Bridging the Last Mile with Open-Source Advancements: Empowering Communities through Fusion of Aerosol Optical Depth (AOD) Products from Multi-Satellite Sensors

Zhaohui Zhang[1,2], J. C. Wei[2], R. C. Levy[2] , S. Zhao[3], N. Gutkin[4], P. Gupta [5], X. Pan [1,2]

1. ADNET Systems Inc, Bethesda, MD 20817 USA 2. NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA 3.Science Systems and Applications, Inc, Lanham, MD 20709, USA 3. Maryland University, College Park, MD 20742, USA 4. Virginia Tech University, Blacksburg, VA 24061, USA 5.Goddard Earth Sciences Technology And Research (GESTAR), University Space Research Association (USRA), Columbia, Maryland, USA

zhaohui.zhang@nasa.gov

ED23D-0802
AGU23 Abstract ID: 1327660

1.Introduction

Aerosol Optical Depth (AOD) is a crucial parameter for understanding atmospheric aerosol distribution and their impact on climate and air quality. With the growing number of Earth observation satellites, there is an abundance of AOD products derived from various sensors onboard both geostationary and low-orbit satellites. The availability of multiple datasets provides an opportunity to harness the strengths of each sensor and create comprehensive and accurate AOD datasets for climate and air quality studies at different temporal and spatial scales.

Our NASA aerosol MEaSURES project has made significant strides in recent years by undertaking the ambitious task of developing an open-source package tailored for fusing AOD products from different sources. The package is based on OOP (Object-Oriented Programming) design and is implemented in Python modules. Generic interfaces enable easy inclusion of large and heterogeneous data. The package may be utilized to produce harmonized AOD datasets with enhanced spatial and temporal coverage. The latest version of the package is able to process and integrate the dark-target AOD data from six different sensors: AHI Himawari-8, ABI GOES-West, ABI GOES-East, MODIS AQUA, MODIS TERRA, and VIIRS SNPP. Rigorous validation and intercomparison studies have been performed to assess the accuracy and reliability of the fused AOD product against ground-based measurements and reference datasets.

The open-source nature of the developed package ensures transparency, reproducibility, and community engagement. The research community and stakeholders can access, contribute to, and further improve the fusion methodology, making it adaptable to other studies, or expanding it to include new satellite data as they become available.

