

## Article

# Aerosol Optical Properties and Type Retrieval via Machine Learning and an All-Sky Imager

Stavros-Andreas Logothetis<sup>1</sup>, Christos-Panagiotis Giannaklis<sup>1</sup>, Vasileios Salamalikis<sup>2</sup>, Panagiotis Tzoumanikas<sup>1</sup>, Panagiotis-Ioannis Raptis<sup>3</sup>, Vassilis Amiridis<sup>4</sup>, Kostas Eleftheratos<sup>3,5</sup> and Andreas Kazantzidis<sup>1,\*</sup>

<sup>1</sup> Laboratory of Atmospheric Physics, Physics Department, University of Patras, GR-26500 Patras, Greece; logothetis\_s@upnet.gr (S.-A.L.); up1057893@upnet.gr (C.-P.G.); tzumanik@ceid.upatras.gr (P.T.)

<sup>2</sup> NILU—Norwegian Institute for Air Research, P.O. Box 100, 2027 Kjeller, Norway; vsal@nilu.no

<sup>3</sup> Department of Geology and Geoenvironment, National and Kapodistrian University of Athens, GR-15784 Athens, Greece; piraptis@meteo.noa.gr (P.-I.R.); kelef@geol.uoa.gr (K.E.)

<sup>4</sup> Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, GR-15236 Athens, Greece; vamoir@noa.gr

<sup>5</sup> Center for Environmental Effects on Health, Biomedical Research Foundation of the Academy of Athens, GR-11527 Athens, Greece

\* Correspondence: akaza@upatras.gr

**Abstract:** This study investigates the applicability of using the sky information from an all-sky imager (ASI) to retrieve aerosol optical properties and type. Sky information from the ASI, in terms of Red-Green-Blue (RGB) channels and sun saturation area, are imported into a supervised machine learning algorithm for estimating five different aerosol optical properties related to aerosol burden (aerosol optical depth, AOD at 440, 500 and 675 nm) and size (Ångström Exponent at 440–675 nm, and Fine Mode Fraction at 500 nm). The retrieved aerosol optical properties are compared against reference measurements from the AERONET station, showing adequate agreement (R: 0.89–0.95). The AOD errors increased for higher AOD values, whereas for AE and FMF, the biases increased for coarse particles. Regarding aerosol type classification, the retrieved properties can capture 77.5% of the total aerosol type cases, with excellent results for dust identification (>95% of the cases). The results of this work promote ASI as a valuable tool for aerosol optical properties and type retrieval.

**Keywords:** all-sky imager; aerosol properties; AERONET; machine learning; aerosol type classification



**Citation:** Logothetis, S.-A.; Giannaklis, C.-P.; Salamalikis, V.; Tzoumanikas, P.; Raptis, P.-I.; Amiridis, V.; Eleftheratos, K.; Kazantzidis, A. Aerosol Optical Properties and Type Retrieval via Machine Learning and an All-Sky Imager. *Atmosphere* **2023**, *14*, 1266. <https://doi.org/10.3390/atmos14081266>

Academic Editors: Madhu Gyawali, Rudra P. Aryal and Yadav Pandit

Received: 28 June 2023

Revised: 29 July 2023

Accepted: 8 August 2023

Published: 10 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Aerosols regulate the Earth's radiative balance by interacting with atmospheric radiation [1]. Under cloudless conditions, the aerosol radiative impact on the climate is described by direct aerosol radiative forcing (DARF), which is often characterized by significant uncertainties due to the high spatiotemporal variabilities of the aerosol's physical and optical properties [2]. Those properties depend on the emission sources; biogenic (e.g., sea salt and dust) and anthropogenic (e.g., biomass burning and fossil fuel combustion), and aerosol composition and size [3–6]. The contiguous and accurate monitoring of aerosol properties at the finest spatiotemporal resolution is desirable in order to (a) better understand the aerosol effect on climate and (b) model the solar irradiance.

Several techniques exist to monitor aerosol properties including ground-based measurements and satellite observations. The AERONET (AERosol RObotic NETwork) provides aerosol optical properties at the highest temporal resolution (5–15 min), with over 600 stations worldwide using CIMEL sun-photometers [7], and it is commonly used for validation studies related to aerosol properties retrieval. Although the AERONET stations provide high-quality aerosol information, most stations are located in landlocked and low-elevated regions. On the other hand, satellite-based remote sensing offers a

near-global spatiotemporal coverage of aerosol optical properties. Over the past few decades, satellite remote sensing using passive imaging radiometers, such as MODIS (MODerate resolution Imaging Spectroradiometer) [8–10], MISR (Multi-angle Imaging SpectroRadiometer) [11], MERIS (MEdium Resolution Imaging Spectrometer) [12], AATSR (Advanced Along Track Scanning Radiometer) [13], GOES (Geostationary Observational Environmental Satellite) [14], POLDER (POLarization and Directionality of the Earth's Reflectances) [15], and VIIRS (Visible Infrared Imaging Radiometer Suite) [16], measure aerosol properties at regional and global scales. Several limitations characterize the aerosol satellite retrievals, such as (i) low temporal resolution, (ii) sensor calibration, (iii) cloud detection, and (iv) surface reflections and brightness, inducing significant uncertainties in the aerosol retrieval [17–19]. Based on the limitations of the currently available aerosol retrievals, it is essential to explore alternative techniques for aerosol monitoring that may work in conjunction with remote sensing instruments.

ASIs (All-sky imagers) constitute automated imaging systems that depict the whole sky (~180° view) by taking hemispherical images. ASIs have been proven very useful for cloud-related applications such as cloud segmentation [20–23], cloud classification [21,24–27], cloud base height estimation [28–30], solar irradiance and solar ramp forecasting [31–38], and solar irradiance estimation [39,40].

Several studies reported the potential use of an ASI system to characterize aerosol properties. Olmo et al. [41] calculated the AOD at 550 nm using the spectral radiances obtained from ASI and radiative transfer model (RTM) simulations. The predicted AOD underestimated the AERONET observations, with 80% (60%) of the AOD deviations being lower than 0.04 (0.02). Cazorla et al. [42] used radiances derived from a calibrated ASI and artificial neural network (ANN) to estimate AOD at three wavelengths (440, 675, and 870 nm). Compared to the AERONET, the predicted AODs revealed adequate performance with a coefficient of determination ( $R^2$ ) higher than 0.90 at all wavelengths. In addition to AOD, they also calculated the Ångström Exponent (AE), revealing a weaker correlation against the AERONET ( $R^2 = 0.77$ ). Huo and Lu [43] retrieved AOD at 500 nm by comparing the Blue (B) to Red (R) ratio derived from an ASI against the relevant radiance ratios (between 440 and 650 nm) from RTM simulations (MODTRAN) [44,45] at the Xianghe Observatory in Hebei Province, China. The retrieved AOD, against sun-photometer measurements, showed a correlation coefficient close to 0.95 and an average retrieval error of around 7%. Kazantzidis et al. [46] proposed a methodology to calculate AOD at 440, 500, and 675 nm using the RGB (Red-Green-Blue channels) intensities, the sun-saturated area (SAT) from an ASI, and RTM radiances (libRadtran) [47] at the Plataforma Solar de Almeria, Spain. A validation against the AERONET observations showed mean/median differences and a standard deviation lower than 0.01 and 0.03, respectively.

An ASI has been used in the GRASP code (Generalized Retrieval of Atmosphere and Surface Properties) [48] to retrieve aerosol properties [49]. The normalized sky radiances (NSRs) [50], extracted from the ASI at three effective wavelengths (467, 536, and 605 nm), were imported to GRASP for retrieving AOD. The GRASP-AOD and AERONET-AOD retrievals revealed  $R^2$  close to 0.87. Generally, the median and standard deviation of the AOD differences were between 0.006–0.010 and 0.024–0.030, respectively. Scarlatti et al. [51] recently proposed a machine learning (ML) approach for AOD and AE retrieval using the smoothing RGB signals towards the principal plane, as captured from a well-calibrated ASI installed at the University of Valencia, Spain. Different distinct ML models with varied input information were implemented. More specifically, AOD and AE were predicted using the color signal relevant to the spectral AOD or using solely the red channel. In addition, the B/R ratio was also applied to retrieve AOD and AE. All the aforementioned approaches revealed adequate retrieval performance against the AERONET, with  $R^2$  exceeding 0.95. A novel point of the Scarlatti et al. [51] study was the implementation of partially clouded images during ML model training.

In cases of clear-sky conditions, solar resource estimation is linked mainly to AOD and total column water vapor (TCWP) [52,53]. The global and direct components of the surface

solar irradiance (GHI and DNI) are important for PV systems and concentrated solar power (CSP) technologies [53–55]. The former two solar parameters under clear skies are primarily affected by the atmospheric aerosol burden and aerosol scattering effects [56]. Thus, quality-assured AOD measurements with low errors are required to model DNI and GHI accurately. Aerosol optical properties are frequently used to characterize aerosol type. Particle type classification is important in aerosol modelling since different aerosol size modes present different chemical compositions and deposition processes [57]. Various aerosol classification schemes have been proposed in the literature, relying on aerosol absorptivity (single scattering albedo, SSA) and AE and/or FMF using AERONET retrievals [58–64].

The principal objectives of the presented study are:

1. The application of a supervised learning technique for retrieving AOD at 440, 500, and 674 nm ( $AOD_{440\text{ nm}}$ ,  $AOD_{500\text{ nm}}$ , and  $AOD_{675\text{ nm}}$ ), Ångström Exponent between 440 and 675 nm ( $AE_{440-675\text{ nm}}$ ), and Fine Mode Fraction at 500 nm ( $FMF_{500\text{ nm}}$ ) using valuable sky information from an ASI;
2. The efficiency of the results in performing aerosol-type classification.

The retrievals of the presented methodology can be used to expand and complement the aerosol properties and type information in non-existing timeframes when sky information from an ASI is available.

This paper is structured as follows: Section 2 describes the measurement site and the measuring instruments. Section 3 presents the methodology for the retrieval of the aerosol properties and the implemented validation metrics. Section 4 presents the results of the study and, lastly, the main conclusions are summarized in Section 5.

## 2. Data

### 2.1. Measurement Site

The data used in this study were collected from an ASI and a CIMEL Sun-photometer installed at the National Observatory of Athens (NOA; 37.97° N, 23.72° E; 130 m above sea level) in Thessio, Greece. The site is characterized as urban, located in the city center of Athens. According to the Köppen–Geiger classification system, the climate is of type Csa with warm and dry summers and wet and mild winters.

Both natural and anthropogenic sources contribute to the aerosol conditions of Athens. For instance, dust episodes from the Sahara Desert are a common natural occurrence, accounting for ~23% of the annually averaged AOD [65]. Other sources are sea-salt aerosols emitted by the surrounding sea bodies [66], biomass burning [67], wildfires [68,69], and photochemical pollution [70,71]. Raptis et al. [72] performed a detailed analysis using decadal measurements of aerosol properties ( $AOD_{440\text{ nm}}$  and  $AE_{440-870\text{ nm}}$ ) from the NOA AERONET station to extract the seasonality and the trends of aerosols in the region. Polluted (27%) and mixed (23%) aerosols are dominant, followed by continental (19%) and dust (16%) particles. Marine (11%) and biomass burning (5%) aerosols have the fewest occurrence frequencies.

