**NASA DEVELOP National Program**



NASA Langley Research Center

*Spring 2016*

LaRC CALIPSO Cross-Cutting III

Interacting with CALIPSO Data through a

Graphical User Interface

 **Technical Report**

Rough Draft – Feb 18, 2016

Kathleen Moore (Project Lead)

Jordan Vaa

Kenton Ross, Ph.D., NASA Langley Research Center (Science Advisor)

Grant Mercer, University of Nevada, Las Vegas (Science Advisor)

Michael Bender, NASA Langley Research Center (Science Advisor)

Previous Contributors:

Grant Mercer

Nathan Qian

Ashna Aggarwal

Courtney Duquette

Jordan Vaa

# 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**

Atmospheric Aerosols, CALIPSO, CALIOP, LiDAR

# II. Introduction

The Earth’s atmosphere is a complex and dynamic system, with many factors contributing to its overall structure. Some of the major factors in this system are clouds and aerosols, which consist of various particles that travel through the various layers of the atmosphere. Understanding how clouds and aerosols form, travel, and extinguish in the atmosphere helps us to understand the composition of the atmosphere with much greater clarity. Aerosols are created by a variety of sources both anthropogenic, human made, and natural, occurring from natural interactions. Because they are in the atmosphere, as aerosols travel, they can affect large geographic areas as opposed to other environmental effects, and thus understanding their life cycle through the atmosphere is important in understanding how various environmental entities can interact.

On April 20, 2006, the Cloud-Aerosol LiDAR Infrared Pathfinder Satellite Observation (CALIPSO) was launched as a follow-up to the shuttle based LiDAR In-Space Technology Experiment (LiTE) to serve as a more long-term platform for continuing this work. Its purpose was to provide measurements of clouds and aerosols and their interactions and roles in the climate system (Winker et al, 2010). CALIPSO carries three instruments as its payload. The primary instrument is the Cloud-Aerosol LiDAR with Orthogonal Polarization (CALIOP), a new-nadir viewing two-wavelength polarization-sensitive LiDAR sensor (Winker et al, 2009). The remaining two are passive sensors: a wide field camera and an infrared imaging radiometer (Winker et al, 2009). CALIPSO can determine an aerosol’s type by measuring backscatter and volume depolarization ratio measurements (Omar et al. 2009). The algorithm takes altitude, integrated attenuated backscatter, location, surface type, and volume depolarization ratio as parameters (Omar et al, 2009). With CALIPSO, researchers have access to more aerosol data and can better track aerosols through the atmosphere as well as determine crucial information about where aerosols travel.

A major challenge facing researchers looking to use CALIPSO data is the data presentation style, which is different from many other commonly used datasets. CALIPSO is collecting vertical profile data rather than wide swath data meaning many common tools for visualizing data have difficulties working with CALIPSO data. To overcome this obstacle, a proprietary visualization tool was developed specifically to view and manipulate CALIPSO data. However, this tool was written in IDL, which is an uncommon and proprietary language, which makes it difficult for new developers to update and maintain. In the spring of 2015, a DEVELOP group created an updated visualization tool in python to alleviate this difficulty. They integrated a Python-based tool known as the CALIPSO and CloudSat Plotter (CCplot) that takes CALIPSO data as input and outputs two-dimensional plots. This tool had no user interface and limited ability to manipulate the data once plotted.  In two terms of DEVELOP, this tool was updated to reflect the needs of researchers, becoming the Visualization Of CALIPSO (VOCAL) tool. Continuing to update this tool to better fit the needs of the researchers, we have enabled the tool to be easily installed in Apple’s Mac OS X, and we have integrated web-based data download. This will allow researchers greater flexibility to work with the tool as well as alleviate the need to download and store the large data sets.

Including the items listed below; write a synopsis of the following information. Be concise. Word count should be between 200-1000 as one to two pages should suffice.

Material to include:

* Background Information: Relevant information to inform the reader of current status, issues, previous studies, etc.
* Project Objectives: These should be short decisive action items.
* Study Area: Describe the geographic location of the study
* Study Period: Explain the time period of data you are looking at (years and dates of data)
* National Application(s) Addressed: Explain which NASA national application areas this project addresses and how it contributes to them
* Project Partners: Explain who the project partners are, why they are interested in this project, how they will use it, what decision making they have to do and is being addressed with this research and methodologies, etc. How will they benefit from this project and methodology?
* Katie notes:
  + Can we reuse (not word for word) Previous background info? ECA: Yes.
  + Grant made the above bulleted list into explicit subsections. Bad or good? ECA: Whichever we’d like. I’m going to **not** break it up.

