Climatology of aerosol optical properties in Cyprus based on aerosol type classification from AERONET and Lidar data

Francesco Scarlatti[1], Rodanthi-Elisavet Mamouri[1,2], Argyro Nisantzi[1,2], Athena Savva[1,2]

[1]Eratosthenes Centre of Excellence, 82 Franklin Roosevelt, 3012, Lemesos, Cyprus
[2]Department of Civil Engineering and Geomatic, Cyprus University of Technology, Limassol, 3036, Cyprus.
Presenting author email: francesco.scarlatti@eratosthenes.org.cy

INTRODUCTION AND SUMMARY

Aerosols are suspended particles in the atmosphere of solid or liquid nature. They affect weather and climate in a way that today is not fully understood. A crucial role on aerosol climate effect is played by their nature, chemical or physics, since these characteristics largely affects the optical (so, heating) properties of the particles themselves. Alike the gaseous components of the atmosphere, aerosols show great spatial and temporal variations also in the troposphere where they mostly reside.
Thus, a classification of the aerosol type along with the fingerprint of their optical properties, is widely needed in scientific study to address local, global and temporal effects of aerosols on weather and climate.

The present study constitutes an attempt to address the difficult but needed task of aerosol type classification based on few measurements from the global ground station network AERONET. The classification is based only on two optical intensive magnitudes retrieved by sun photometers, namely the Single Scattering Albedo (SSA at 440nm) and the Angstrom Exponent (AE at 440-870nm). The methodology relies on the strength of a machine learning technique (Random Forest Classifier: RF) in the typical task of classification. The power of prediction result is evaluated on the test (30%) part of the built dataset and shows an overall performance of 87% of accuracy on balanced data. Then, A test for the desert dust type if performed on Arabic peninsula station to show that this method (obtained on Saharan desert dust) is highly generalizable, showing an 87% of cases in Solar Village station to be classified as desert dust in the period literature says the cases there are mainly of dust nature. After that, a statistics of the available events in two Cyprus AERONET stations (in Limassol and Agia Marina) show a good agreement in the overall climatology of the two stations even if the trend evolution in times of relative abundances of the different types is different for the two cases. In the same stations a characterization of the optical aerosol magnitude SSA with the wavelength is also performed to show good similarities in the detected types of aerosols. In the remaining some cases of desert dust have been evaluated in the Limassol station with Poliphon method in other to confirm the desert dust classification (not shown here).
This study, finally, wants to conclude that a simple model based on two optical magnitudes widely available in many ground-based stations, would be helpful in characterizing the type of aerosol in the atmosphere and the relative optical properties for climatological studies.

DATA AND METHODS

To create the random forest classifier that will predict the aerosol types based on unseen data, we collected aerosol optical properties (SSA and AE) from AERONET site, on stations around the world in specific times. This because the scientific literature (CATRALL et al. 2005) reports that in specific location and times AERONET stations are characterized by one type of aerosol events (Table 1). The philosophy here is to train a machine learning model in supervised way: SSA and AE are given as features or input of the model, and output label data (the aerosol type) is inferred based on the location and time of measurements. The final dataset has been split in train (70%) and test (30%) sets. The train set has been further cut in 5-95 quantile due to the presence of unwanted outliers (fraction of the data that belongs to other types). The result for the train set in depicted in figure 1a.