### 2.2. Measuring Instruments

#### 2.2.1. AERONET Station

The CE318 sun-sky photometer (CIMEL Electronique, Paris, France) is the standard instrument of all AERONET stations. It measures the sun-collimated direct beam irradiance and sky radiance to provide high-quality aerosol optical and microphysical properties [7,73]. More specifically, sun measurements with a 1.2° field of view were obtained at ~5 to 15 min and AOD [74] was calculated at eight (8) standard wavelengths between 340 and 1640 nm. Precipitable water was also retrieved via the 940 nm channel [74,75]. Sky radiances in the almucantar geometry (zenith angle is set equal to solar zenith angle, and ±180 azimuthal sweeps), at 440, 675, 870, and 1020 nm, also gave aerosol properties such as SSA, aerosol volume distribution, phase function, FMF, refractive index, etc., via an inverse algorithm [7,76]. Typically, the total AOD uncertainty under clear skies is less than

$\pm 0.01$  for  $\lambda > 440$  nm and less than  $\pm 0.02$  for shorter wavelengths [74], reflecting the variations in atmospheric conditions and instrumental calibrations.

The AERONET database is divided into three data quality levels: Level 1.0, Level 1.5, and Level 2.0. In the latest Version 3, Level 1.0 is for prescreened data; Level 1.5 represents data with near-real-time automatic cloud screening and automatic instrument anomaly quality controls; and Level 2.0 includes pre-field and post-field calibrations [74]. The current work uses AERONET Level 2.0 Version 3 (L2V3) measurements and, more specifically, the following parameters:  $AOD_{440\text{ nm}}$ ,  $AOD_{500\text{ nm}}$ ,  $AOD_{675\text{ nm}}$ ,  $AE_{440-675\text{ nm}}$ , and  $FMF_{500\text{ nm}}$ . The data were downloaded directly from the network's website (<https://aeronet.gsfc.nasa.gov/>; accessed on 28 June 2023).

### 2.2.2. All-Sky Imager

In this work, a commercial Mobotix Q24M ASI ([www.mobotix.com](http://www.mobotix.com); accessed on 28 June 2023) was installed on NOAA station. This hemispheric camera is a weatherproof IP network dome camera with a special hemispheric lens (fisheye). It provides sky images every 640  $\mu$ s, stored in 24-bit JPEG format with a spatial resolution of  $1024 \times 768$  pixels. The sensor has a RGB filter. These channels correspond to 440, 500, and 675 nm wavelengths with color intensities ranging from 0 to 255. The whole sky is represented as a circle and has no shadowing mechanism in the system. Inevitably, at the images' edges, close to the horizon, obstacles, such as other instruments installed on the rooftop of the measurement site, nearby buildings, and trees, are detected and excluded from further analysis.

The aerosol measurements from the CIMEL sun-photometer were temporally synchronized with the nearest ASI image. The study period ranges between 1 January 2021 and 18 November 2021 and the measurements were collected during the experimental campaign of the ASPIRE project (Atmospheric parameters affecting SPectral solar IRradiance and solar Energy) (<https://aspire.geol.uoa.gr/>; accessed on 28 June 2023). Possible cloud-contaminated aerosol data that were not flagged as cloudy instances during the cloud-screening approach were further detected and removed from the dataset based on the ASI images. Moreover, instances with a solar zenith angle (SZA) greater than  $70^\circ$  were discarded to avoid cases with the sun at the edges of the image, where obstacles may interfere. Finally, 3212 images were retained for subsequent analysis.

## 3. Methodology

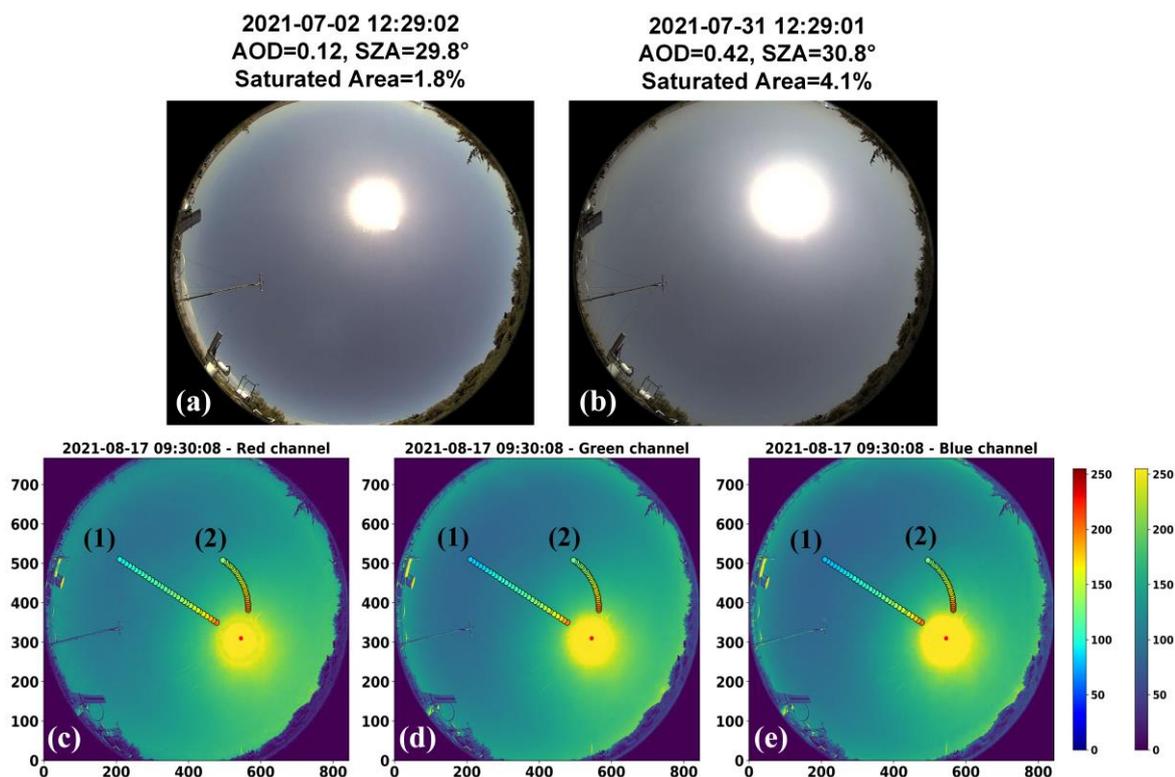
### 3.1. Machine Learning Approach

This work uses an updated Gradient Boosting Machine (GBM) [77] algorithm, the Light GBM (LGBM) [78], for regression. GBM is an ensemble technique that performs iterative decision trees in an additive and sequential way. The processing of the decision trees is the principal difference between the LGBM and GBM, where LGBM implements a leaf-wise tree growth, and GBM applies a depth-wise tree growth. The former reflects lower memory usage, better accuracy, and easy handling of large datasets. Although LGBM constitutes a suitable algorithm, several other supervised machine learning (ML) techniques are also investigated (see Section 4.1). Finally, LGBM is selected regarding the overall performance, accuracy, and total training time.

LGBM is trained and tested using a 70/30 random splitting procedure. Several parameters called hyperparameters describe the LGBM architecture (e.g., learning rate, max depth). The 'optimal' hyperparameter combination is selected through a randomized searching approach and a 10-fold cross-validation process. The mean square error (MSE) is a commonly used fitness function measuring the amount of error in regression analyses, and the lower MSE extracts the 'optimal' hyperparameter combination.

Different LGBM models were designed to predict  $AOD_{440\text{ nm}}$ ,  $AOD_{500\text{ nm}}$ ,  $AOD_{675\text{ nm}}$ ,  $AE_{440-675\text{ nm}}$ , and  $FMF_{500\text{ nm}}$  using the sky information from ASI images as well as the SZA and total column water vapor (TCWV) as auxiliary information. Regarding ASI sky information, two different sources of data were extracted. Firstly, RGB values from 60 pixels around the sun are selected (Figure 1c–e), where 30 out of 60 pixels consider a con-

stant zenith angle and variable azimuth angle (almucantar: curved line (2) in Figure 1c–e) and vice versa (principal plane: straight line (1) in Figure 1c–e). Torres et al. [79] stated that almucantar retrievals, combined with principal plane inversions, using ground-based radiometer observations could result in satisfying aerosol retrievals. Almucantar observations offer the great advantage of symmetry, a decisive factor in minimizing errors and cloud-screening processes. At the same time, principal plane retrievals provide accuracy during the day, which is crucial for aerosol studies.



**Figure 1.** Two all-sky imager (ASI) images with similar solar zenith angle (SZA) and different aerosol loads on (a) 2 July 2021 at 12:29:02 UTC with SZA = 29.8, aerosol optical depth (AOD) = 0.12, and sun-saturated area (SAT) = 1.8%, and (b) 31 July 2021 at 12:29:01 UTC with SZA = 30.8, AOD = 0.42, and SAT = 4.1%. The three heatmap images illustrate the 60 selected pixels in (c) red, (d) green, and (e) blue color scale. Both color bars indicate the intensity of the image color. The different color bars are applied to easily distinguish the selected pixels on the graph. Principal plane: (1) the straight line refers to pixels with constant azimuth and varying zenith angles (30 pixels with a 2° step). Almucantar: (2) the curved line refers to pixels with constant zenith and varying azimuth angles (30 pixels with a 2° step).

Secondly, the saturation area (SAT in %) is defined as the ratio between the number of pixels around the sun that includes sunlight and the total number of image pixels. In order to explain the significance of SAT, two cases with different AOD and almost identical solar geometries are represented in Figure 1a,b. The increased solar disk in Figure 1b explains that increases in SAT (from 1.8% to 4.1%) are reflected in AOD increases (from 0.12 to 0.42). In particular, a higher SAT corresponds to a higher aerosol burden within the atmosphere, according to the Mie theory [80].

TCWV strongly correlates to AOD in various aerosol environments (biomass burning, heavy pollution, and dust) [81]. Therefore, TCWV is also included as model input. TCWV data were obtained through CAMS (Copernicus Atmosphere Monitoring Service, <https://ads.atmosphere.copernicus.eu/>; accessed on 28 June 2023) Reanalysis product [82], and were linearly interpolated from their native temporal resolution (3 h) to fit AERONET

and ASI measurements. TCWC from CAMS encompasses adequately good performance against AERONET TCWV, with  $R^2 = 0.94$  and  $RMSE = 0.22$  cm (Figure S1).

Finally, five LGBM models are trained, one for each aerosol parameter. SAT, SZA, and TCWV parameters are standard input parameters for all models. On the other hand, models for  $AOD_{440\text{ nm}}$ ,  $AOD_{500\text{ nm}}$ , and  $AOD_{675\text{ nm}}$  use each image’s blue, green, and red channels relevant to each wavelength. The aerosol size properties use all RGB values. Before the training process, all input parameters are normalized to extend between 0 and 1 through the Min-Max normalization process, except for SZA, where the cosine function is applied. Hereafter, the retrievals of the proposed methodology will be abbreviated as ‘ML-ASI’.