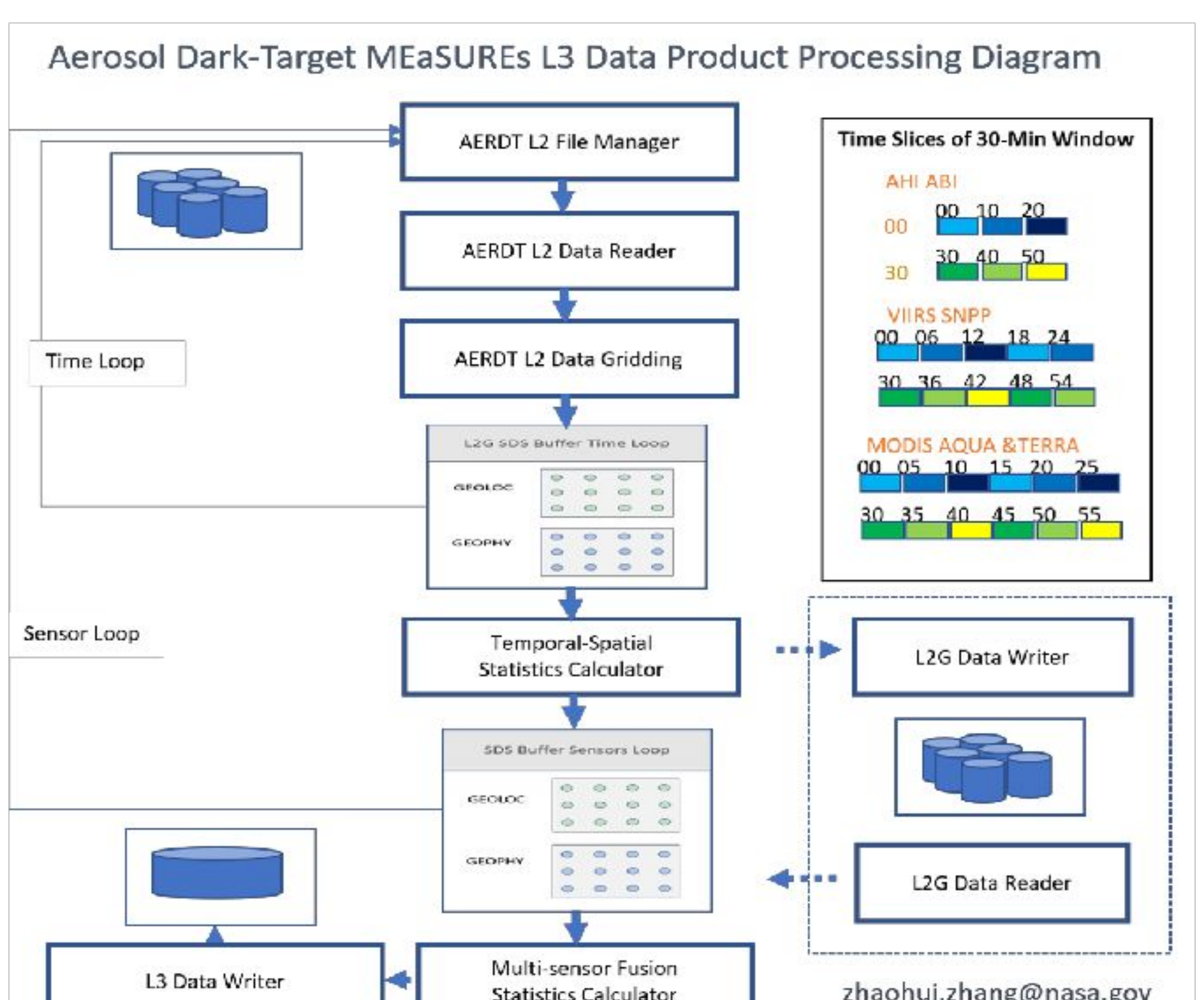
In this poster presentation, we will introduce the accomplishments and challenges faced during the development of the open-source package for AOD data fusion, and demonstrate the advantages of combining AOD products from the six aforementioned satellite sensors. The presentation aims to foster discussions, collaborations, and future directions in integrating Earth observation and remote sensing data, which may contribute to a better understanding of atmospheric aerosols and their impacts on our environment.

2. Aerosol Optical Depth (AOD) Level-2 Data Fusion Diagram

The AOD fusion package is implemented in Python to utilize its extensive support libraries. Python itself is open source and free, which enables easy community collaborations and software distribution. The current AOD fusion package version is developed for fusing the dark-target AOD Level 2 (L2) data from the six satellites of the NASA aerosol MEaSUREs project.