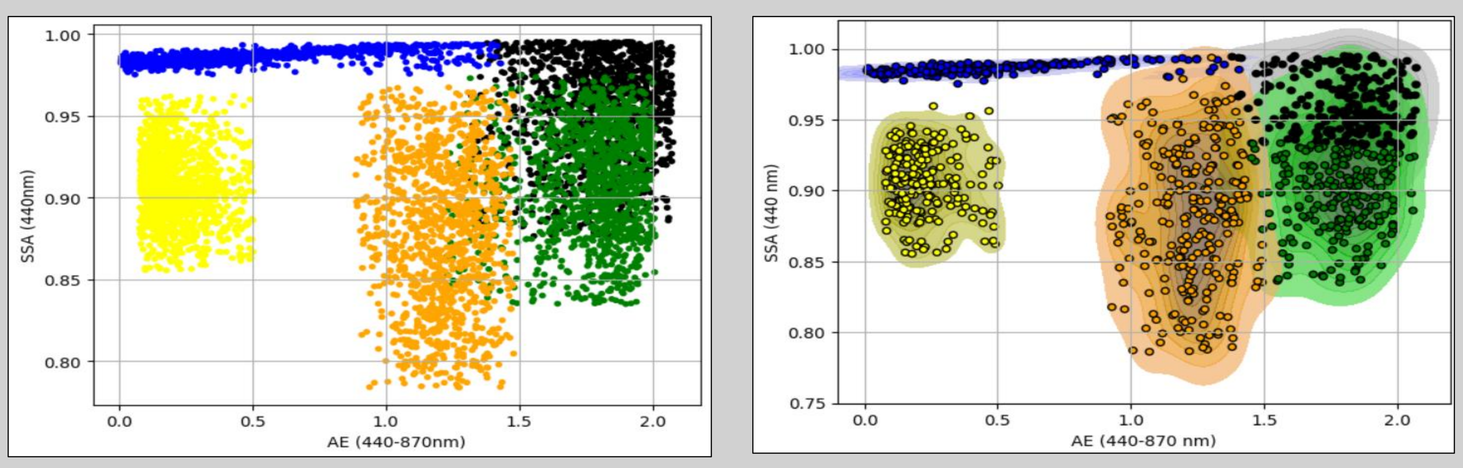
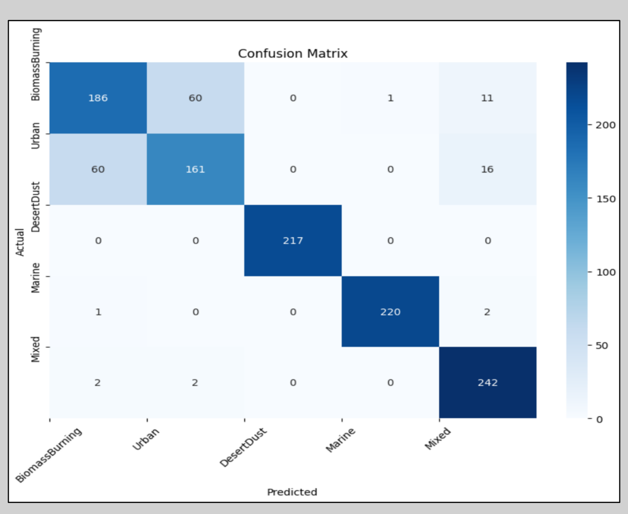