The proposed methodology can be applied in any instance if AERONET retrievals and ASI data are jointly available to train the machine learning algorithm. The CIMEL sun photometer can be installed for a certain period (in this study, that period was 11 months) for recording the month-to-month variability of aerosol and climate characteristics. After the model training, the ML-ASI can be used to reproduce the aerosol optical properties without the presence of the CIMEL. It should be mentioned that model transferability is possible at sites with similar aerosol and climate climatology without requiring the presence of a sun photometer.

The performance of the applied methodology could be affected by the earth’s location in different ways. Firstly, the proposed methodology is applied only to clear-sky ASI measurements. In regions with relatively high cloud presence (e.g., North Europe), it is expected to be more difficult to train the machine learning algorithm due to the suppression of data sampling. Secondly, regarding the model input parameter, TCWV, it is also expected to be related to the earth’s location. In higher latitudes ( $>60^\circ$  N and  $>60^\circ$  S), due to lower water vapor levels, the TCWV-AOD correlation may be eradicated. Nevertheless, the decrease in TCWV-AOD correlation is not expected to reduce the performance of the applied methodology but rather the importance level of TCWV as a feature parameter to machine learning algorithms.

### 3.2. Validation Metrics

The predicted aerosol parameters are compared against ground observations in terms of the Mean Bias Error (MBE), relative MBE (rMBE), Root Mean Square Error (RMSE), relative RMSE (rRMSE), and Pearson’s correlation coefficient (R):

$$MBE = \frac{1}{N} \sum_{i=1}^N (y_{ML-ASI, i} - y_{AERONET, i}) \tag{1}$$

$$rMBE = 100 \times \frac{MBE}{\bar{y}_{AERONET}} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{ML-ASI, i} - y_{AERONET, i})^2} \tag{3}$$

$$rRMSE = 100 \times \frac{RMSE}{\bar{y}_{AERONET}} \tag{4}$$

$$R = \frac{\sum_{i=1}^N (y_{ML-ASI, i} - \bar{y}_{ML-ASI}) (y_{AERONET, i} - \bar{y}_{AERONET})}{\sqrt{\sum_{i=1}^N (y_{ML-ASI, i} - \bar{y}_{ML-ASI})^2} \sqrt{\sum_{i=1}^N (y_{AERONET, i} - \bar{y}_{AERONET})^2}} \tag{5}$$

where  $y_{ML-ASI}$  and  $y_{AERONET}$  are the retrieved aerosol parameters from the ML model and AERONET, respectively.

The performance of ML-ASI-predicted aerosol retrievals for the aerosol type classification is achieved in terms of total accuracy and precision:

$$\text{Total Accuracy} = \frac{\text{True cases}}{\text{Total cases}} \quad (6)$$

$$\text{Precision} = \frac{\text{Predicted cases per aerosol type}}{\text{Observed cases per aerosol type}} \quad (7)$$

All the above statistical metrics are calculated solely using the testing dataset as described in Section 3.1. The testing dataset is also used to interpret the results of the proposed methodology (Section 4).

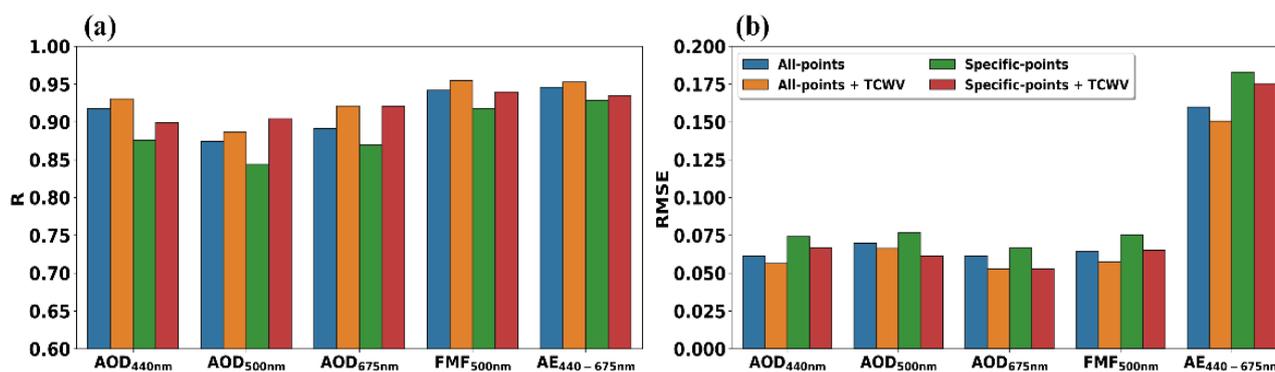
## 4. Results

### 4.1. Performance of the Retrieved Aerosol Optical Properties

The LGBM model's efficiency in reproducing aerosol parameters is presented in the following sections. A sensitivity analysis of the model input parameters is presented in Section 4.1.1. The  $\text{AOD}_{440 \text{ nm}}$ ,  $\text{AOD}_{500 \text{ nm}}$ , and  $\text{AOD}_{675 \text{ nm}}$  accuracy is shown in Section 4.1.2, and the performance of the  $\text{AE}_{440-675 \text{ nm}}$  and  $\text{FMF}_{500 \text{ nm}}$  is discussed in Section 4.1.3.

#### 4.1.1. Sensitivity Analysis on Model Input Parameters and ML Application

Two sensitivity exercises are performed to achieve the best model accuracy. Firstly, the model performance is evaluated using (a) sky information from all 60 pixels as already described in Figure 1c–e (blue bars in Figure 2) or (b) pixels at specific zenith angles ( $25\text{--}35^\circ$  and  $51\text{--}61^\circ$ ) (green bars in Figure 2). Figure S2 shows the channels' intensity variations against the zenith angle. The first range of the specific zenith angles ( $25\text{--}35^\circ$ ) forms an area close to the sun, with intense scattering effects, while the second group of specific zenith angles ( $51\text{--}61^\circ$ ) is located at a distant area from the sun, where the channel intensity still reduces with increasing zenith angle. At higher zenith angles, the color intensity shows an almost stable pattern (Figure S2). The model performance for all aerosol parameters reduces as zenith angle reduces. R and RMSE ranges are 0.87–0.94 and 0.06–0.16 using all pixels, while for the specific pixels, the corresponding ranges are 0.84–0.93 and 0.07–0.18.



**Figure 2.** (a) Correlation coefficient, R, and (b) root mean square error (RMSE) at four different model scenarios. “All-points” refer to the 60 pixels of Figure 1, and “Specific-points” indicate only a confined range of zenith angles ( $25\text{--}35^\circ$  and  $51\text{--}61^\circ$ ). The term “+ TCWV” means that the light gradient boosting machine (LGBM) model includes the total column water vapor (TCWV).

Furthermore, the model's performance is also evaluated (a) with TCWV (Figure 2; orange and red bars) and (b) without TCWV (Figure 2; blue and green bars) as input in the training stage. LGBM, including TCWV, shows increasing accuracy with R and RMSE ranges of 0.89–0.95 and 0.05–0.15 (Figure 2). Lower values are expected for models without TCWV as input (R: 0.88–0.94 and RMSE: 0.06–0.18). Generally, the user can select which

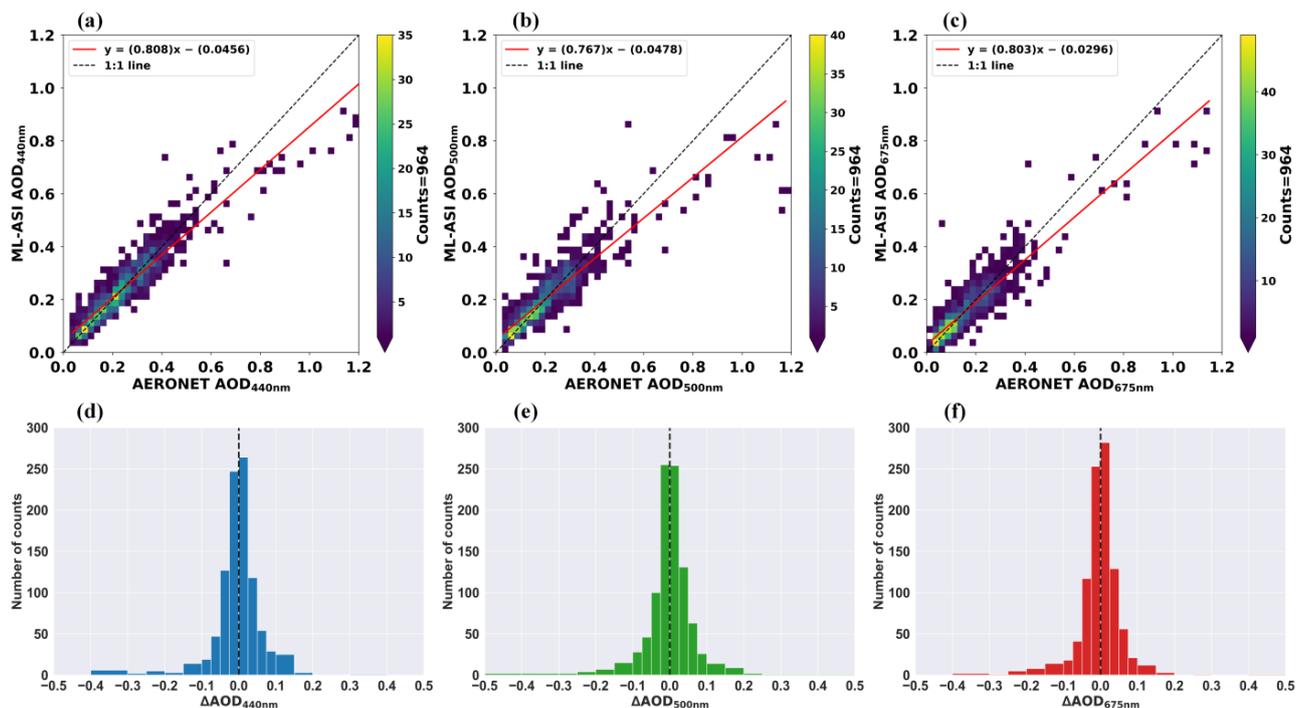
scenario fits the application. The model slightly improves by including the TCWV, but ASIs alone provide accurate estimates that can be used for further analysis.

Based on the above sensitivity analysis, the following sections only discuss the scenario with 60 all-sky pixels of Figure 1c–e and TCWV.

In addition, the performance of other ML algorithms is also investigated using the best input scenario described above. Eight ML models with different prediction mechanisms and algorithmic structures (Table S1) are evaluated, namely, (1) linear-based: MARS (Multivariate Adaptive Regression Splines) and ANN (Artificial Neural Network), (2) tree-based: GBM, LGBM, XGBoost (Extreme Gradient Boosting Machine), and RF (Random Forest), (3) distance-based: KNN (K-Nearest Neighbors), and (4) kernel-based: SVM (Support Vector Machine). The ‘best’ ML model is selected based on overall performance, accuracy, and tuning/training time. The ML algorithms are evaluated on the same computer system following the hyperparameter tuning strategy described in Section 3.1. According to Figure S3, the model’s performances vary between 0.81–0.93 and 0.05–0.09 in terms of R (Figure S3a) and RMSE (Figure S3b). LGBM encompasses the highest R and lowest RMSE values. Its high performance and the relatively low tuning/training time (~2 min; Figure S3c) explain the selection of LGBM in this study.

