**NASA DEVELOP National Program**



Langley Research Center

*Spring 2015*

CALIPSO: Global Health & Air Quality

*Creating a Tool with the ability to Classify Aerosols within CALIPSO Data to Help Improve Future Research and Decision Making*

 **Technical Report** 

Rough Draft - Feb 13, 2015

Jordan Vaa

Courtney Duquette

Ashna Aggarwal

Dr. Kenton Ross (National Science Advisor)

Nathan Owen (Langley DEVELOP Center Lead)

Chad Smith (Langley DEVELOP Assistant Center Lead)

# I. Abstract

[Placeholder - do not put anything here until the final draft submission. The abstract in the project summary is where the working draft of the abstract should “live”]

**Keywords**

CALIPSO, CALIOP, Remote Sensing, Lidar, Atmospheric Composition, Atmospheric Aerosols, Smoke, Clouds, Dust, Air Quality, Python

# II. Introduction

**Background Information**

On April 20, 2006 the Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observation (CALIPSO), was launched from Vandenberg Air Force base in California (Winker et al, 2007). CALIPSO was the first satellite-based lidar systems in orbit, following the Lidar In-space Technology Experiments (LITE). Before LITE , lidar observations had only been made with ground based installations and some small number of airplane based systems, lidar in space was only theoretical. Where LITE was a temporary test of the viability of using lidar in space to observe the atmosphere, CALIPSO is a semi-permanent system (Winker et al, 1996). Its purpose is to observe clouds and aerosols in the atmosphere, allowing researchers to track and understand their effects. Due to the relatively new quality of lidar data, formatting data and accessibility can be enhanced in many ways (Winker et al., 2013). CALIPSO’s instrument package is comprised principally of the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), along with a Wide Field Camera (WFC) and an Imaging Infrared Radiometer (IIR) (Winker et al, 2007). CALIOP sends pulses of 532nm and 1064nm into the atmosphere. By utilizing two receiver channels for the 532 nm wavelength and a total attenuated backscatter at the 1064 nm wavelength CALIPSO has the ability to identify several different types of aerosol, both anthropogenic (biomass burning and urban pollution) and biogenic ( desert dust, natural, sea salt, and volcanic) (Omar et al, 2009). With the data taken from the measurements, scientists are able to further chemically classify the aerosols (black carbon, organic carbons, sulfates, etc.) (Omar et al, 2009). CALIPSO determines the type based on an algorithm consisting of both backscatter measurements and volume depolarization ratio measurements (Omar et al, 2009). The algorithm uses at least five input parameters to make this determination: altitude, integrated attenuated backscatter measurements, location, surface type, and volume depolarization ratio (Omar et al, 2009). Although this algorithm does generate the aerosol subtype, it cannot discriminate between the sources of the aerosol (Omar et al, 2009). One aerosol that requires such attention is smoke. Because of its various compositions and change in structure over time, it is difficult to track smoke plumes as they move through the environment from source fires. Generally, smoke particles begin as chains of nanometer sized particles. However, “with aging these chains tend to collapse into irregular aggregates of small particles,” (Sun et al, 2013). These chains are more difficult to detect with the algorithms listed above. Our proposed tool will help researchers document smoke plume objects with a human eye and track them as they move across the environment.

Before the DEVELOP spring 2015 term, the CALIPSO science team viewed this data using an IDL program that was developed in early 2007 for the team. Although this program has undergone updates and improvements over the years, due to the proprietary nature of IDL it has remained difficult for researchers outside of NASA to use. A program, called the CloudSat and CALIPSO plotting tool (ccplot) was built in 2010 that somewhat mimics the IDL program used by the CALIPSO science team. This tool was developed in order to allow for an easy to acquire and simple to use method of visualizing CloudSat and CALIPSO data (Kuma, 2010). Ccplot allows a user to specify HDF files, specifically the CloudSAT or CALIPSO data files, and create an image based on those files. This image provides data based on the file provided, such as depolarization or aerosol types detected by the lidar, but does not allow for any manipulation on the part of the user. Additionally, it only exists as a command line program, and requires several Python libraries and modules in order to install correctly (Kuma, 2010). It is the goal of this project to update ccplot so that it is easier to install and use, as well as add the ability to manipulate the images created by it, such as selecting elements

Project Objectives

 To increase ease of access in one respect, the project’s objective is to improve upon a pre-existing IDL tool. The current tool allows a user to upload an HDF file containing L1 and L2 data products from CALIPSO, and visualize these objects as images on a grid with axes of time and altitude. The user can then select cloud and aerosol objects of interests by freehand in order to classify the data. However, the IDL platform provides no method of exporting this selection, meaning it cannot be easily shared with other researchers. The current projects goal is to create a tool with the capability of creating and exporting these selections for use in a database and ultimately a web application.