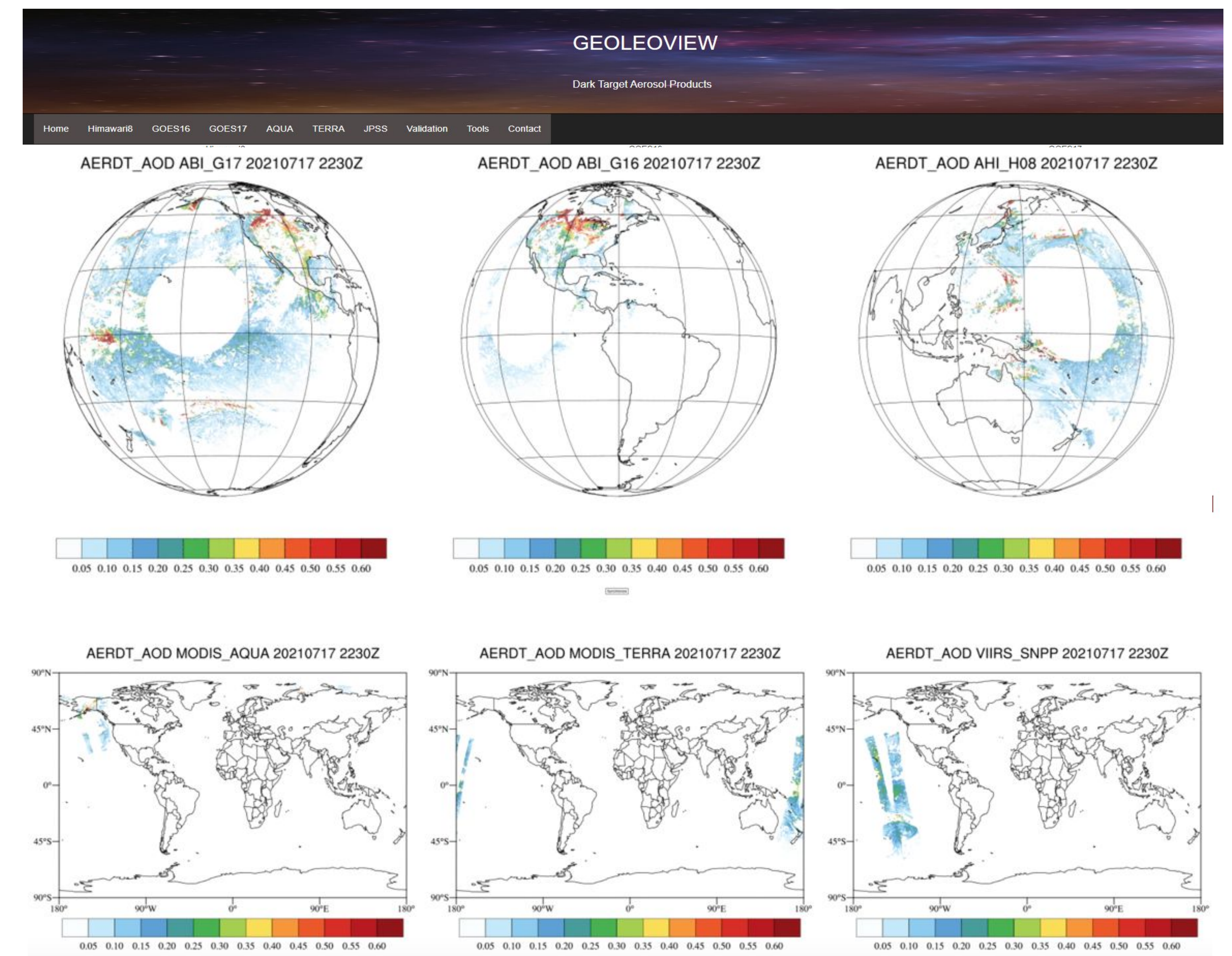
The fusion package is able to process data in NetCDF, HDF and HDF5 formats, and in different temporal and spatial resolutions.The fused data product (Level-3) is in NetCDF4, enabling easy re-processing and visualization with other Python support libraries. The package also allows users to customize their own input time and space domains and the output time and space resolutions.

The fusion package serves as the prototype for inclusion of many other AOD data sources in the future, e.g, the aerosol Deep-Blue L2 data.