#### 4.1.2. Aerosol Optical Depth Retrieval Performance

The ML-ASI AOD retrievals are compared against reference AOD measurements from the AERONET. Figure 3a–c display the scatter density plots of ML-ASI AOD against the AERONET AOD at each wavelength. The ML-ASI AOD correlates well with the AERONET, with R values extending from 0.89 to 0.93 at all wavelengths (Table 1). The negative values of MBE show an overall underestimation of ML-ASI for AOD<sub>440 nm</sub> and AOD<sub>675 nm</sub>, while ML-ASI slightly overestimates the AERONET for AOD<sub>500 nm</sub> (Table 1). The results in Figure 3 explain an overall underestimation of the estimated AODs, especially under high aerosol burden conditions (AOD > 0.5). However, such high AOD cases were limited in the available dataset. In addition, the dispersion error is relatively low, with RMSEs lying between 0.05 and 0.07, revealing the adequate retrieval accuracy of the ML model.



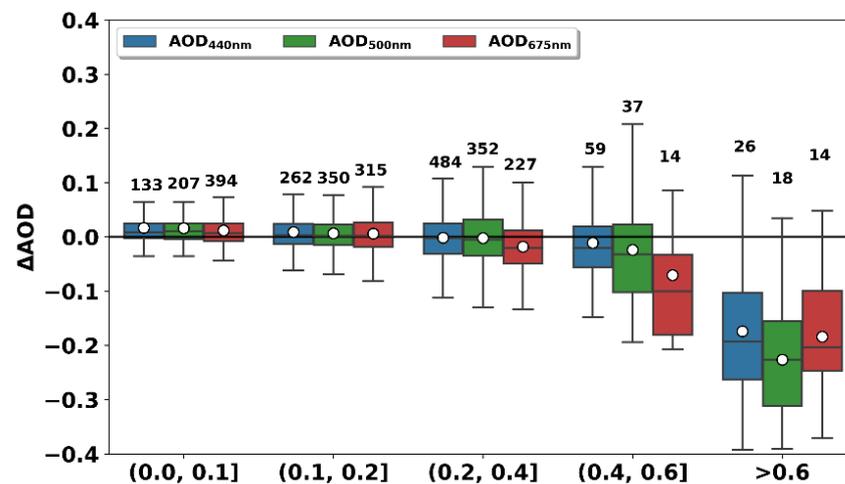
**Figure 3.** Density scatter plots of aerosol optical depth (AOD) retrieved from ML-ASI as a function of the AERONET AOD at (a) 440, (b) 500, and (c) 675 nm. Frequency distributions of  $\Delta$ AOD at (d) 440, (e) 500, and (f) 675 nm.  $\Delta$ AOD corresponds to the AOD difference between ML-ASI and AERONET.

**Table 1.** Mean bias error (MBE), root mean square error (RMSE), and correlation coefficient (R) metrics for aerosol optical depth (AOD) at the three wavelengths, along with their relative values (in parenthesis).

Variable	MBE (rMBE)	RMSE (rRMSE)	R
AOD <sub>440 nm</sub>	−0.0012 (−0.51%)	0.056 (23.1%)	0.93
AOD <sub>500 nm</sub>	0.000069 (0.033%)	0.066 (32.4%)	0.89
AOD <sub>675 nm</sub>	−0.0011 (−0.71%)	0.053 (33.9%)	0.92

Figure 3d–f represent the distribution of the differences between the AERONET and ML-ASI AODs ( $\Delta\text{AOD} = \text{ML-ASI AOD} - \text{AERONET AOD}$ ). The peak of the distribution at 440 and 675 nm is slightly biased to negative values, explaining the overall underestimation of ML-ASI AOD to the AERONET. Regarding the frequency distribution at 500 nm, the peak is perfectly centered to zero. The shape of the distribution implies that ML-ASI slightly overestimates the AERONET AOD (more positive AOD difference at the right tail of the statistical distribution). The AERONET's estimated AOD uncertainties in the visible spectrum are 0.01 [74]. More than 40% of ML-ASI AOD retrievals revealed differences lower than the AERONET's uncertainties, whereas 85% of ML-ASI AOD retrievals were below 0.05 against the AERONET (Figure 3).

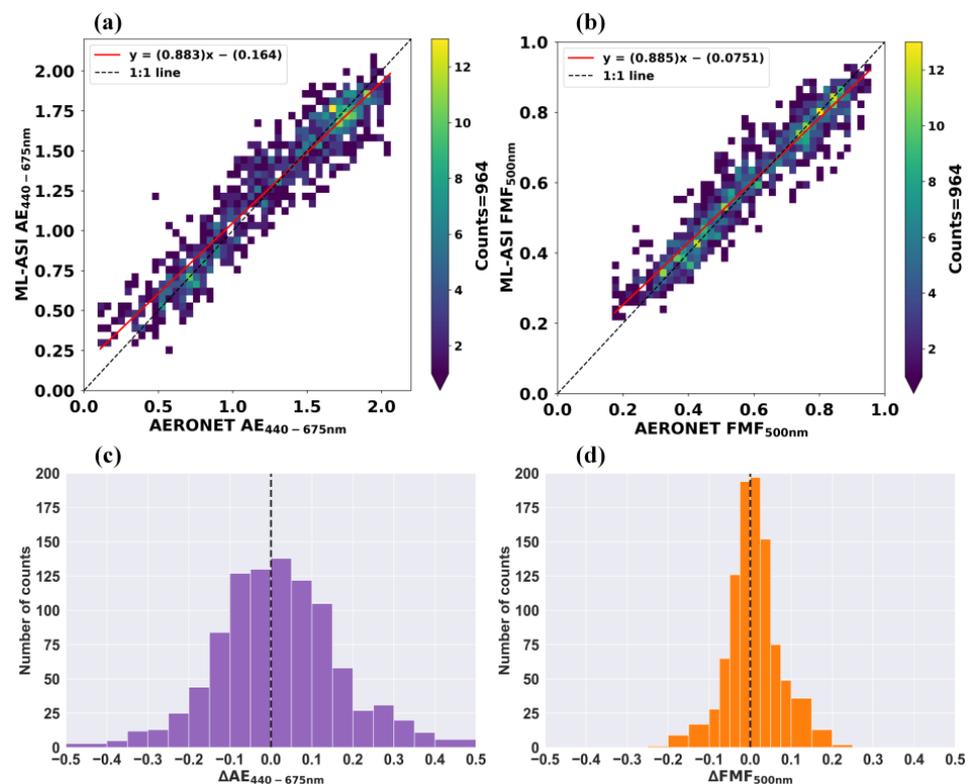
$\Delta\text{AOD}$  against AOD ranges are shown in Figure 4. Apparently, the amount of available data per bin is much higher for lower AODs. For  $\text{AOD} > 0.6$ , only 14–26 data points are detected within the testing dataset.  $\Delta\text{AOD}$  is relatively small for  $\text{AOD} < 0.4$ , with the boxplot median (and mean) value close to zero.  $\Delta\text{AOD}$  increases with AOD magnitude, where the median (and mean) is negative, indicating the substantial underestimation of ML-ASI AOD at high AERONET AODs.



**Figure 4.** Boxplots of the differences between ML-ASI and AERONET ( $\Delta\text{AOD}$ ) at 440 nm (blue), 500 nm (green), and 675 nm (red) for specific AOD ranges relying on AERONET data. The numbers above each boxplot correspond to the total number of measurements.

#### 4.1.3. $\text{AE}_{440-675 \text{ nm}}$ and $\text{FMF}_{500 \text{ nm}}$ Retrieval Performance

This subsection investigates the  $\text{AE}_{440-675 \text{ nm}}$  and  $\text{FMF}_{500 \text{ nm}}$  parameters. For the AERONET, the Ångström power law is used to calculate the spectral-dependent AE. In addition,  $\text{FMF}_{500 \text{ nm}}$  is derived through the fraction of the fine-mode AOD to the total AOD, accounting for the amount of anthropogenic aerosols in the atmospheric column. Figure 5 shows the scatter density plots for ML-ASI  $\text{AE}_{440-675 \text{ nm}}$  and  $\text{FMF}_{500 \text{ nm}}$  against the AERONET.



**Figure 5.** Density scatter plots of retrieved (a) Ångström Exponent between 440 and 675 nm ( $AE_{440-675\text{ nm}}$ ) and (b) Fine Mode Fraction at 500 nm ( $FMF_{500\text{ nm}}$ ) from ML-ASI as a function of AERONET values. Frequency distributions of (c)  $\Delta AE_{440-675\text{ nm}}$  and (d)  $\Delta FMF_{500\text{ nm}}$ .  $\Delta AE_{440-675\text{ nm}}$  and  $\Delta FMF_{500\text{ nm}}$  correspond to the  $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$  difference between ML-ASI and AERONET.

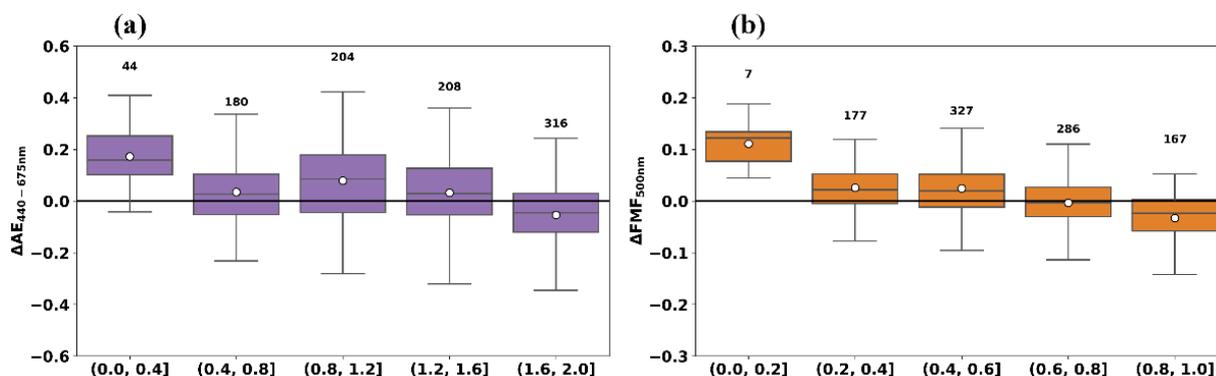
Both parameters are in good agreement with the observations.  $R$  is equal to 0.92 and 0.95 for  $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$ , respectively (Table 2). The positive values of MBE (0.007–0.017) indicate a slight overestimation of ML-ASI  $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$  to the AERONET. In addition, the low RMSEs, 0.15 and 0.057 for  $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$ , verify the good accuracy of the ML-ASI-derived aerosol parameters.

**Table 2.** Mean bias error (MBE), root mean square error (RMSE), and correlation coefficient ( $R$ ) metrics for Ångström Exponent between 440 and 675 nm ( $AE_{440-675\text{ nm}}$ ) and Fine Mode Fraction at 500 nm ( $FMF_{500\text{ nm}}$ ), along with their relative values (in parenthesis).