# III. Methodology

VOCAL is a graphical user interface (GUI) that enables a user to load a Hierarchical Data Format (HDF) file, which is the output of the CALIPSO satellite.  The user can then visualize the data and interact with it by tagging target regions in the image in a “lasso”-style.  Because the previous DEVELOP teams emphasized the object-oriented, organized, and expandable nature of the code, much of the infrastructure and functionality of the current VOCAL version have been rolled into the current term.  Upon this code base, our team has improved the system in two major spaces: cross-platform functionality and efficiency/scalability.  Here, we give a brief overview of how VOCAL works, followed by our specific additions to the project and their importance to the study of Earth observations as a whole.

**A. System Overview**

At the core of the VOCAL system is the interaction between three general components: shapes (and their manager), the database (and its manager), and the GUI.  A user that wants to view CALIPSO data launches the VOCAL GUI and either loads in an HDF file or previously created shapes from a JavaScript Object Notation (JSON).  These shapes are associated with a particular HDF file so that shapes cannot become “orphaned” data.

Whether or not the user loads previous shapes, the user can then draw further on the image, creating a polygon, incrementally with each click.  On completion of the drawing, VOCAL associates attributes with the shape, including the polygon’s coordinates and the color assigned to that shape for identification purposes.  The user can draw an unlimited number of shapes, but if these crowd the image space, the user can hide selected shapes from view while continuing to work in the same view.  However, because VOCAL also accommodates zooming in/out and panning, the user has multiple options for focusing on a desired region of the image.  Lastly, shapes can be moved and erased.  They are purely a means for denoting regions of interest on the image and enabling this data to persist after the current VOCAL session.

To this end, VOCAL has an integrated database for storing and loading shapes.  Utilizing the SQLAlchemy database manager, VOCAL saves previous shapes, each in a JSON file, and this file is also associated with the HDF file from which it was originally taken.  Because attributes such as the local image coordinates are saved, the user can immediately return to work on this particular shape or share it with another user of VOCAL.  Herein lies the one of the best features of VOCAL: the enabling of collaboration.

The current features of VOCAL are in place largely to serve the purpose of unambiguous discussion of atmospheric aerosol data.  Earth observations such as those produced by the CALIPSO satellite are pinned down in location on a global scale, but this vantage potentially loses accuracy and definitely loses intuitiveness when one starts to work on a subset of a cross-section of the atmosphere.  Now, multiple scientists can be remotely located and be certain that they are discussing the same region, taken from the same file.  Our current task, then, became a matter of how to increase the collaborative capabilities of the software and its data.

**B. Cross-platform Compatibility**

Before the current term, we had an involved, multi-step process for installing VOCAL on its primary operating system, Windows 7.  Limited by the CCPlot library, the installation on Windows is 32-bit.  To increase the chances that a more novice computer-user could install and use the software, we created a VOCAL installation executable (EXE) file that would streamline the installation process.

The primary users of the VOCAL system are expected to be the CALIPSO science team, many of whom work on Apple computers running Mac OS X, so we worked to integrate Mac functionality into the VOCAL installation. Similarly to the Windows installation, the major obstacle was CCPlot and its associated dependencies. As the methods and means of installation on a Mac differ from those on a Windows system, the process had to be fine-tuned and tested for operability.  [Also, may have to talk about converting to 64-bit code here.]

**C. OPeNDAP Integration**

One of the bottlenecks in using VOCAL is the need for locally-stored HDF files, which can oftentimes be very large. This both hinders storage capacity by taking up a large amount of memory, as well as creates length download speeds for gathering data. A collaboration with the Atmospheric Science Data Center (ASDC), the organization responsible for collecting and storing CALIPSO data, has led to the discovery of the Open-source Project for a Network Data Access Protocol (OPeNDAP). OPeNDAP is a protocol for accessing scientific data from over the Internet. OPeNDAP would allow researchers using VOCAL to select only the part of the dataset in which they are interested directly from the VOCAL interface, rather than having to download the data and ahead of time, and then load it into the program in bulk. This would reduce the storage and download requirements on the researcher's part, as well as facilitate easier sharing of findings and data between researchers.