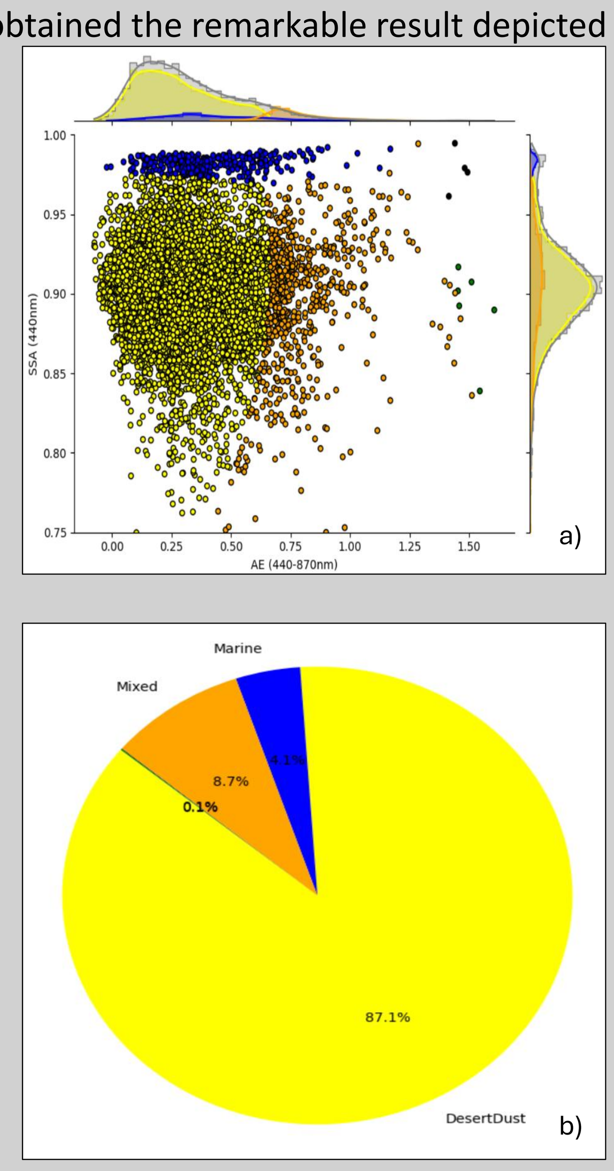
Figure 1. a) Train set graphical representation based on SSA, AE and types. The label types of aerosols are in different colours: Desert dust (yellow), Mixed (orange), Maritime (blue), Urban (black), Biomass Burning (green).
b) Same colours for the test set, the points are reported alongside the probability density function for the likelihood of the prediction of that specific type.



In figure 1b one can appreciate in the graphics the goodness of the prediction in the case of the test set. The same result is reported in figure 2 where a confusion matrix of the prediction is reported to confirm the good result.

Figure 2. The confusion matrix of the prediction of aerosol types
on the test set. The overall accuracy on the balanced set is 87%, lowered principally
by the cluster overlap of Biomass Burning and Urban types.

Finally, we confirm the goodness and robustness of the methodology applying the model to the data of Solar Village station of Arabic Peninsula, qualitatively different from Saharan dust. We obtained the remarkable result depicted in figures 3.



SITE	COUNTRY	TYPE	MONTH	YEAR	LAT	LONG
ALTA FLORESTA	BRASIL	BIOMASS BURNING	AUG-OCT	1993-2003	9.87 S	56.10 W
BANIZOUMBOU	NIGERIA	DUST	MARCH-JULY	1995-2004	13.55 N	2.67 E
LANAI	HAWAII	MARITIME	ALL	1996-2004	20.74 N	156.92 E
GSFC	MARYLAND	URBAN	JUNE-SEPT	1993-2025	38.99 N	76.84 W
BEIJING	CHINA	MIXED	ALL	2001-2025	39.98 N	116.38 E

Table 1. AERONET sites used in this study as data provider for the training phase. Each site is characterized by a specific type of aerosol based on literature (Cattrall et al. 2005), where the type is predominant in specific part of the year.

Figure 3. a) Graphics of the aerosol type prediction based on the model and data from Solar Village station in Saudi Arabia, along with types relative abundances on the x and y axis.
b) Relative abundances of types in Solar Village based on the model. The Desert Dust type is, as expected, the largely dominant one.

RESULTS AND ANALYSIS

We took advantage of the classification model to realize a annual and monthly analysis of the dataset of the two AERONET stations in Limassol and Agia Marina reported in the satellite map in figure 4.
The results of the analysis is depicted in figure5.

Figure 4. Satellite view of Cyprus island with the two AERONET sotes of Limassol and Agia Marina reported as yellow stars.