Variable	MBE (rMBE)	RMSE (rRMSE)	R
$AE_{440-675\text{ nm}}$	0.017 (1.4%)	0.15 (12.0%)	0.92
$FMF_{500\text{ nm}}$	0.007 (1.2%)	0.057 (9.67%)	0.95

The frequency distributions of  $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$  differences between ML and the AERONET ( $\Delta AE_{440-675\text{ nm}}$  and  $\Delta FMF_{500\text{ nm}}$ ) are shown in Figure 5c,d. The frequency distribution of  $\Delta AE_{440-675\text{ nm}}$  reveals the concentration of high differences in the distribution tails, while the slightly skewed pattern explains the minimal overestimation. The distribution of  $\Delta FMF_{500\text{ nm}}$  shows that most values are around the central tendencies (mean and median), similar to the AOD distributions of Figure 3, with a slight overestimation (MBE = 0.007). The AERONET's estimated  $FMF_{500\text{ nm}}$  and  $AE_{440-675\text{ nm}}$  uncertainties are around 0.1 and 0.2, respectively [83]. More than 90% and 80% of ML-ASI  $FMF_{500\text{ nm}}$  and  $AE_{440-675\text{ nm}}$  retrievals, respectively, lie within the AERONET's uncertainties (Figure 5).

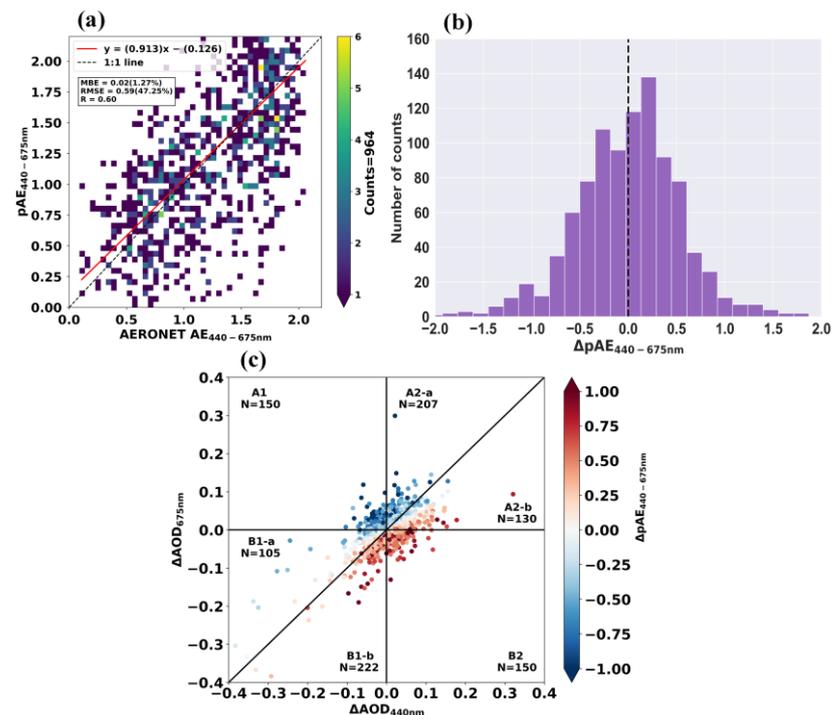
As in Section 4.1.2 for AOD,  $\Delta AE_{440-675\text{ nm}}$  and  $\Delta FMF_{500\text{ nm}}$  are calculated at specific ranges (Figure 6). The median of the boxplots is positive at almost all ranges, highlighting that ML-ASI predictions tend to overestimate the AERONET. The higher values of ML-ASI  $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$  explain that the ML-ASI model also overestimates the real size of aerosol particles. Only for  $AE_{440-675\text{ nm}} > 1.8$  and  $FMF_{500\text{ nm}} > 0.8$  are the medians negative, showing a minimal underestimation of ML-ASI retrievals to the AERONET.



**Figure 6.** Boxplots of the differences between ML-ASI and AERONET for (a) Ångström Exponent between 440 and 675 nm ( $AE_{440-675\text{ nm}}$ ) and (b) Fine Mode Fraction at 500 nm ( $FMF_{500\text{ nm}}$ ). The numbers above each boxplot correspond to the total number of measurements.

According to Section 3.1, an LGBM model is created for retrieving ML-ASI  $AE_{440-675\text{ nm}}$ . Alternatively, the spectral ML-ASI AOD retrievals at 440 and 675 nm can calculate  $AE_{440-675\text{ nm}}$  by applying the Ångström power law (hereafter abbreviated as  $pAE_{440-675\text{ nm}}$ ). Based on Figure 7a,  $pAE_{440-675\text{ nm}}$  provided relatively low accuracy (RMSE = 0.59) and overall performance ( $R = 0.60$ ) against the AERONET  $AE_{440-675\text{ nm}}$ . The latter can also be observed by the wider dispersion of the corresponding differences ( $\Delta pAE_{440-675\text{ nm}}$ ) against the AERONET  $AE_{440-675\text{ nm}}$  (histogram of Figure 7b). For instance, in several cases (Figure 7b), the ML retrievals differ by more than one value from the reference measurements. The poor performance of the  $pAE_{440-675\text{ nm}}$  could be attributed to: (a) the independence of the LGBM models in the AOD retrievals and (b) the sensitivity of the Ångström power law to small changes in the spectrally determined AOD. For example, a raw measurement can be overestimated from the ML-ASI  $AOD_{440\text{ nm}}$  and underestimated from the ML-ASI  $AOD_{675\text{ nm}}$  and vice versa, revealing significant differences between the  $pAE_{440-675\text{ nm}}$  and the AERONET  $AE_{440-675\text{ nm}}$ .

The scatterplot of Figure 7c represents the differences between the ML-ASI and AERONET AODs and how they are related to the corresponding  $\Delta pAE_{440-675\text{ nm}}$ . The scatterplot is divided into six areas (from A1 to B2) based on the  $AE_{440-675\text{ nm}}$  difference signs. For cases with different signs in the  $\Delta AOD$ , like in A1 (ML-ASI  $AOD_{675\text{ nm}}$  overestimates and ML-ASI  $AOD_{440\text{ nm}}$  underestimates the AERONET),  $pAE_{440-675\text{ nm}}$  substantially underestimates the AERONET  $AE_{440-675\text{ nm}}$  (blue area in the color bar of Figure 7c). A contrasting result is observed in the B2 area, which is characterized by substantial overestimation. For  $\Delta AOD$ s differences with similar signs (A2 and B1 in Figure 7c),  $pAE_{440-675\text{ nm}}$  reveals lower biases, and the bias sign depends on the  $\Delta AOD$  sign. If  $\Delta AOD_{440\text{ nm}}$  is higher than  $\Delta AOD_{675\text{ nm}}$ ,  $pAE_{440-675\text{ nm}}$  overestimates the AERONET (B1-b and A2-b are of Figure 7c). The opposite behavior is documented for the B1-a and A2-a areas of Figure 7c.

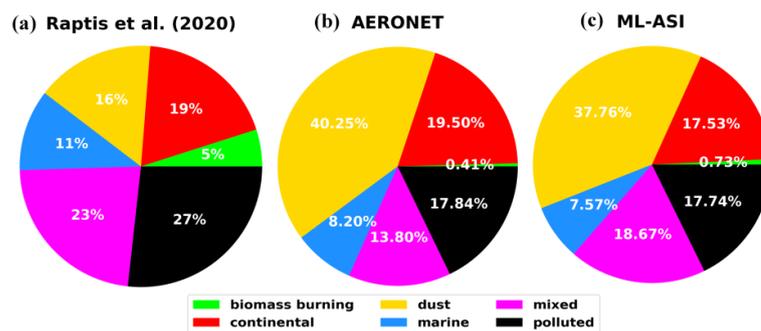


**Figure 7.** (a) Density scatter plot of the ML-ASI Ångström Exponent between 440 and 675 nm ( $AE_{440-675\text{ nm}}$ ) by applying the Ångström power law (pAE) as a function of AERONET  $AE_{440-675\text{ nm}}$ ; (b) Frequency distributions of the differences between pAE and AERONET  $AE_{440-675\text{ nm}}$  ( $\Delta pAE_{440-675\text{ nm}}$ ); (c) Scatter plot between the differences of ML-ASI and AERONET aerosol optical depth (AOD) at 440nm (x-axis;  $\Delta AOD_{440\text{ nm}}$ ) and 675nm (y-axis;  $\Delta AOD_{675\text{ nm}}$ ). The color bar represents  $\Delta pAE_{440-675\text{ nm}}$ .

#### 4.2. Aerosol Type Classification

This section employs the predicted aerosol properties to perform aerosol-type classification. The aerosol classification scheme of Raptis et al. [72] is adapted which is applied for the same AERONET station (see Section 2.1). More specifically, aerosols are classified into six (6) main classes, (1) biomass burning, (2) continental, (3) dust, (4) marine, (5) mixed, and (6) polluted, relying on pre-defined fixed threshold limits of  $AOD_{440\text{ nm}}$  and  $AE_{440-870\text{ nm}}$ . The applied aerosol classification scheme defines the prevailing aerosol type. The latter does not secure the purest aerosol-type conditions, which requires additional information about aerosol absorptivity [62]. Both pie charts in Figure 8a,b used the AERONET retrievals at different study periods. More specifically, Figure 8a,b were generated by including the AERONET retrievals for the 2008–2018 and 01/2021–11/2011 time frames, respectively. Figure 8c includes the ML-ASI retrievals of the testing dataset during the study period (January 2021–November 2011).

In this study, a relatively higher percentage of dust particles was observed, highlighting the increasing dust activity (~40%; Figure 8b) within the study period compared to the climatological average value (16%; Figure 8a). In addition, lower percentages for the mixed (~14%; Figure 8b) and polluted (~18%; Figure 8b) aerosol types are encompassed compared to climatological average values (23% and 27%, respectively; Figure 8a). The rest of the aerosol types provided small differences (<5%). In regions with both anthropogenic and natural aerosol sources, differences in the yearly averages of aerosol types occurrence are likely to be documented. For example, aeolian dust particles from the Sahara Desert often disperse towards the Mediterranean basin, reaching the Southern European regions, but the seasonal dust concentration levels are strongly related to the variability of cyclones that occur during the year [84].



**Figure 8.** Pie charts of aerosol type classification based on Raptis et al.’s [72] aerosol classification scheme at the station of the National Observatory of Athens (NOA; 37.97° N, 23.72° E) in Thissio, Greece using (a) the Raptis et al. approach [72] (May 2008–September 2018), (b) AERONET, and (c) ML-ASI retrievals during the study period of this work, which is the testing dataset within the 1 January 2021–18 November 2021 time frame.

