastern Po Valley (Bologna, Italy). *Environ. Sci. Pollut. Res.* **2013**, *21*, 872–890. [CrossRef] [PubMed]
39. Gupta, P.; Christopher, S.A.; Wang, J.; Gehrig, R.; Lee, Y.; Kumar, N. Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmos. Environ.* **2006**, *40*, 5880–5892. [CrossRef]
40. Crilley, L.R.; Shaw, M.; Pound, R.; Kramer, L.J.; Price, R.; Young, S.; Lewis, A.C.; Pope, F.D. Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. *Atmos. Meas. Tech.* **2018**, *11*, 709–720. [CrossRef]
41. Brattich, E.; Bracci, A.; Zappi, A.; Morozzi, P.; Di Sabatino, S.; Porcù, F.; Di Nicola, F.; Tositti, L. How to Get the Best from Low-Cost Particulate Matter Sensors: Guidelines and Practical Recommendations. *Sensors* **2020**, *20*, 3073. [CrossRef] [PubMed]
42. Liu, H.; Shah, S.; Jiang, W. On-line outlier detection and data cleaning. *Comput. Chem. Eng.* **2004**, *28*, 1635–1647. [CrossRef]
43. Renard, J.B.; Thaury, C.; Mineau, J.L.; Gaubicher, B. Small-angle light scattering by airborne particulates: Environnement S.A. continuous particulate monitor. *Meas. Sci. Technol.* **2010**, *21*, 085901. [CrossRef]
44. Hess, M.; Koepke, P.; Schult, I. Optical Properties of Aerosols and Clouds: The Software Package OPAC. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 831–844. [CrossRef]
45. Qin, W.; Fang, H.; Wang, L.; Wei, J.; Zhang, M.; Su, X.; Bilal, M.; Liang, X. MODIS high-resolution MAIAC aerosol product: Global validation and analysis. *Atmos. Environ.* **2021**, *264*, 118684. [CrossRef]
46. Rogozovsky, I.; Ohneiser, K.; Lyapustin, A.; Ansmann, A.; Chudnovsky, A. The impact of different aerosol layering conditions on the high-resolution MODIS/MAIAC AOD retrieval bias: The uncertainty analysis. *Atmos. Environ.* **2023**, *309*, 119930. [CrossRef]
47. Münkel, C.; Eresmaa, N.; Räsänen, J.; Karppinen, A. Retrieval of mixing height and dust concentration with lidar ceilometer. *Bound. Layer Meteorol.* **2007**, *124*, 117–128. [CrossRef]
48. Liu, B.; Ma, X.; Ma, Y.; Li, H.; Jin, S.; Fan, R.; Gong, W. The relationship between atmospheric boundary layer and temperature inversion layer and their aerosol capture capabilities. *Atmos. Res.* **2022**, *271*, 106121. [CrossRef]
49. Eresmaa, N.; Härkönen, J.; Joffre, S.M.; Schultz, D.M.; Karppinen, A.; Kukkonen, J. A Three-Step Method for Estimating the Mixing Height Using Ceilometer Data from the Helsinki Testbed. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 2172–2187. [CrossRef]
50. Haman, C.L.; Lefer, B.; Morris, G.A. Seasonal Variability in the Diurnal Evolution of the Boundary Layer in a Near-Coastal Urban Environment. *J. Atmos. Ocean. Technol.* **2012**, *29*, 697–710. [CrossRef]
51. Li, D.; Wu, Y.; Gross, B.; Moshary, F. Dynamics of Mixing Layer Height and Homogeneity from Ceilometer-Measured Aerosol Profiles and Correlation to Ground Level $PM_{2.5}$ in New York City. *Remote Sens.* **2022**, *14*, 6370. [CrossRef]
52. Durre, I.; Vose, R.S.; Wuertz, D.B. Overview of the Integrated Global Radiosonde Archive. *J. Clim.* **2006**, *19*, 53–68. [CrossRef]
53. Durre, I.; Yin, X.; Vose, R.S.; Applequist, S.; Arnfield, J. Enhancing the data coverage in the Integrated Global Radiosonde Archive. *J. Atmos. Ocean. Technol.* **2018**, *35*, 1753–1770. [CrossRef]
54. Arpae Emilia-Romagna. Available online: https://www.arpae.it/it (accessed on 30 June 2025).
55. Aerosol Robotic Network (AERONET). Available online: https://aeronet.gsfc.nasa.gov/new_web/index.html (accessed on 30 June 2025).
56. Holben, B.; Eck, T.; Slutsker, I.; Tanré, D.; Buis, J.; Setzer, A.; Vermote, E.; Reagan, J.; Kaufman, Y.; Nakajima, T.; et al. AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. *Remote Sens. Environ.* **1998**, *66*, 1–16. [CrossRef]

57. Mazzola, M.; Lanconelli, C.; Lupi, A.; Busetto, M.; Vitale, V.; Tomasi, C. Columnar aerosol optical properties in the Po Valley, Italy, from MFRSR data. *J. Geophys. Res. Atmos.* **2010**, *115*, D17206 . [CrossRef]
58. Wang, Z.; Chen, L.; Tao, J.; Zhang, Y.; Su, L. Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method. *Remote Sens. Environ.* **2010**, *114*, 50–63. [CrossRef]

https://doi.org/10.3390/atmos17070627