Previous Studies

In the recent years, there has been a number of algorithms developed to automatically discern cloud and aerosol layers. For example, the selective, iterated boundary location (SIBYL) algorithm scans the backscatter signals through multiple iterations and averages the signal depending on the background noise (Vaughan et al, 2009). Then, additional algorithms such as the Scene Classification Algorithm (SCA) attempt to identify what type of aerosol or cloud the SIBYL detected. The Hybrid Extinction Retrieval Algorithm (HERA) uses known lidar ratios and scattering ratios to further create backscatter profiles for clouds and aerosols (Winker et al, 2006). There is an existing tool used to analyze spectrometer imaging by combining high spectral resolution and spatial data presentation. The Spectral Image Processing System (SIPS) is an integrated system and user interface developed in IDL that allowed analysis, utilities for formatting, and visualizing data sets from AVIRIS, GERIS, and Eos HIRIS (Kruse et. al., 1993). The user interface of this tool had the ability to zoom and edit the image created based off the images in addition to saving the image in a BSQ format (Kruse et. al., 1993). In addition to the built in User Interface, SIPS could export any data to a GIS software in order to allow the user to perform further analysis (Kruse et. al., 1993).

**Objective**

The goal of this project is to replace current IDL scripts with a tool in the open source Python programming environment to give it a much wider distribution base. Once the tool is created it will be able to allow selected objects to be exported into a web-based database, allowing for greater ease of access for many users. Additionally this will allow for a consolidation of all prerequisite code and tools, which proved challenging to collect and run due to its disparate nature. An additional goal of this project is to simplify the selection of objects within CALIOP data, specifically smoke plumes, which have significant impact on the air quality on a region. Smoke is difficult to track over time because of its various possible compositions and visual similarity to clouds. With an organized, accessible database, researchers will be able to better recognize the impact of smoke as it travels further from source fires.

**Area of Study**

This project has a global area of study, as the tool can be used with any CALIPSO data.

**Period of Study**

2006-Present

**National Application Addressed**

This project addresses the health and air quality application area. It will allow researchers to better document the lifecycles and effects of various atmospheric aerosols throughout the globe which affect the overall quality of the air.

**Project Partners**

Our partners include the NASA CALIPSO science team, with the point of contacts Dr. Charles Trepte and Dr. Amber Soja. Dr. Trepte is hoping for a centralized location for organized CALIOP data. However, Dr. Amber Soja would like to see the database assist the effort of detecting smoke plumes as they travel through the environment.

# III. Methodology

**Data Used**

CALIPSO provides 6 L1 and 10 L2 data products. These data projects are provided in the Hierarchical Data Format (HDF) which is useful for storing large scientific data samples, but difficult to access. This is compounded by the fact that CALIPSO data is meant to viewed in a graphical format, showing the “curtain view” of the lidar reflections. The CALIPSO science team uses an IDL command-line program known as PDFreader which allows them to view the data and select objects within it for further consideration as well as active manipulation. However, IDL is a proprietary language making it difficult for users to get access outside of NASA. To this end the current project is written in Python, which is a free to use language, in contrast to proprietary languages such as IDL and MatLab. It was released under a two part BSD license, meaning it may be used and integrated into other products, so long as the original license is retained (Kuma, 2010). Over the five years, Python has been used more frequently in the atmospheric science due to the added features of array handling and other data structures (Lin, 2012). Python also has a clear and natural syntax that is easy to read by resembling pseudo-code. Finally since Python is an open-source language, the users have to ability to use or create their own customized packages (Lin, 2012). Since Python can utilize several different packages with a single interpreted environment, the code is more concise and flexible (Lin, 2012).