At first glance, ML retrievals seem to reproduce the aerosol classification scheme of the AERONET quite well, providing minor differences (<3%) for all aerosol types (Figure 8c). However, in the case of mixed aerosol type, the difference is almost 5.0%. Apart from the pie charts, the confusion matrix of Figure 9 gives more insights into the aerosol classification outcomes by presenting the correctly and falsely assigned aerosol types using the ML-ASI results. The last column of the confusion matrix displays the Precision of each aerosol type. In particular, the green/red percentage corresponds to the correctly/false assigned aerosol type from the ML-ASI divided by the true assigned aerosol type from the AERONET. Above each percentage the total number of aerosol type cases is presented, as assigned by ML-ASI. The last row of the last column is the Total Accuracy. Overall, the ML-ASI method captures 77.5% of the total cases. The percentage contribution of each aerosol type to Total Accuracy is presented by the diagonal elements of the matrix, with dust and continental types encompassing the highest contribution. The last row of the table shows the true number of cases for each aerosol type as revealed using the AERONET retrievals. The percentages of the last row are calculated accordingly, as discussed above, for the last column of the confusion matrix, evaluating the AERONET retrievals against ML-ASI.

		Confusion matrix						
ML-ASI	biomass burning	2				1	4	7 28.6% 71.4%
	continental		136		21	5	7	169 80.5% 19.5%
	dust	3	350	3	8			364 96.2% 3.8%
	marine	8	8	47	10			73 64.4% 35.6%
	mixed	1	20	30	8	86	35	180 47.8% 52.2%
	polluted	1	21			23	126	171 73.7% 26.3%
		4	188	388	79	133	172	964 77.5% 22.5%
		50.0%	72.3%	90.2%	59.5%	64.7%	73.3%	
		50.0%	27.7%	9.8%	40.5%	35.3%	26.7%	
		biomass burning	continental	dust	marine	mixed	polluted	
		AERONET						

**Figure 9.** The confusion matrix includes the 6 possible aerosol types based on the research of Raptis et al. [72]. The diagonal elements show the aerosol types that are correctly predicted, while the off-diagonal elements indicate false predictions. The 7th column of the confusion matrix represents the Precision (Equation (7)) of ML-ASI aerosol type classification against the AERONET, with the last row referring to the total ML-ASI Accuracy (Equation (6)).

A great advantage of this confusion matrix is that it can detect the falsely assigned aerosol types. For instance, ML-ASI has the tendency to falsely document the polluted aerosol type as mixed (35 out of 964 total cases, 3.6%), which is related to the underestimation of ML-ASI AE against the AERONET retrievals. A promising finding refers to the number of true predicted dust aerosol types (Precision > 95.0% of the cases). In addition, the continental, polluted, and marine aerosol types are adequately extracted with a precision exceeding 60%. Moderate precision (~50.0%) is calculated for the mixed aerosol type and low precision (<29.0%) for the biomass-burning aerosols.

## 5. Conclusions

The current study investigates the feasibility of using the detailed sky information from an all-sky imager to retrieve various aerosol optical properties and types using machine learning. The presented retrieval methodology uses the RGBs and the percentage of saturated pixels near the sun extracted from the all-sky imager along with the solar zenith angle and the total column water vapor to train the machine learning algorithm (here, it is Light GBM, LGBM). Five individual models were trained to retrieve aerosol properties analogous to aerosol burden ( $AOD_{440\text{ nm}}$ ,  $AOD_{500\text{ nm}}$ , and  $AOD_{675\text{ nm}}$ ) and size ( $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$ ). Then, the retrieved aerosol properties were used to perform an aerosol type classification.

A comprehensive analysis of the factors affecting the performance of the proposed methodology has been conducted by (i) varying the sky information in terms of sky pixel intensity, (ii) including the total column water vapor, and (iii) applying various machine learning algorithms. The best model performance was obtained by including more information about the sky, such as zenith angles distant from the sun and total column water vapor. Among the various algorithms, the LGBM revealed the highest performance and the best accuracy at a relatively low training time.

The retrieved aerosol properties (ML-ASI) correlated well with reference ground-based measurements (AERONET), recording R values of 0.89–0.93 for AOD and 0.92–0.95 for  $AE_{440-675\text{ nm}}$  and  $FMF_{500\text{ nm}}$ . The differences between the ML-ASI and AERONET AOD follow an increasing tendency with AOD magnitude. Regarding the properties for aerosol size determination, the highest  $\Delta AOD$  was revealed for the coarser particles.

The retrieved aerosol optical properties were also implemented to classify the prevailing aerosol type. The ML-ASI aerosol classification results indicated that the proposed retrieval methodology could predict the dominant aerosol types with relatively high precision (>60.0% in 4 out of 6 aerosol clusters).

The findings from the presented study highlight the feasibility of an ASI to retrieve aerosol optical properties accurately with the synergy of supervised learning under clear skies. Those ASI-based aerosol properties revealed the ability to identify the prevailing aerosol type. The applicability of the ML-ASI system to further retrieve aerosol properties under partial cloud conditions is under investigation.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/atmos14081266/s1>, Figure S1: (a) Scatter plot between CAMS and AERONET total column water vapor retrievals; Figure S2: (a) Red, (b) Green, and (c) Blue channel intensity against zenith angle. The zenith angle points are shown in Figure 1 for the principal plane (straight line 1), ranging from sun's center point (zenith angle = 0) and to sun's area up to 75°. Different colors refer to three different AOD ranges. Blue, red, and green colors represent AOD range values that are relatively low (0.1), moderate (0.1–0.3), and high (>0.3). The shaded areas around the lines correspond to  $\pm 1$  standard deviation bands. The two rectangles refer to the specific zenith angles (25–35° & 51–61°) which are used in Section 4.1; Figure S3: (a) R correlation coefficient and (b) RMSE for the eight different machine learning models by using the 60 pixels of Figure 1 and the total column water vapor as model inputs to retrieve  $AOD_{440\text{ nm}}$ . (c) Execution time for models training procedure including the tuning; Table S1: Machine Learning Architecture including their hyperparameters that are tuned during the training procedure.

**Author Contributions:** Conceptualization, S.-A.L., A.K. and V.S.; methodology, S.-A.L. and V.S.; software, S.-A.L.; validation, S.-A.L.; formal analysis, S.-A.L.; investigation, S.-A.L., V.S. and A.K.; data curation, S.-A.L., P.T. and C.-P.G.; writing—original draft preparation, S.-A.L.; writing—review and editing, S.-A.L., C.-P.G., V.S., K.E. and A.K.; visualization, S.-A.L.; resources, P.-I.R., K.E. and V.A.; supervision, A.K.; funding acquisition, A.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** The publication of this article has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH—CREATE—INNOVATE (project code: T2EDK—00681, MIS 5067617). The publication of this article has also been financed by the Research Committee of the University of Patras.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We acknowledge support of this work by the project DeepSky, co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH—CREATE—INNOVATE (project code: T2EDK—00681, MIS 5067617). We acknowledge support of this work by the project D3D, supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “2nd Call for H.F.R.I. Research Projects to support Faculty Members & Researchers” (Project Number: 4129). Kostas Eleftheratos acknowledges support from the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant” (Atmospheric parameters affecting Spectral solar IRradiance and solar Energy (ASPIRE), project number 300). The publication of this article has been financed by the Research Committee of the University of Patras.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

AOD	Aerosol optical depth
AE	Ångström exponent
AERONET	AERosol RObotic NETwork
ANN	Artificial neural network
ASI	All-sky imager
DNI	Direct normal irradiance
FMF	Fine mode fraction
GBM	Gradient boosting machine
GHI	Global horizontal irradiance
KNN	K-Nearest neighbors
LGBM	Light gradient boosting machine
MARS	Multivariate adaptive regression splines
MBE	Mean bias error
ML	Machine learning
ML-ASI	ML-ASI retrievals
pAE	AE calculated using Ångström power formula based on ML-ASI AODs
rMBE	Relative mean bias error
rRMSE	Relative root mean square error
RF	Random forest
RGB	Red-green-blue
RMSE	Root mean square error
RTM	Radiative transfer model
R	Pearson’s correlation coefficient
R <sup>2</sup>	Coefficient of determination
SAT	Sun-saturated area
SVM	Support vector machines

SZA	Solar zenith angle
TCWV	Total column water vapor
XGBoost	Extreme gradient boosting machine
$\Delta$	Difference between ML-ASI and AERONET