3. Example AOD Level-2 Data Inputs of the Fusion Package

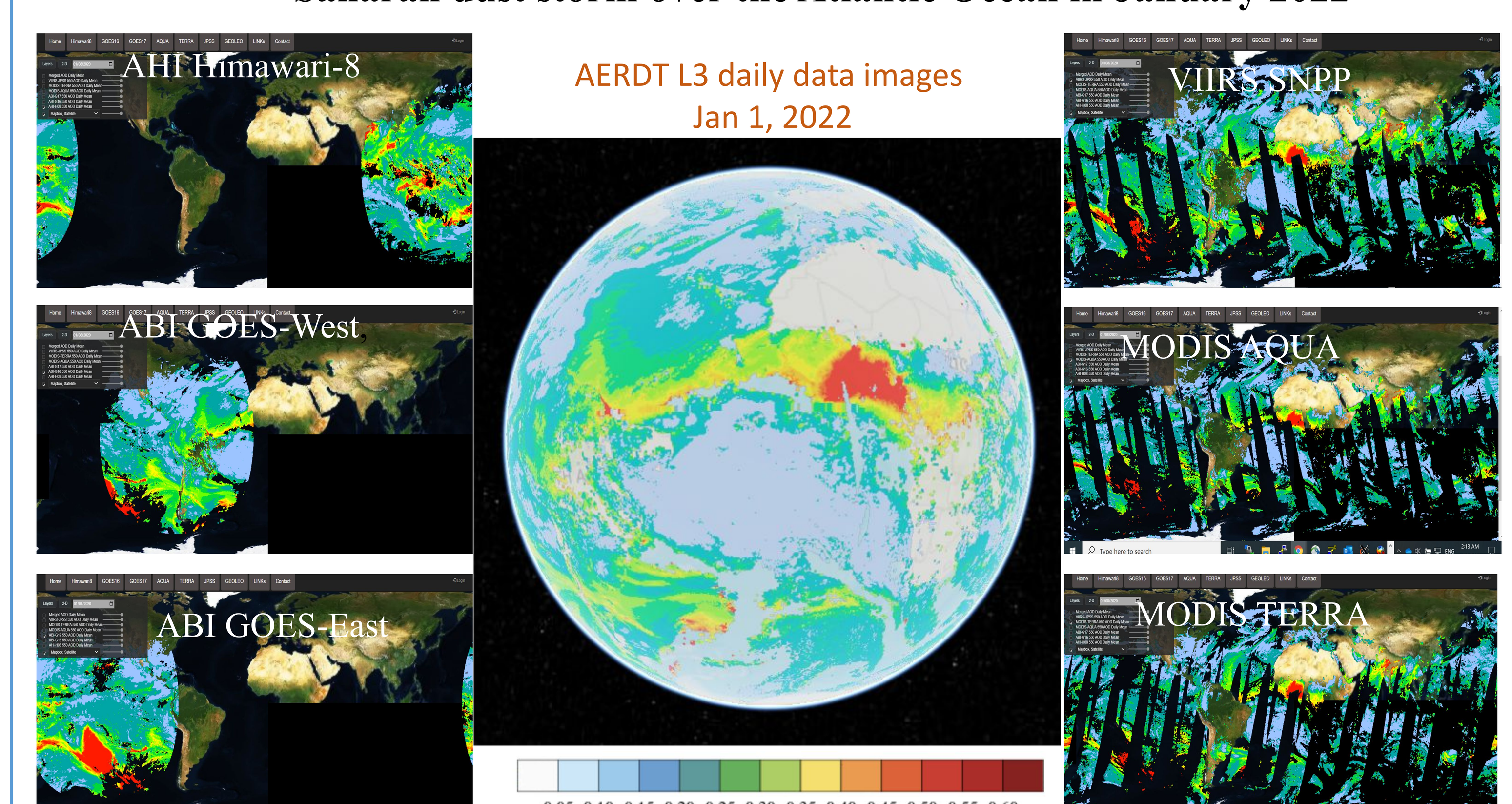
This website snapshot shows an example dark-target AOD Level-2 data (XAERDT) of three geostationary satellite sensors and three low-orbiting satellite sensors, which are the inputs for the fusion processing package.



4. Example of the Merged Data Product (Level-3)

The fused XAERDT data provide global AOD coverage at daily scale. The daily diurnal cycle of global AOD data enables analysis and monitoring of global air quality and its relationship to climate change.

