Name: Alec Brooks
Institution: UNLV, HDRFS Internship

Project Title: Developing a Data Pipeline for CASS-Based Aerosol Analysis

Email Contact: Alec_brooks@mail.tmcc.edu

Mentor(s)/Co-Author(s): *Antony Chen Ph. D., Olusanya Onamuti M.S.*

Abstract:

Carbonaceous aerosols (CA) are complex mixtures with significant implications for atmospheric chemistry, climate dynamics, and human health. Accurately apportioning these aerosols into distinct components is critical for understanding their sources and environmental impacts. Differentiating between natural and anthropogenic contributions, such as those from wildfires and vehicular emissions, offers valuable insight into the role of carbonaceous aerosols in air quality and climate systems.

Traditional methods for aerosol speciation are often time-consuming and poorly suited for real-time analysis. The Carbonaceous Aerosol Speciation Sampler (CASS) improves this process by enabling near real-time measurement of Total Carbon (TC) and Black Carbon (BC) using two core instruments: the Total Carbon Analyzer (TCA) and the Aethalometer (AE33). While the instrumentation provides high-resolution data, the raw outputs are not readily accessible for immediate interpretation.

To address this limitation, a cross-platform, Python-based application was developed to automate the data pipeline from ingestion to analysis. The pipeline consists of three stages: cleaning raw instrument outputs, storing them in a structured SQL database, and retrieving relevant datasets for automated analysis. AE33 data, recorded every minute and stored in daily .dat files, and TCA data, appended every 30 minutes to a single .csv file, are preprocessed to remove formatting inconsistencies. Cleaned datasets are indexed by timestamp to ensure integrity and avoid duplication.

Users can define custom time ranges and averaging intervals. Afterward, the software generates a comprehensive report that includes aggregated and calculated values, R-squared correlation metrics, and both diurnal and time series plots. By automating the workflow, the tool improves data accessibility, consistency, and processing speed, facilitating both real-time and historical insights into carbonaceous aerosol behavior.