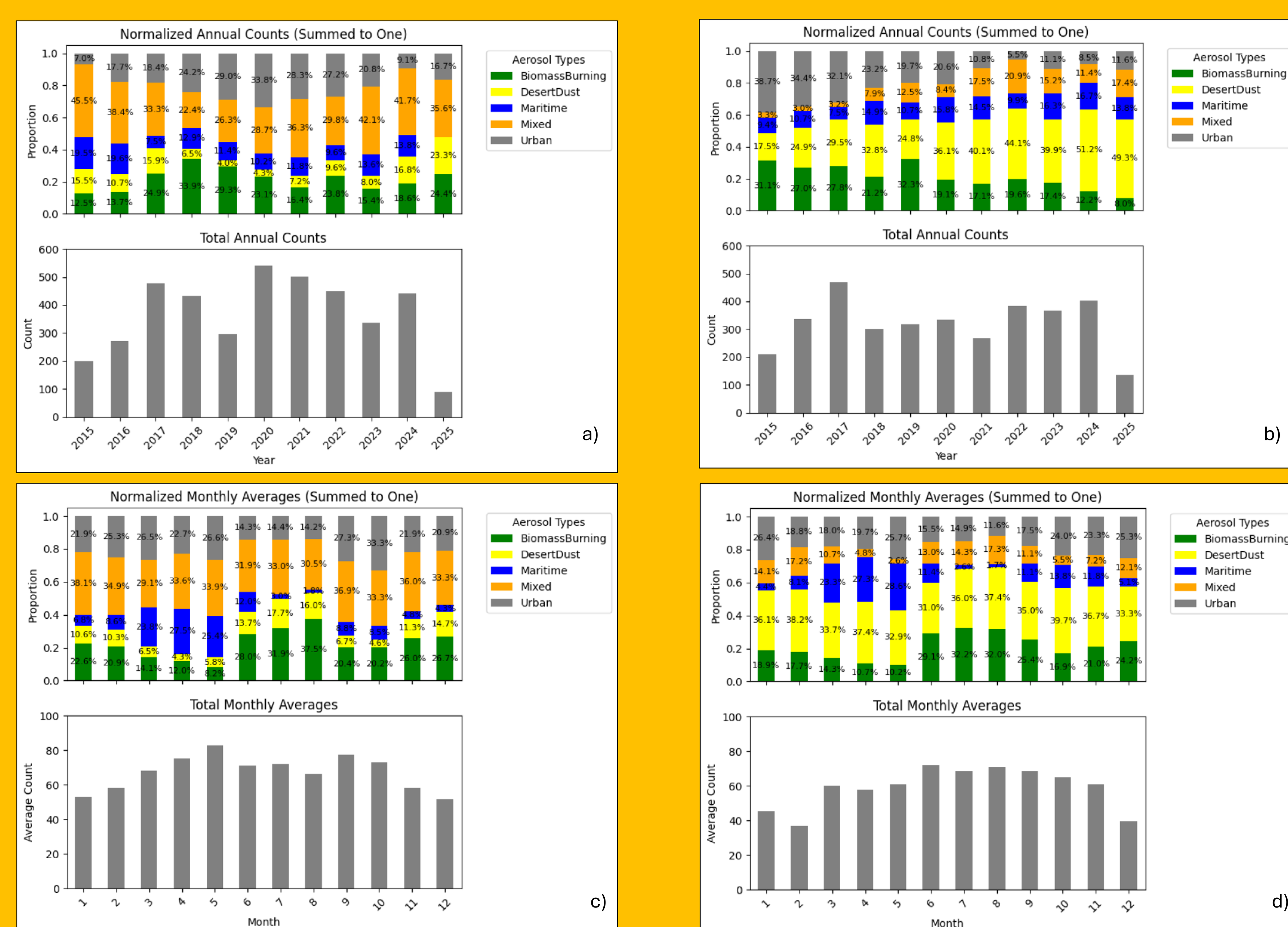
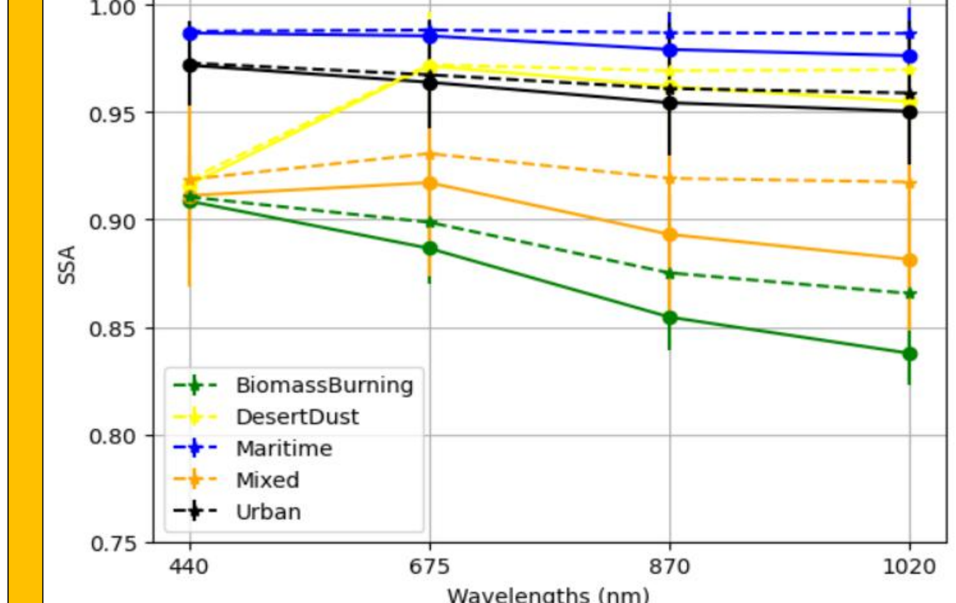