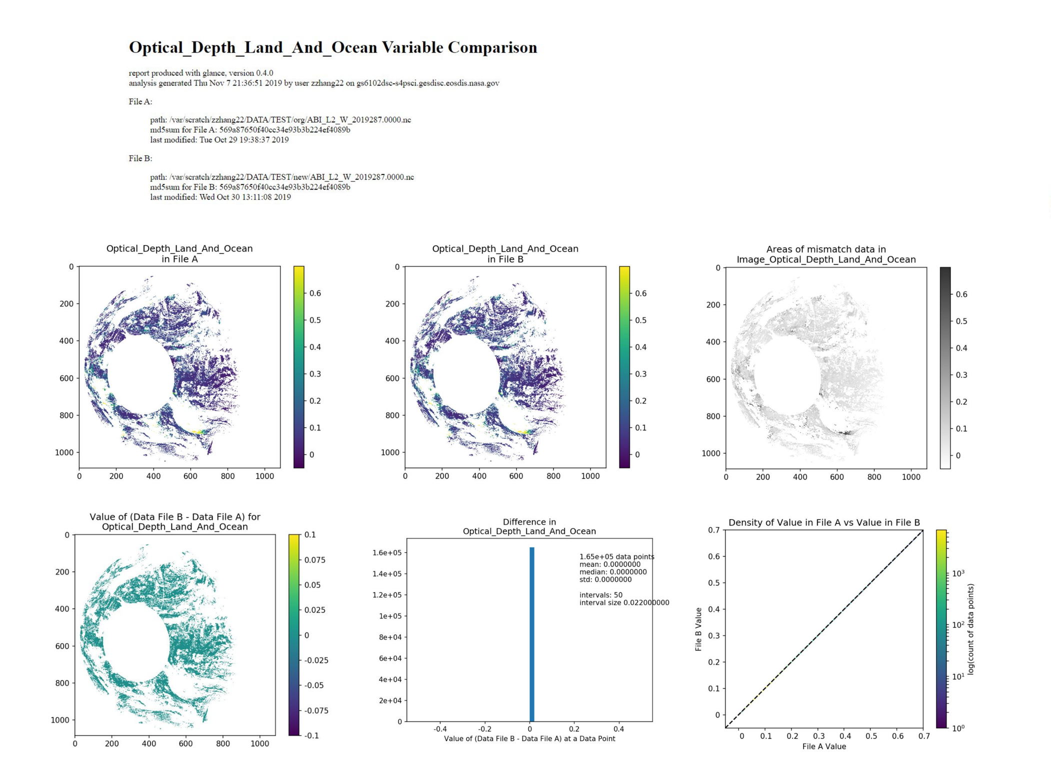
Saharan dust storm over the Atlantic Ocean in January 2022