This should be the focus of the paper - concise, yet explanatory, and highlight the NASA Earth observations utilized and its/their capabilities. Include a paragraph or more for each of the following items. No word cap, but be thoughtful and keep it in the two to six page range.

Content to include:

* Data Acquisition: What data did you get, what level products are they, for what dates did you get images, where did you get the images from, etc.
* Data Processing: What did you do to the data? Were there conversions needed to be able to analyze it? Did you have to mosaic images? Did you have to normalize anything to fit other datasets? Did you run an NDVI, change detection, etc?
* Data Analysis: How did you analyze the data? What methods did you use?
* Katie notes:
  + Data Acquisition: CALIPSO curtain images from May 2006 to present (theoretically), available through the ASDC?
  + Data Processing: VOCAL displays the data in its raw format, enables zooming in/out (does it yet??? Doesn’t work for my installation.)
  + Data Analysis: VOCAL enables the tagging of shapes with attributes and storage to a database.
  + Grant largely just goes through all of the program’s classes and describes what each one does and how it relates to the others. I’m compacting that section and then adding our content.

# IV. Results & Discussion

This is currently the third generation of this project. Since the first term, the CALIPSO Cross-Cutting teams have built and continued to add functionality to the visualizer. It continues to become an increasingly useful program through which scientists can interact with image data.

Insert images, graphs, maps, charts, etc. here. Choose the most important results to highlight here. No word cap, but two to six pages is a good range.

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?
* Katie:
  + Grant uses this section to briefly review what a user can do with the tool, and he shows a few screenshots of the tool in action. He also includes a screenshot of the github logger, which we may not include. Instead, we could show a shot of the installation screen? Since that would be a new contribution? Potentially boring though.
  + The next major portion that Grant includes is a section on future work. For us, this could include more of the database hosting possibilities, and whether this tool could be converted to a web-based application.
  + Talk about difficulties.

# V. Conclusions

Final conclusions. Word count: 200-600 (~a page).

# 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

Omar, Ali H., David M. Winker, Mark A. Vaughan, Yongxiang Hu, Charles R. Trepte, Richard A. Ferrare, Kam-Pui Lee, Chris A. Hostetler, Chieko Kittaka, Raymond R. Rogers, Ralph E. Kuehn, and Zhaoyan Liu. "The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm." *Journal of Atmospheric and Oceanic Technology J. Atmos. Oceanic Technol.* 26.10 (2009): 1994-2014. *Web of Science*. Web. 17 June 2015.

Winker, David M., Mark A. Vaughan, Ali Omar, Yongxiang Hu, Kathleen A. Powell, Zhaoyan Liu, William H. Hunt, and Stuart A. Young. "Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms.” *Journal of Atmospheric and Oceanic Technology J. Atmos. Oceanic Technol.* 26.11 (2009): 2310-323. *Web of Science*. Web. 17 June 2015.

Insert references here. Only include articles/content cited in the body of text above. It’s great if you read many other articles, but they should not all be listed here unless they are being cited in this report.

Use whatever style you want - here are some options:

<http://www.dovepress.com/author_guidelines.php?folder_id=208>

<http://en.wikipedia.org/wiki/Citation>

<http://www.agu.org/pubs/pdf/AuthorRefSheet.pdf>

<http://linguistics.byu.edu/faculty/henrichsenl/apa/apa01.html>

# VIII. Content Innovation

In preparation for DEVELOP’s coming microjournal, please select two content innovation features to support your paper. For each item, please list the name of the feature, and include the tool itself if possible (eg. glossary terms and definitions). If the tool does not work in Microsoft Word (eg. Interactive MATLAB Figure Viewer), please list the file name and upload the related file to the microjournal folder on the DEVELOP Exchange. If you choose to use Inline Supplementary Material, please also include where the material should appear in the text.

**Some options include:**

AudioSlides

Database Linking Tool

Data Profile

Executable Papers

Featured Author Videos

Featured Multimedia for this Article (video and podcast options)

Glossary Viewer

Inline Supplementary Material (figures, tables, computer code)

Interactive Map Viewer

Interactive MATLAB Figure Viewer

Interactive Plot Viewer

Nomenclature Viewer

# IV. Appendices

Insert here