Aerosol Classification using AERONET Optical Properties: A Case Study at the Manila Observatory

AI201 Mini-Project

Hans Jarett Ong
Rossjyn Fallorina

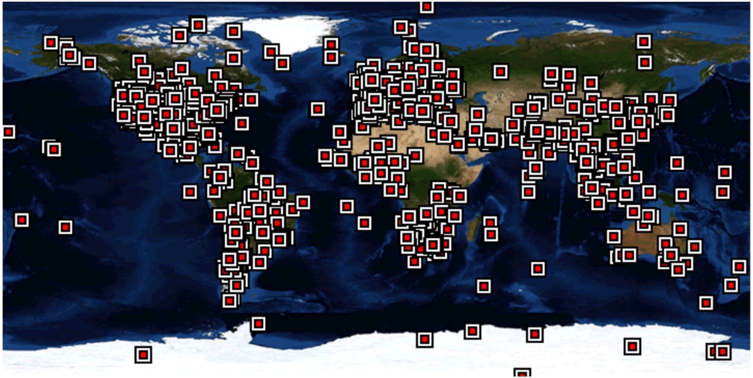
Outline

1. Background and Introduction
2. Objectives
3. Instrumentation and Data
4. Methodology
5. Results and Discussion
6. Conclusion

Problem Statement

1. The SEA region is known to be vulnerable to climate impacts and aerosols play a crucial role in earth and climate systems (IPCC).
2. The 7-SEAS mission was established to facilitate interdisciplinary research into the aerosol environment in SEA. 7-SEAS mainly uses remote sensing data to study aerosols.
3. AERONET is a global network of ground-based sun photometers that can continuously measure aerosol optical properties.

AERONET



Source: `https://aeronet.gsfc.nasa.gov/`

Objectives

- Create a model that can classify aerosols using their optical properties
- Accurately identify types of aerosols present at the Manila Observatory and understand their temporal variability
- Compare results with known events and trends; investigate relationship between aerosol properties and human activities

Instrumentation

- CIMEL Electronique CE-318 sun-sky radiometer/ sun photometer
  - 1.2° field of view
  - 2 measurement modes: Direct Sun and Sky Radiance
- Beer-Lambert-Bouguer Law:

$$V_\lambda = V_{0\lambda} d^2 e^{-\tau_\lambda m} \cdot t_y$$

Direct Sun Measurements

- Aerosol Optical Thickness

$$\tau_a = \tau - \tau_{H_2O} - \tau_{Rayleigh} - \tau_{O_3} - \tau_{NO_2} - \tau_{CO_2} - \tau_{CH_4}$$

- Angstrom Exponent

$$\alpha = -\frac{d \ln \tau_a}{d \ln \lambda}$$

Inversion Products

Optical Parameter	Symbol	Units/ Range of Values
Asymmetry Parameter	g	$0 \leq g(\lambda) \leq 1$
Effective Radius	r_{eff}	μm
Imaginary Refractive Index	$k(\lambda)$	$0.0005 \leq k(\lambda) \leq 0.5$
Real Refractive Index	$n(\lambda)$	$1.33 \leq n(\lambda) \leq 1.6$
Single Scattering Albedo	$\omega(\lambda)$	$0 \leq \omega(\lambda) \leq 1$
Volume Concentration	C_V	$\mu m^3/\mu m^2$
Volume Mean Radius	r_V	μm
Volume Size Distribution	$dV(r)/dlnr$	$\mu m^3/\mu m^2$

Aerosol Reference Clusters

Class Name	Abbreviation	Site	Period
Mineral Dust	MD	Solar Village, Saudi Arabia	Mar-Jul (1999-2015)
Polluted Dust	PD	Beijing, China	Whole Year (2001-2013)
Biomass Burning, Dark Smoke	BB-D	Mongu, Nigeria	Aug-Nov (1995-2009)
Biomass Burning, White Smoke	BB-W	Alta Floresta, Brazil	Aug-Oct (1995-2013)
Urban/Industrial (Developed Economy)	UI	GSFC, Maryland, USA	Jun-Sept (1993-2013)
Urban/Industrial (Developing Economy)	UI-D	Chen Kung Univ., Tainan, Taiwan	Whole Year (2002-2014)

Supervised Learning

- Uses labeled datasets to train algorithms to either classify data or predict outcomes
- For this study, three supervised learning algorithms were used:
  - Mahalanobis Distance Classifier (MC)

$$d_M(x) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

  - k-Nearest Neighbors Classifier (KNN)
  - Naive Bayes Classifier (NB)

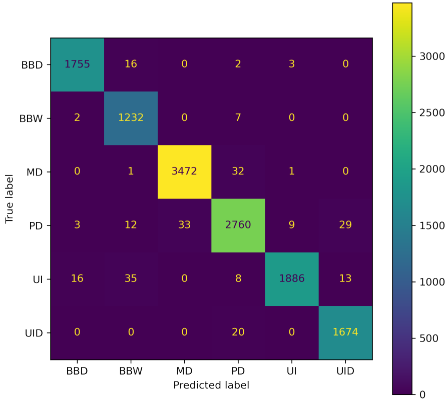
$$P(y|x_1, x_2, ..., x_n) = \frac{P(y)\prod_{i=1}^{n} P(x_i|y)}{\prod_{i=1}^{n} P(x_i)}$$

Model Performance

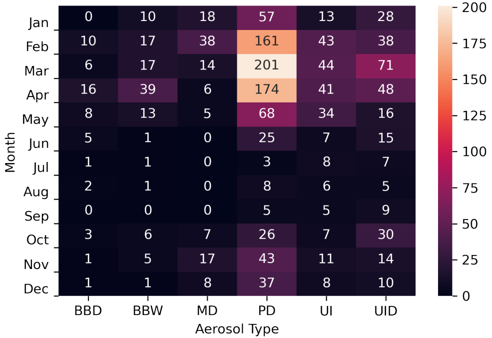
	MC	KNN	NB
Accuracy	0.812	**0.981**	0.727
Precision (Macro)	0.806	**0.979**	0.695
Precision (Weighted)	0.820	**0.982**	0.734
Recall (Macro)	0.796	**0.982**	0.711
Recall (Weighted)	0.812	**0.981**	0.728
F1 Score (Macro)	0.799	**0.980**	0.695
F1 Score (Weighted)	0.814	**0.981**	0.723

A comparison of the performance metrics for the three classifier models

Confusion Matrix for KNN Classifier



Monthly Seasonality of Aerosols in Manila Observatory



Comparing Aerosol Types Before and During COVID

Aerosol Type	Pre-COVID-19 (%)	COVID-19 (%)
BBD	3.378	0.000
BBW	6.947	5.882
MD	7.075	5.882
PD	50.988	23.529
UI	13.384	50.000
UID	18.228	14.706

Aerosol Types Before and During COVID-19

Conclusions

1. We tested 3 different classifiers and found KNN to perform the best for aerosol classification.
2. The aerosols observed in Manila Observatory are predominantly PD, UI, and UID.
3. There was a change in aerosol composition during COVID-19: with a decrease in PD and an increase in UI.
4. Next steps/ Recommendations:
   - Tree-based Algorithms
   - Extend the analysis to other sites