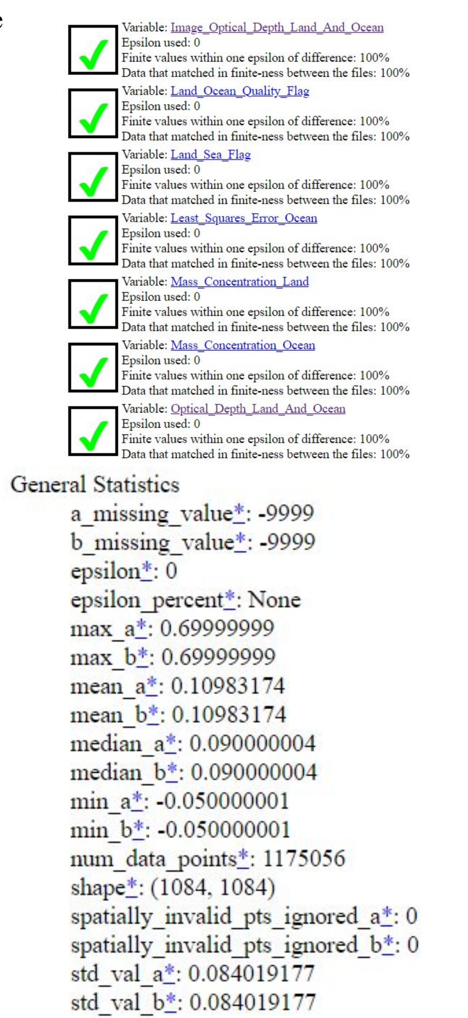


https://s4psci.gesdisc.eosdis.nasa.gov/geoleo/geoleoview/

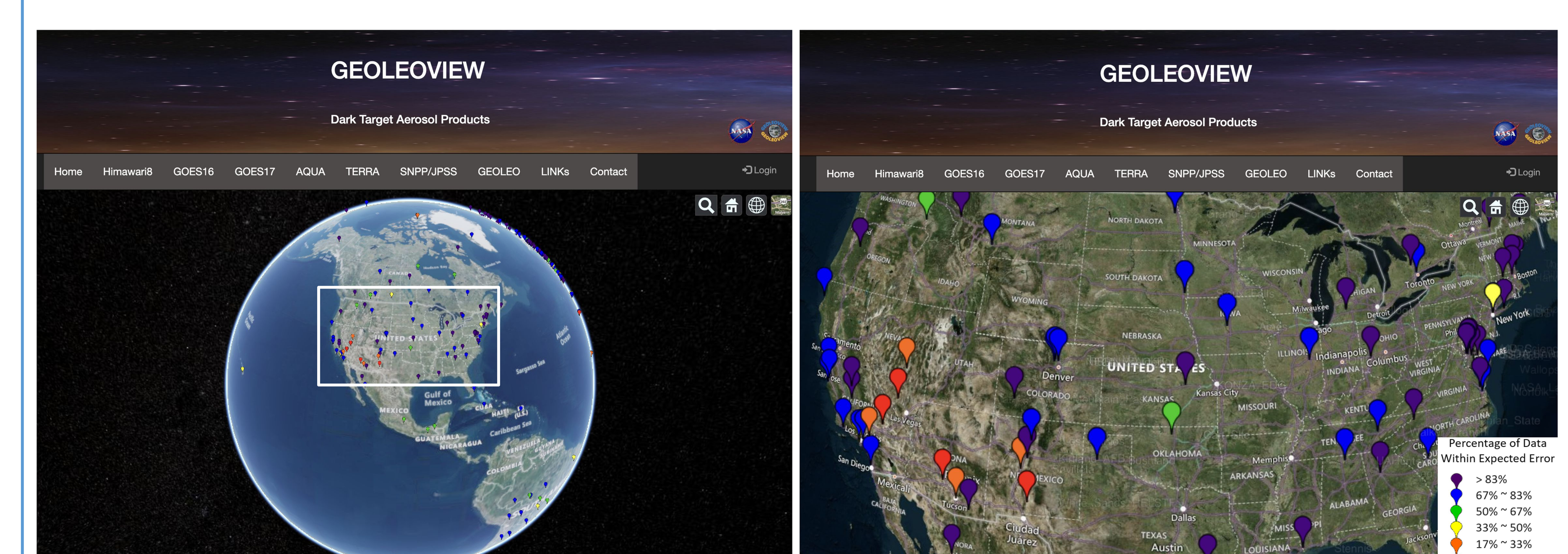
5. Web-Based Visualization Tool for Version Control and Validation

