EPFL Design Project – SIE 2023

AeroSpec

Development of Tools to Understand Air Pollution Measurements From Novel Instrumentation

AeroSpec:
Nikunj Dundani

EPFL Supervisor:
Satoshi Takahama

Students:
Ali Waseem
Marie Lou Hirschy

Context

Aerosols play an important role in the lives around us, from negative health impacts to changing meteorological conditions[1,2]. The *State of Global Air* found that air pollution is the 4th leading risk of early death worldwide[3]. This makes aerosols and air pollution an important area of study. The chemical composition of particles can provide information about the particle (age, emission source) as well as how they can interact with their surrounds (oxidative stress on humans, likelihood to be cloud condensing nuclei, likely atmospheric reactions).

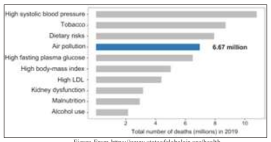


Figure From https://www.stateofglobalair.org/health

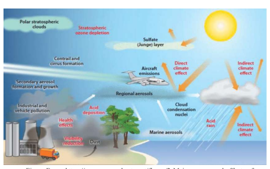


Figure From: https://www.researchgate.net/figure/2-Main-sources-and-effects-of-atmospheric-aerosol-particles-28_fig3_348931267

Chemical compositions are often obtained by sending samples to laboratory leading to high costs and low time resolution[4]. AeroSpec is engineering a product that can take continuous chemical composition measurements on-site using Fourier transform infrared spectrometry with the goal of **empowering research labs and environmental agencies to improve air quality and global health.**

IR Spectrometry Process

1. Particles collection: AeroSpec uses reusable ZnSe crystals with low spectroscopic interference, unlike commonly used Polytetrafluoroethylene (PTFE) filters[5].
2. FTIR analysis A beam of light is split in two. One beam is shone through the sample and the other used as a reference. A movable mirror varies the path of light changing the optical distance. The signal intensity as a function of optical distance is Fourier transformed and deconvoluted to create an absorbance spectra

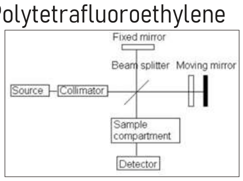
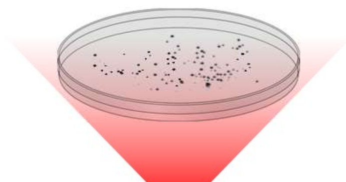
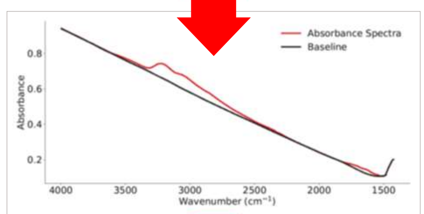
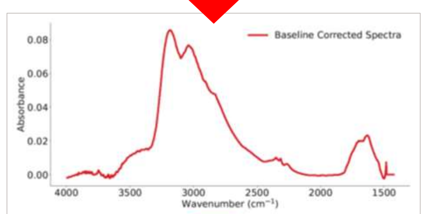


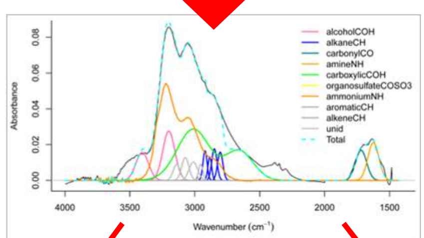
Figure from https://chem.libretexts.org/@go/page/432180.

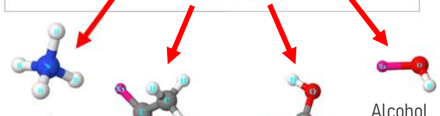
3. Variability in the collection environment causes a baseline effect where peaks in the absorbance caused by the particles become harder to analyze. To solve this **“Baseline Correction”** is applied











4. The baseline is subtracted from the absorbance, the baseline corrected spectra is analyzed to obtain a potential collection of functional groups that could possibly recreate the absorbance spectra that is observed

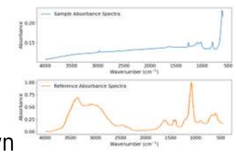
Objectives

- Analyze existing codes for use on new IR-AM instrument.
- Update provided python code to improve baseline correction on new IR-AM instrument compared to current code base.
- Test several baseline fitting methods on a variety of spectral data and obtain results on code performance for a variety of sample cases.
- **Deliver an optimized baseline correction method for use in the AeroSpec product using provided data.**

Methodology

To deliver an optimized baseline correction method, the existing baseline correction code was extended after which 4 functions were tested from the pybaselines package.

Data



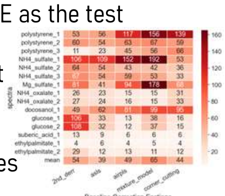
The data used for optimization was 46 spectra samples each containing 1 of 8 known compounds. Data was split into 70/30 train and test sets. The splitting was random while ensuring at least 1 of each compound was in the test set. Reference spectral data was found from Nistwebbook[6] and SpectraBase[7].

Optimization

The existing baseline correction code was optimized using a metropolis search algorithm due to its large parameter space. Other baseline functions were optimized using scipy.optimize with Leave One Out cross validation due to the small size of the train set. Root mean squared error (RMSE) between the reference spectra and baseline corrected spectra was used as an objective function to minimize. The data required normalization due to the unknown concentration of compound in the spectra.

Results

The optimized baseline functions were tested using mean overall RMSE as the test statistic. It was determined that the best performing baseline function was the asymmetric least squares from pybaselines.



Conclusion

An optimized baseline correction method was provided to AeroSpec furthering them on their goal to provide comprehensive data to the world. Baseline correction is important when designing a product to test air quality in variable environmental conditions. Creating a product that can provide continuous chemical composition measurements is another step forward towards better understanding the intricacies that make up air pollution's impact on human health and our world at large.

Sciences et ingénierie de l'environnement

References
[1] Exposure health impacts of air pollution. url: https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/exposure-air-pollution.

[2] Aerosols and Climate. url: https://www.gfdl.noaa.gov/aerosols-and-climate/.

[3] Impacts on your Health. url: https://www.stateofglobalair.org/health.

[4] Paul A. Solomon et al. U.S. National PM2.5Chemical Speciation Monitoring Net-works—CSN and IMPROVE: Description of networks. en. Nov. 2014. doi: 10.1080/10962247.2014.956904. url: http://dx.doi.org/10.1080/10962247.2014.956904.5

[5] N. Dudani and S. Takahama. "Design and fabrication of an electrostatic precipitator for infrared spectroscopy". In: Atmospheric Measurement Techniques 15.16 (2022), pp. 4693–4707. doi: 10.5194/amt-15-4693-2022. url: https://amt.copernicus.org/articles/15/4693/2022/.

[6] NIST Office of Data and Informatics. NIST Chemistry Webbook. url: https://webbook.nist.gov/chemistry/

[7] Online Spectral Database. url: https://spectrabase.com/