## References

1. Masson-Delmotte, V.; Zhai, P.; Pirani, A.; Connors, S.L.; Péan, C.; Berger, S.; Caud, N.; Chen, Y.; Goldfarb, L.; Gomis, M.I.; et al. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2021. [\[CrossRef\]](#)
2. Mishchenko, M.; Cairns, B.; Kopp, G.; Schueler, C.F.; Fafaul, B.; Hansen, J.; Hooker, J.; Itchkawich, T.; Maring, H.; Travis, L.D. Accurate monitoring of terrestrial aerosols and total solar irradiance: Introducing the Glory Mission. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 677–691. [\[CrossRef\]](#)
3. Cuneo, L.; Ulke, A.G.; Cerne, B. Advances in the characterization of aerosol optical properties using long-term data from AERONET in Buenos Aires. *Atmos. Pollut. Res.* **2002**, *13*, 101360. [\[CrossRef\]](#)
4. Kaufman, Y.; Tanré, D.; Boucher, O. A satellite view of aerosols in the climate system. *Nature* **2002**, *419*, 215–223. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Eck, T.F.; Holben, B.N.; Sinyuk, A.; Pinker, R.T.; Goloub, P.; Chen, H.; Chatenet, B.; Li, Z.; Singh, R.P.; Tripathi, S.N.; et al. Climatological aspects of the optical properties of fine/coarse mode aerosol mixtures. *J. Geophys. Res. Atmos.* **2010**, *115*. [\[CrossRef\]](#)
6. Liu, Y.; Yi, B. Aerosols over East and South Asia: Type Identification, Optical Properties, and Implications for Radiative Forcing. *Remote Sens.* **2022**, *14*, 2058. [\[CrossRef\]](#)
7. Holben, B.N.; Eck, T.F.; Slutsker, I.; Tanré, D.; Buis, J.P.; Setzer, A.; Vermote, E.; Reagan, J.A.; Kaufman, Y.J.; Nakajima, T.; et al. AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. *Remote Sens. Environ.* **1998**, *66*, 1–16. [\[CrossRef\]](#)
8. Remer, L.A.; Kleidman, R.G.; Levy, R.C.; Kaufman, Y.J.; Tanré, D.; Mattoo, S.; Vanderlei, M.; Ichoku, C.; Koren, I.; Yu, H.; et al. Global aerosol climatology from the MODIS satellite sensors. *J. Geophys. Res.* **2008**, *113*, D14. [\[CrossRef\]](#)
9. Levy, R.C.; Mattoo, S.; Munchak, L.A.; Remer, L.A.; Sayer, A.M.; Patadia, F.; Hsu, N.C. The Collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech.* **2013**, *6*, 2989–3034. [\[CrossRef\]](#)
10. Sayer, A.M.; Hsu, N.C.; Bettenhausen, C.; Jeong, M.J. Validation and uncertainty estimates for MODIS Collection 6 "deep Blue" aerosol data. *J. Geophys. Res. Atmos.* **2013**, *118*, 7864–7872. [\[CrossRef\]](#)
11. Kahn, R.A.; Gaitley, B.J.; Garay, M.J.; Diner, D.J.; Eck, T.F.; Smirnov, A.; Holben, B.N. Multiangle Imaging Spectroradiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. *J. Geophys. Res. Atmos.* **2010**, *115*, D23. [\[CrossRef\]](#)
12. Mei, L.; Rozanov, V.; Vountas, M.; Burrows, J.P.; Levy, R.C.; Lotz, W. Retrieval of aerosol optical properties using MERIS observations: Algorithm and some first results. *Remote Sens. Environ.* **2017**, *197*, 125–140. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Riffler, M.; Popp, C.; Hauser, A.; Fontana, F.; Wunderle, S. Validation of a modified AVHRR aerosol optical depth retrieval algorithm over Central Europe. *Atmos. Meas. Tech.* **2010**, *3*, 1255–1270. [\[CrossRef\]](#)
14. Prados, A.I.; Kondragunta, S.; Ciren, P.; Knapp, K.R. GOES Aerosol/Smoke Product (GASP) over North America: Comparisons to AERONET and MODIS observations. *J. Geophys. Res.* **2007**, *112*, D15201. [\[CrossRef\]](#)
15. Dubovik, O.; Herman, M.; Holdak, A.; Lapyonok, T.; Tanré, D.; Deuzé, J.L.; Ducos, F.; Sinyuk, A.; Lopatin, A. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. *Atmos. Meas. Tech.* **2011**, *4*, 975–1018. [\[CrossRef\]](#)
16. Sawyer, V.; Levy, R.C.; Mattoo, S.; Cureton, G.; Shi, Y.; Remer, L.A. Continuing the MODIS Dark Target Aerosol Time Series with VIIRS. *Remote Sens.* **2020**, *12*, 308. [\[CrossRef\]](#)
17. King, M.D.; Kaufman, Y.J.; Tanré, D.; Nakajima, T. Remote sensing of tropospheric aerosols from space: Past, present, and future. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 2229–2259. [\[CrossRef\]](#)
18. Li, Z.; Zhao, X.; Kahn, R.; Mishchenko, M.; Remer, L.; Lee, K.H.; Wang, M.; Laszlo, I.; Nakajima, T.; Maring, H. Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: A review and perspective. *Ann. Geophys.* **2009**, *27*, 2755–2770. [\[CrossRef\]](#)
19. Bilal, M.; Nichol, J.E.; Chan, P.W. Validation and accuracy assessment of a Simplified Aerosol Retrieval Algorithm (SARA) over Beijing under low and high aerosol loadings and dust storms. *Remote Sens. Environ.* **2014**, *153*, 50–60. [\[CrossRef\]](#)
20. Ghonima, M.S.; Urquhart, B.; Chow, C.W.; Shields, J.E.; Cazorla, A.; Kleissl, J. A method for cloud detection and opacity classification based on ground based sky imagery. *Atmos. Meas. Tech.* **2012**, *5*, 2881–2892. [\[CrossRef\]](#)
21. Kazantzidis, A.; Tzoumanikas, P.; Bais, A.F.; Fotopoulos, S.; Economou, G. Cloud detection and classification with the use of whole-sky ground-based images. *Atmos. Res.* **2012**, *113*, 80–88. [\[CrossRef\]](#)
22. Hasenbalg, M.; Kuhn, P.; Wilbert, S.; Nouri, B.; Kazantzidis, A. Benchmarking of six cloud segmentation algorithms for ground-based all-sky imagers. *Sol. Energy* **2020**, *201*, 596–614. [\[CrossRef\]](#)
23. Fabel, Y.; Nouri, B.; Wilbert, S.; Blum, N.; Triebel, R.; Hasenbalg, M.; Kuhn, P.; Zorzalejo, L.F.; Pitz-Paal, R. Applying self-supervised learning for semantic cloud segmentation of all-sky images. *Atmos. Meas. Tech.* **2022**, *15*, 797–809. [\[CrossRef\]](#)

24. Cheng, H.Y.; Yu, C.C. Block-based cloud classification with statistical features and distribution of local texture features. *Atmos. Meas. Tech.* **2015**, *8*, 1173–1182. [[CrossRef](#)]
25. Taravat, A.; Del Frate, F.; Cornaro, C.; Vergari, S. Neural networks and support vector machine algorithms for automatic cloud classification of whole-sky ground-based images. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 666–670. [[CrossRef](#)]
26. Zhang, J.; Liu, P.; Zhang, F.; Song, Q. CloudNet: Ground-Based Cloud Classification with Deep Convolutional Neural Network. *Geophys. Res. Lett.* **2018**, *45*, 8665–8672. [[CrossRef](#)]
27. Ye, L.; Cao, Z.; Xiao, Y.; Yang, Z. Supervised Fine-Grained Cloud Detection and Recognition in Whole-Sky Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7972–7985. [[CrossRef](#)]
28. Blanc, P.; Massip, P.; Kazantzidis, A.; Tzoumanikas, P.; Kuhn, P.; Wilbert, S.; Schüler, D.; Prah, C. Short-term forecasting of high resolution local DNI maps with multiple fisheye cameras in stereoscopic mode. *AIP Conf. Proc.* **2017**, *1850*, 140004. [[CrossRef](#)]
29. Kuhn, P.; Nouri, B.; Wilbert, S.; Prah, C.; Kozonek, N.; Schmidt, T.; Yasser, Z.; Ramirez, L.; Zorzalejo, L.; Meyer, A.; et al. Validation of an all-sky imager-based nowcasting system for industrial PV plants. *Prog. Photovoltaics Res. Appl.* **2018**, *26*, 608–621. [[CrossRef](#)]
30. Nouri, B.; Kuhn, P.; Wilbert, S.; Hanrieder, N.; Prah, C.; Zorzalejo, L.; Kazantzidis, A.; Blanc, P.; Pitz-Paal, R. Cloud height and tracking accuracy of three all sky imager systems for individual clouds. *Sol. Energy* **2019**, *177*, 213–228. [[CrossRef](#)]
31. Crisosto, C.; Hofmann, M.; Mubarak, R.; Seckmeyer, G. One-hour prediction of the global solar irradiance from all-sky images using artificial neural networks. *Energies* **2018**, *11*, 2906. [[CrossRef](#)]
32. Kamadinata, J.O.; Ken, T.L.; Suwa, T. Sky image-based solar irradiance prediction methodologies using artificial neural networks. *Renew. Energy* **2019**, *134*, 837–845. [[CrossRef](#)]
33. Jiang, J.; Lv, Q.; Gao, X. The ultra-short-term forecasting of global horizontal irradiance based on total sky images. *Remote Sens.* **2020**, *12*, 3671. [[CrossRef](#)]
34. Nouri, B.; Blum, N.; Wilbert, S.; Zorzalejo, L.F. A hybrid solar irradiance nowcasting approach: Combining all sky imager systems and persistence irradiance models for increased accuracy. *Sol. RRL* **2021**, *6*, 2100442. [[CrossRef](#)]
35. Logothetis, S.A.; Salamalikis, V.; Nouri, B.; Remund, J.; Zorzalejo, L.; Xie, Y.; Wilbert, S.; Ntavelis, E.; Nou, J.; Hendriks, E.; et al. Benchmarking of solar irradiance nowcast performance derived from all-sky imagers. *Renew. Energy* **2022**, *199*, 246–261. [[CrossRef](#)]
36. Chu, Y.; Pedro, H.T.C.; Li, M.; Coimbra, C.F.M. Real-time forecasting of solar irradiance ramps with smart image processing. *Sol. Energy* **2015**, *114*, 91–104. [[CrossRef](#)]
37. Caldas, M.; Alonso-Suárez, R. Very short-term solar irradiance forecast using all-sky imaging and real-time irradiance measurements. *Renew. Energy* **2019**, *143*, 1643–1658. [[CrossRef](#)]
38. Logothetis, S.A.; Salamalikis, V.; Wilbert, S.; Remund, J.; Zorzalejo, L.; Xie, Y.; Nouri, B.; Ntavelis, E.; Nou, J.; Hendriks, E.; et al. Solar irradiance ramp forecasting based on all-sky imagers. *Energies* **2022**, *15*, 6191. [[CrossRef](#)]
39. Dev, S.; Savoy, F.M.; Lee, Y.H.; Winkler, S. Estimating solar irradiance using sky imagers. *Atmos. Meas. Tech.* **2019**, *12*, 5417–5429. [[CrossRef](#)]
40. Chu, Y.; Li, M.; Pedro, H.T.C.; Coimbra, C.F.M. A network of sky imagers for spatial solar irradiance assessment. *Renew. Energy* **2022**, *187*, 1009–1019. [[CrossRef](#)]
41. Olmo, F.J.; Cazorla, A.; Alados-Arboledas, L.; López-Álvarez, M.A.; Hernández-Andrés, J.; Romero, J. Retrieval of the optical depth using an all-sky CCD camera. *Appl. Opt.* **2008**, *47*, 34. [[CrossRef](#)]
42. Cazorla, A.; Shields, J.E.; Karr, M.E.; Olmo, F.J.; Burden, A.; Alados-Arboledas, L. Technical Note: Determination of aerosol optical properties by a calibrated sky imager. *Atmos. Chem. Phys.* **2009**, *9*, 6417–6427. [[CrossRef](#)]
43. Huo, J.; Lu, D.R. Preliminary retrieval of aerosol optical depth from all-sky images. *Adv. Atmos. Sci.* **2010**, *27*, 421–426. [[CrossRef](#)]
44. Berk, A.; Anderson, G.; Acharya, P.; Bernstein, L.; Muratov, L.; Lee, J.; Fox, M.; Adler-Golden, S.; Chetwynd, J.; Hoke, M.; et al. MODTRAN5: A reformulated atmospheric band model with auxiliary species and practical multiple scattering options. *Proc. SPIE* **2005**, *5655*, 88–95. [[CrossRef](#)]
45. Servera, J.V.; Rivera-Caicedo, J.P.; Verrelst, J.; Muñoz-Marí, J.; Sabater, N.; Berthelot, B.; Camps-Valls, G.; Moreno, J. Systematic Assessment of MODTRAN Emulators for Atmospheric Correction. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 4101917. [[CrossRef](#)]
46. Kazantzidis, A.; Tzoumanikas, P.; Nikitidou, E.; Salamalikis, V.; Wilbert, S.; Prah, C. Application of Simple All-sky Imagers for the Estimation of Aerosol Optical Depth. *AIP Conf. Proc.* **2017**, *1850*, 140012. [[CrossRef](#)]
47. Mayer, B.; Kylling, A. Technical note: The libRadtran software package for radiative transfer calculations e description and examples of use. *Atmos. Chem. Phys.* **2005**, *5*, 1855–1877. [[CrossRef](#)]
48. Dubovik, O.; Lapyonok, T.; Litvinov, P.; Herman, M.; Fuertes, D.; Ducos, F.; Lopatin, A.; Chaikovsky, A.; Torres, B.; Derimian, Y.; et al. GRASP: A versatile algorithm for characterizing the atmosphere. *SPIE Newsroom* **2014**, *25*, 2-1201408. [[CrossRef](#)]
49. Román, R.; Antuña-Sánchez, J.C.; Cachorro, V.E.; Toledano, C.; Torres, B.; Mateos, D.; Fuertes, D.; López, C.; González, R.; Lapyonok, T.; et al. Retrieval of aerosol properties using relative radiance measurements from an all-sky camera. *Atmos. Meas. Tech.* **2022**, *15*, 407–433. [[CrossRef](#)]
50. Antuña-Sánchez, J.C.; Román, R.; Cachorro, V.E.; Toledano, C.; López, C.; González, R.; Mateos, D.; Calle, A.; de Frutos, Á.M. Relative sky radiance from multi-exposure all-sky camera images. *Atmos. Meas. Tech.* **2021**, *14*, 2201–2217. [[CrossRef](#)]
51. Scarlatti, F.; Gómez-Amo, J.L.; Valdelomar, P.C.; Estellés, V.; Utrillas, M.P. A Machine Learning Approach to Derive Aerosol Properties from All-Sky Camera Imagery. *Remote Sens.* **2023**, *15*, 1676. [[CrossRef](#)]