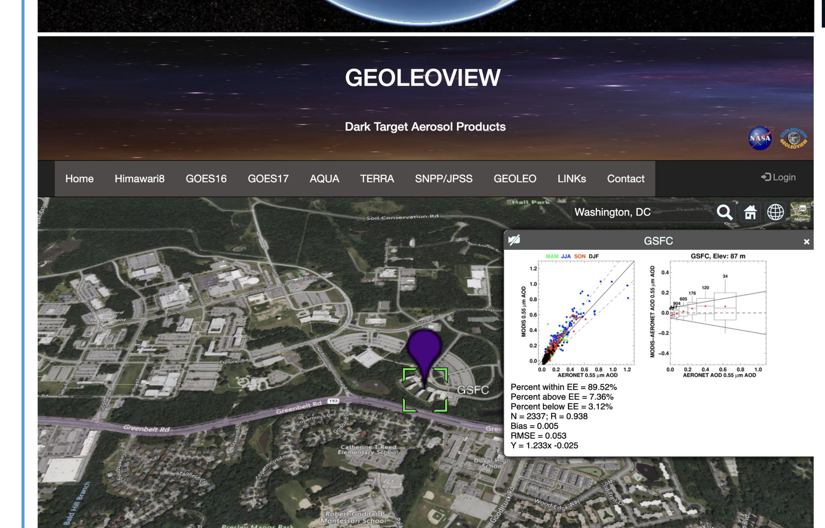
1) Version control is essential in the software life cycle. The AOD fusion package utilizes the GitHub branch technique for basic code version control and the GLANCE utility tool for result comparisons among different versions. GLANCE is an open-source tool (https://gitlab.ssec.wisc.edu/evas/UW-Glance/-/tree/master) for comparing complex datasets stored in HDF4, HDF5, and NetCDF formats. It is made up of a collection of Python routines that can be run at the command-line on Unix systems. It is implemented in our web-based tool GEOLEOVIEW (https://s4psci.gesdisc.eosdis.nasa.gov/geoleo/geoleoview/) to render the comparisons of the AOD fusion package IOs of different versions.





2) CesiumJS (https://github.com/CesiumGS/cesium) is an open-source JavaScript library for creating 3D globes and 2D maps in a web browser. Cesium is implemented in GEOLEOVIEW to facilitate the comparison of the AOD fusion package outputs with the ground-based AOD measures. Shown in this example, MODIS Aqua dark-target aerosol products for land and ocean are validated by comparing Aerosol Optical Depth (AOD) values against AOD measurements made by AERONET ground-based instruments. With the Cesium globe 3D zooming techniques, our GEOLEOVIEW tool allows users to quickly navigate over the AERONET sites and the associated validation results.





Validation maps and plots can indicate AERONET sites globally. Colored pin icons show the locations of AERONET sites, with the color indicating the performance of a specific site. A higher percentage indicates better performance of the retrieval algorithm at that particular location.

Scatter plots & box and whisker plots display the comparison of MODIS Aqua 550 nm AOD and AERONET 550 nm AOD.

6. Summary and Further Ongoing Development

- An AOD data fusion package has been developed with dark-target AOD data from three geostationary satellite sensors and three low-orbiting satellite sensors. It may be easily expanded to process data from other sources, including cloud-based data.
- The open-source nature of our AOD data fusion package ensures transparency, and community engagement. The package provides the interfaces for the efficient handling of large-size input datasets in different formats, different temporal and spatial resolutions and from different sources.
- It is user-friendly and easy to learn, with clear documentation and tutorials that can guide them through the analysis process of satellite data and remote sensing products.
- The published version is available at https://pypi.org/project/pyroscopegridding. Users are encouraged to try out this PYPI version to get familiar with the fundamental software interfaces and code structure. The version controlled repository is hosted at https://github.com/jwei-openscapes/aerosol-data-fusion

Performance Issues in Python:

- Python is a popular language for functional and convenient programming but can suffer from performance issues.
- When compared to lower-level languages like C, Python can be slower for computationally complex tasks.
- Optimization of data and calculations is crucial to minimize run time for computationally complex packages in Python.

Optimization Solutions: Multi-threading Parallelization

- GPU-based:
  - Custom CUDA kernels used for parallelization on NVIDIA GPUs.
  - Achieved 72x speedup on local hardware with GPU-based parallelization.
- CPU-based:
  - CPU-based parallelization implemented for users without GPUs.