**Project Process**

The DEVELOP Team has endeavored to expand the ease of use and functionality of ccplot. In its native form, ccplot requires the acquisition and installation of numerous disparate python libraries, since it was optimized for use in a UNIX (meaning Linux or MACos) environment. However, most users are coming from a Windows user system. By consolidating ccplot with each of the required libraries included, users will be able to simply download and install ccplot, rather than acquiring each library separately. Acquiring a library separately can cause issues with versioning and incorrect library usage, leading ccplot to fail. In addition, the team has developed a graphical user interface (GUI) to make ccplot more usable. Natively, ccplot uses a command line interface, which requires an understanding of command line usage as well as directory navigation. A GUI both gives access to a much wider user base and speeds up the process of inputting data and outputting results. Eventually the DEVELOP team would like to introduce tools that will allow researchers to manipulate the output images, allowing them to select aerosol elements of interest for further consideration. Once an element is selected, it can be flagged, categorized, and exported into a MySQL database so other scientists can view various objects organized by similar characteristics. The ultimate goal of this project is to build an internet based application that will give users access to this database of objects created by various researchers, allowing them to collaborate and share information with ease, giving a much wider audience the ability to benefit from the CALIPSO data and the information it can provide about aerosol movement in the atmosphere.

# IV. Results & Discussion

Insert images, graphs, maps, charts, etc. here. Choose the most important results to highlight here. Things to discuss:

* Analysis of Results: What can you tell from your graphs, images, etc? What does this mean for your project?
* Errors & Uncertainty: What factors could you not account for, what things didn’t work out like you expected they would, etc
* Future Work: If this project was to be selected for another term, what would be the focus? What other areas would be of interest?

# V. Conclusions

Final conclusions. Word count: 200-600.

# VI. Acknowledgments

Insert here. Keep to a concise paragraph or bullets of names. End with the following sentence.

This material is based upon work supported by NASA through contract NNL11AA00B and cooperative agreement NNX14AB60A.

# VII. References

Kruse, F., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P., & Goetz, A. (1993). The spectral image processing system (sips): Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2-3). 145-163. doi: [10.1016/0034-4257(93)90013-N](http://dx.doi.org/10.1016/0034-4257%2893%2990013-N)

Kuma, P. (2010). Visualising Data from CloudSat and CALIPSO Satellites. Bachelors Thesis. *Bratislava Faculty of Mathematics, Physics, and Informatics, Comenius University*.

Lin, W. (2012): Why Python Is the Next Wave in Earth Sciences Computing. *Bull. Amer. Meteor. Soc.*, 93, 1823–1824. doi: http://dx.doi.org/10.1175/BAMS-D-12-00148.1

Omar, A., Winker, D., Vaughan, M., Hu, Y., Trepte, C., Ferrare, R., Lee, K., Hostetler, C., Kittaka, C., Rogers, R., Kuehn, R., & Liu, Z, (2009) The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm. *J. Atmos. Oceanic Technol.*,*26*(10). doi: <http://dx.doi.org/10.1175/2009JTECHA1231.1>

Setvák, M., Bedka, K., Lindsey, D., Sokol, A., Charvát, Z., Šťástka, J., & Wang, P. (2013). A-Train observations of deep convective storm tops. *Atmospheric Research,* *123*, 229-248. doi: [10.1016/j.atmosres.2012.06.020](http://dx.doi.org/10.1016/j.atmosres.2012.06.020)

Sun, W., Liu, Z., Videen, G., Fu, Q., Muinonen, K., Winker, D., Lukashin, C., Jin, Z., Lin, B., Huang, J., (201) For the depolarization of linearly polarized light by smoke particles. *Journal of Quantitative Spectroscopy and Radiative Transfer.* doi: [10.1016/j.jqsrt.2012.03.031](http://dx.doi.org/10.1016/j.jqsrt.2012.03.031)

Vaughan, M., Powell, K., Winker, D., Hostetler, C., Kuehn, R., Hunt, W.... Mcgill, M. (2009). Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements. *Journal of Atmospheric and Oceanic Technology, 26*(10), 2034-2050. doi: <http://dx.doi.org/10.1175/2009JTECHA1228.1>

Winker, D., Couch, R., & Mccormick, M. (1996). An overview of LITE: NASA's Lidar In-space Technology Experiment.*Proceedings of the IEEE,* *84*(2), 164-180. doi: [10.1109/5.482227](http://dx.doi.org/10.1109/5.482227)

Winker, D., Hostetler, C., Vaughan, M., Omar, A., (2006) CALIOP Algorithm Theoretical Basis Document. *PC-SCI-202 Part 1 Release 2.0*

Winker, D., Hunt, W., & Mcgill, M. (2007). Initial performance assessment of CALIOP.*Geophysical Research Letters,* *34*(19). doi: 10.1029/2007GL030135

Winker, D, Pelon, J, Trepte, C, & Barroso, C. (2013) CALIPSO: Proposal to the 2013 Earth Science Division Senior Review Panel.

# VIII. Appendices

Insert here