

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

F. Scarlatti<sup>1</sup>, R.E. Mamouri<sup>1</sup>, A. Nisantzi<sup>1</sup>, A. Savva<sup>1</sup>

<sup>1</sup>Eratosthenes Centre of Excellence, 82 Franklin Roosevelt, 3012, Lemesos, Cyprus

Associated conference topics: 5.8, 5.9.

Presenting author email: francesco.scarlatti@eratosthenes.org.cy

A model for aerosol typing based on optical AERONET measurements is presented. It takes advantage of two main optical properties retrieved at several wavelengths by CIMELs of the network: the Single Scattering Albedo (SSA) and the Angstrom Exponent (AE). It is obtained by a Machine Learning (ML) methodology for classification purposes, a typical supervised task: given the inputs (optical properties) and specified outputs (relative aerosol type), it returns a parametric model that can relate new inputs to specific belonging classes. Different ML methodologies are presented and benchmarked to obtain the best one among the candidates. Principal component analysis (PCA) methodology has also been applied to each method to check if fewer input variables can perform with reasonable loss of accuracy with the benefit of less information needed. The input variables are selected in the AERONET sites and periods that the wide scientific literature states belong to a specific aerosol type. AERONET version 3 and level 1.5 had been chosen for data. The 1.5 level returns a higher amount of data with respect to level 2, which allowed us to train the ML model in a more reliable way. Also, the less refinement of the data is either managed from the model itself and a previous selection of the data based on the quantiles relative to the input data variability. Thus, we created a dataset of wavelength dependent SSA and AE related to aerosol type: the inputs and outputs to feed the ML candidate model. The chosen model passing the benchmark test is a Random Forest Classifier (RFC) with PCA down scaling the input variables to two: SSA at 440 nm and AE at 440-870 nm. The overall accuracy of the classification model on unseen data reached 95%.

The final ML model is then applied to the data of CIMEL situated in Limassol (Cyprus), belonging to AERONET and maintained by Cyprus University of Technology (CUT). The results show the relative presence of various aerosol types in the location using all available data, annual and monthly data, and other statistics.

To cross-validate this outcome of the ML method, we use the optical data of a collocated Polarized Raman Lidar and a sound existing methodology to classify aerosol type. The used input measurements from Lidar are the Particle Depolarization Ratio (PDR) and Lidar Ratio (LR). Once the methods have been cross validated, the two have also been used to obtain a climatology of other optical properties from AERONET and Lidar based

on the classification made by the two methodologies on different input data from the ones used in the generation of the models. Figure 1 shows a scheme of the workflow involved to obtain the climatology of the aerosol optical properties retrieved in Limassol ground-based stations.

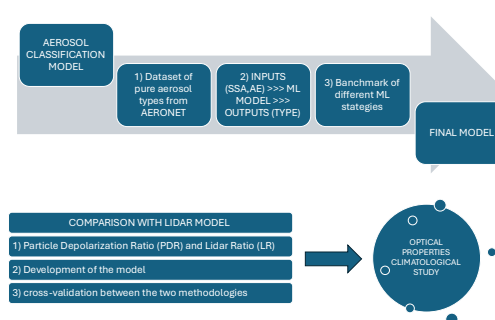


Figure 1. Workflow of the research: from the aerosol typing model to a climatology of its optical properties in Cyprus.

**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](http://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.

Catrrall, C., J. Reagan, K. Thome, and O. Dubovik (2005), Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations, *J. Geophys. Res.*, 110, D10S11, doi:10.1029/2004JD005124.