Figure 5. a) Normalized annual counts for the measurements classified accordingly to one type based on the model in the Limassol station. On the bottom, the corresponding total counts of the measurements. B) same as a) but for the site of Agia Marina. c) Normalized monthly averages for the measurements classified accordingly to one type based on the model in the Limassol station. On the bottom, the corresponding total counts averages of the measurements for that month. d) same as c) but for the site of Agia Marina.

We obtain an useful insight from the statistics in figure 5. Indeed, even if the all Cyprus island is in general affected completly by a dust intrusion, the isolated site of Agia Marina captures better the desert dust events whereas the Limassol urban area is more contaminated by mixed and indeed urban cases that seem to hide the desert component. Also, the annual reported evolution of desert dust from 2020 is appreciable only in Agia Marina in figure 4b.



Furthermore, we process and classify all the available occurrences of the two aerosol features of the model with the corresponding predicted type.
In this way we were able to obtain the characterization or the fingerprint of the SSA with the AERONET wavelengths based on our classification. The result is depicted in figure 6, where the comparison of the SSA trend with the type is shown between the two stations. The trends are highly comparable with the scientific literature, and the two sites share evident similarities: The average SSA of both sites lies in the uncertainty of the other average.

Figure 6. Comparison of the SSA trends with AERONET wavelengths for different aerosol types based on our model in two Cyprus sites of Limassol (--) and Agia Marina (-).

Acknowledgements: The study is supported by the 'EXCELSIOR': ERATOSTHENES: Excellence Research Centre for Earth Surveillance and Space-Based Monitoring of the Environment H2020 Widespread Teaming project (www.excelsior2020.eu). The 'EXCELSIOR' project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510, from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development and the Cyprus University of Technology. The authors acknowledge the ATARRI project funded by the European Union's Horizon Europe Twinning Call (HORIZON-WIDERA-2023-ACCESS-02) under the grant agreement No 101160258. AERONET data were obtained from the NASA AERONET program. We thank the Principal Investigators and site managers for maintaining the AERONET sites used in this work.

Cattrall, C., J. Reagan, K. Thome, and O. Dubovik (2005), Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types

Funded by the European Union

CONCLUSIONS

- The Random Forest classification model based on SSA and AE for aerosol typing shows robustness in validation session.
- The annual and monthly typing in Cyprus based on the model shows the site of Agia Marina to be more sensitive to desert dust intrusion than urban site of Limassol.
- The pure desert dust events compared with POLIPHON method are both perfectly detected by our method and POLIPHON. The AERONET data respect columnar retrievals, next step will be separate multiple layers cases that could lead to mixed aboundances in Limassol maritime and urban area , to check that the two stations see the same events.