52. Gueymard, C.A. Temporal variability in direct and global irradiance at various time scales as affected by aerosols. *Sol. Energy* **2012**, *86*, 3544–3553. [[CrossRef](#)]
53. Vamvakas, I.; Salamalikis, V.; Benitez, D.; Al-Salaymeh, A.; Bouaichaoui, S.; Yassaa, N.; Guizani, A.; Kazantzidis, A. Estimation of global horizontal irradiance using satellite-derived data across Middle East-North Africa: The role of aerosol optical properties and site-adaptation methodologies. *Renew. Energy* **2020**, *157*, 312–331. [[CrossRef](#)]
54. Ruiz-Arias, J.A.; Gueymard, C.A. Worldwide inter-comparison of clear-sky solar radiation models: Consensus-based review of direct and global irradiance components simulated at the earth surface. *Sol. Energy* **2018**, *168*, 10–29. [[CrossRef](#)]
55. Book, H.; Poikonen, A.; Aarva, A.; Mielonen, T.; Pitkanen, M.R.A.; Lindfors, A.V. Photovoltaic system modeling: A validation study at high latitudes with implementation of a novel DNI quality control method. *Sol. Energy* **2020**, *204*, 316–329. [[CrossRef](#)]
56. Abreu, E.F.; Gueymard, C.A.; Canhoto, P.; Costa, M.J. Performance assessment of clear-sky solar irradiance predictions using state-of-the-art radiation models and input atmospheric data from reanalysis or ground measurements. *Sol. Energy* **2023**, *252*, 309–321. [[CrossRef](#)]
57. Seinfeld, J.H.; Pandis, S.N. *Atmospheric Chemistry and Physics from Air Pollution to Climate Change*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006. [[CrossRef](#)]
58. Dubovik, O.; Holben, B.; Eck, T.F.; Smirnov, A.; Kaufman, Y.G.; King, M.D.; Tanre, D.; Slutsker, I. Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations. *J. Atmos. Sci.* **2002**, *59*, 590–608. [[CrossRef](#)]
59. Lee, J.; Kim, J.; Song, C.H.; Kim, S.B.; Chun, Y.; Sohn, B.J.; Holben, B.N. Characteristics of aerosol types from AERONET sunphotometer measurements. *Atmos. Environ.* **2010**, *44*, 3110–3117. [[CrossRef](#)]
60. Giles, D.M.; Holben, B.N.; Eck, T.F.; Sinyuk, A.; Smirnov, A.; Slutsker, I.; Dickerson, R.R.; Thompson, A.M.; Schafer, J.S. An analysis of AERONET aerosol absorption properties and classifications representative of aerosol source regions. *J. Geophys. Res. Atmos.* **2012**, *117*. [[CrossRef](#)]
61. Che, H.; Qi, B.; Zhao, H.; Xia, X.; Eck, T.F.; Goloub, P.; Dubovik, O.; Estelles, V.; Cuevas-Agulló, E.; Blarel, L.; et al. Aerosol optical properties and direct radiative forcing based on measurements from the China Aerosol Remote Sensing Network (CARSNET) in eastern China. *Atmos. Chem. Phys.* **2018**, *18*, 405–425. [[CrossRef](#)]
62. Logothetis, S.-A.; Salamalikis, V.; Kazantzidis, A. Aerosol classification in Europe, Middle East, North Africa and Arabian Peninsula based on AERONET Version 3. *Atmos. Res.* **2020**, *239*, 104893. [[CrossRef](#)]
63. Logothetis, S.-A.; Salamalikis, V.; Kazantzidis, A. The impact of different aerosol properties and types on direct aerosol radiative forcing and efficiency using AERONET Version 3. *Atmos. Res.* **2021**, *250*, 105343. [[CrossRef](#)]
64. Lin, J.; Zheng, Y.L.; Shen, X.; Xing, L.; Che, H. Global Aerosol Classification Based on Aerosol Robotic Network (AERONET) and Satellite Observation. *Remote Sens.* **2021**, *13*, 1114. [[CrossRef](#)]
65. Gerasopoulos, E.; Amiridis, V.; Kazadzis, S.; Kokkalis, P.; Eleftheratos, K.; Andreae, M.O.; Andreae, T.W. Three-year ground-based measurements of aerosol optical depth over the Eastern Mediterranean: The urban environment of Athens. *Atmos. Chem. Phys.* **2011**, *11*, 2145–2159. [[CrossRef](#)]
66. Eleftheriadis, K.; Colbeck, I.; Housiadas, C.; Lazaridis, M.; Mihalopoulos, N.; Mitsakou, C.; Smolik, J.; Ždímal, V. Size distribution, composition and origin of the submicron aerosol in the marine boundary layer during the eastern Mediterranean “SUB-AERO” experiment. *Atmos. Environ.* **2006**, *40*, 6245–6260. [[CrossRef](#)]
67. Athanasopoulou, E.; Speyer, O.; Brunner, D.; Vogel, H.; Vogel, B.; Mihalopoulos, N.; Gerasopoulos, E. Changes in domestic heating fuel use in Greece: Effects on atmospheric chemistry and radiation. *Atmos. Chem. Phys.* **2017**, *17*, 10597–10618. [[CrossRef](#)]
68. Amiridis, V.; Zerefos, C.; Kazadzis, S.; Gerasopoulos, E.; Eleftheratos, K.; Vrekoussis, M.; Stohl, A.; Mamouri, R.E.; Kokkalis, P.; Papayannis, A.; et al. Impact of the 2009 Attica wild fires on the air quality in urban Athens. *Atmos. Environ.* **2012**, *46*, 536–544. [[CrossRef](#)]
69. Masoom, A.; Fountoulakis, I.; Kazadzis, S.; Raptis, I.-P.; Kampouri, A.; Psiloglou, B.; Kouklaki, D.; Papachristopoulou, K.; Marinou, E.; Solomos, S.; et al. Investigation of the effects of the Greek extreme wildfires of August 2021 on air quality and spectral solar irradiance. *EGUsphere* **2023**. [[CrossRef](#)]
70. Eleftheriadis, K. Long term variability of the air pollution sources reflected on the state of mixing of the urban aerosol. *Environ. Sci. Health* **2019**, *8*, 36–39. [[CrossRef](#)]
71. Tsiflikiotou, M.A.; Kostenidou, E.; Papanastasiou, D.K.; Patoulas, D.; Zarnmpas, P.; Paraskevopoulou, D.; Diapouli, E.; Kalt-sonoudis, C.; Florou, K.; Bougiatioti, A.; et al. Summertime particulate matter and its composition in Greece. *Atmos. Environ.* **2019**, *213*, 597–607. [[CrossRef](#)]
72. Raptis, I.-P.; Kazadzis, S.; Amiridis, V.; Gkikas, A.; Gerasopoulos, E.; Mihalopoulos, N. A Decade of Aerosol Optical Properties Measurements over Athens, Greece. *Atmosphere* **2020**, *11*, 154. [[CrossRef](#)]
73. Barreto, A.; Cuevas, E.; Granados-Munoz, M.-J.; Alados-Arboledas, L.; Romero, P.M.; Grobner, J.; Kouremeti, N.; Almansa, A.F.; Stone, T.; Toledano, C.; et al. The new sun-sky-lunar Cimel CE318-T multiband photometer—A comprehensive performance evaluation. *Atmos. Meas. Tech.* **2016**, *9*, 631–654. [[CrossRef](#)]
74. Giles, D.M.; Sinyuk, A.; Sorokin, M.G.; Schafer, J.S.; Smirnov, A.; Slutsker, I.; Eck, T.F.; Holben, B.N.; Lewis, J.R.; Campbell, J.R.; et al. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database—Automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. *Atmos. Meas. Tech.* **2019**, *12*, 169–209. [[CrossRef](#)]

75. Smirnov, A.; Holben, B.N.; Lyapustin, A.; Slutsker, I.; Eck, T.F. AERONET processing algorithms refinement. In Proceedings of the AERONET2004 Workshop, El Arenosillo, Spain, 10–14 May 2004; pp. 10–14.
76. Dubovik, O.; King, M.D. A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements. *J. Geophys. Res.* **2000**, *105*, 20673–20696. [[CrossRef](#)]
77. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
78. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Adv. Neural Inf. Process. Syst.* **2017**, *9*, 3147–3155. Available online: <https://dl.acm.org/doi/10.5555/3294996.3295074> (accessed on 28 June 2023).
79. Torres, B.; Dubovik, O.; Toledano, C.; Berjon, A.; Cachorro, V.E.; Lapyonok, T.; Litvinov, P.; Goloub, P. Sensitivity of aerosol retrieval to geometrical configuration of ground-based sun/sky radiometer observations. *Atmos. Chem. Phys.* **2014**, *14*, 847–875. [[CrossRef](#)]
80. Wriedt, T. Mie Theory: A Review. In *The Mie Theory: Basics and Applications*; Springer Series in Optical Sciences; Springer: Berlin/Heidelberg, Germany, 2012; p. 169. [[CrossRef](#)]
81. Rubin, J.I.; Reid, J.S.; Xian, P.; Selman, C.M.; Eck, T.F. A global evaluation of daily to seasonal aerosol and water vapor relationships using a combination of AERONET and NAAPS reanalysis data. *Atmos. Chem. Phys.* **2023**, *23*, 4059–4090. [[CrossRef](#)]
82. Inness, A.; Ades, M.; Agustí-Panareda, A.; Barré, J.; Benedictow, A.; Blechschmidt, A.M.; Dominguez, J.J.; Engelen, R.; Eskes, H.; Flemming, J.; et al. The CAMS reanalysis of atmospheric composition. *Atmos. Chem. Phys.* **2019**, *19*, 3515–3556. [[CrossRef](#)]
83. O'Neill, N.T.; Dubovik, O.; Eck, T.F. Modified Ångström Exponent for the Characterization of Submicrometer Aerosols. *Appl. Opt.* **2001**, *40*, 2368–2375. [[CrossRef](#)]
84. Gkikas, A.; Basart, S.; Hatzianastassiou, N.; Marinou, E.; Amiridis, V.; Kazadzis, S.; Pey, J.; Querol, X.; Jorba, O.; Gassó, S.; et al. Mediterranean intense desert dust outbreaks and their vertical structure based on remote sensing data. *Atmos. Chem. Phys.* **2016**, *16*, 8609–8642. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.