Creating user-friendly tools for data analysis and visualization in K-12 classrooms: A Fortran dinosaur meets Generation Y

L. H. Chambers[1], S. Chaudhury[2], M. T. Page[2], A. J. Lankey[3], J. Doughty[4], Steven Kern[5], Tina M. Rogerson[6]
[1]NASA Langley Research Center, Hampton, VA 23681-2199
[2]Christopher Newport University, Newport News, VA
[3]Purdue University, West Lafayette, IN
[4]Virginia Polytechnic & State University, Blacksburg, VA
[5]Ocean Lakes High School, Virginia Beach, VA
[6]Science Systems & Applications Inc., Hampton, VA

TOPICS: Tools to Facilitate Data Analysis in the K-12 Classroom

During the summer of 2007, as part of the second year of a NASA-funded project in partnership with Christopher Newport University (Chaudhury) called SPHERE (Students as Professionals Helping Educators Research the Earth), a group of undergraduate students spent 8 weeks in a research internship at or near NASA Langley Research Center. Three students from this group (Page, Lankey, and Doughty) formed the Clouds group along with a NASA mentor (Chambers), and the brief addition of a local high school student (Kern) fulfilling a mentorship requirement.

The Clouds group was given the task of exploring and analyzing ground-based cloud observations obtained by K-12 students as part of the Students' Cloud Observations On-Line (S'COOL) Project, and the corresponding satellite data. This project began in 1997. The primary analysis tools developed for it were in FORTRAN, a computer language none of the students were familiar with. While they persevered through computer challenges and picky syntax, it eventually became obvious that this was not the most fruitful approach for a project aimed at motivating K-12 students to do their own data analysis. Thus, about halfway through the summer the group shifted its focus to more modern data analysis and visualization tools, namely spreadsheets and Google™ Earth.

The result of their efforts, so far, is two different Excel spreadsheets and a Google™ Earth file. The spreadsheets are set up to allow participating classrooms to paste in a particular dataset of interest, using the standard S'COOL format, and easily perform a variety of analyses and comparisons of the ground cloud observation reports and their correspondence with the satellite data. This includes summarizing cloud occurrence and cloud cover statistics, and comparing cloud cover measurements from the two points of view. A visual classification tool is also provided to compare the cloud levels reported from the two viewpoints. This provides a statistical counterpart to the existing S'COOL data visualization tool, which is used for individual ground-to-satellite correspondences. The Google™ Earth file contains a set of placemarks and ground overlays to show participating students the area around their school that the satellite is measuring. This approach will be automated and made interactive by the S'COOL database expert (Rogerson) and will also be used to help refine the latitude/longitude location of the participating schools.

Once complete, these new data analysis tools will be posted on the S’COOL website for use by the project participants in schools around the US and the world.

P1.25 CREATING USER-FRIENDLY TOOLS FOR DATA ANALYSIS AND VISUALIZATION IN K-12 CLASSROOMS: A FORTRAN DINOSAUR MEETS GENERATION Y

L H. Chambers[1], S. Chaudhury[2], M. T. Page[2], A. J. Lankey[3], J. Doughty[4], S. Kern[5], and T. M. Rogerson[6]
[1]NASA Langley Research Center, Hampton, VA
[2]Christopher Newport University, Newport News, VA
[3]Purdue University, West Lafayette, IN
[4]Virginia Polytechnic & State University, Blacksburg, VA
[5]Ocean Lakes High School, Virginia Beach, VA
[6]Science Systems and Applications, Inc., Hampton, VA

1. INTRODUCTION

During the summer of 2007, as part of the second year of a NASA-funded education project in partnership with Christopher Newport University (Chaudhury) called SPHERE (Students as Professionals Helping Educators Research the Earth), a group of undergraduate students spent 8 weeks in a research internship at or near NASA Langley Research Center in southeastern Virginia. Three students from this group (Page, Lankey, and Doughty) formed the Clouds group along with a NASA mentor (Chambers), and the brief addition of a local high school student (Kern) fulfilling a mentorship requirement.

The Clouds group was given the task of exploring and analyzing ground-based cloud observations obtained by K-12 students as part of the Students' Cloud Observations On-Line (S'COOL; http://scool.larc.nasa.gov) Project, along with the corresponding satellite data. The S'COOL project began in 1997 (Chambers et al, 2003). The primary analysis tools developed for it were in FORTRAN, a computer language commonly used by scientists (especially older ones), but that none of the students were familiar with. While they persevered through computer challenges and picky syntax, it eventually became obvious that this was not the most fruitful approach for a project aimed at motivating K-12 students to do their own data analysis. Thus, about halfway through the summer the group shifted its focus to more modern data analysis and visualization tools, namely spreadsheets and Google™ Earth.

The final result of the Cloud group's efforts is two different Excel spreadsheets and a Google™ Earth file. These are discussed in more detail in sections 3 and 4 of this paper.

* *Corresponding author address:* Lin H. Chambers, MS 420, NASA Langley Research Center, Hampton, VA 23681-2199; e-mail: Lin.H.Chambers@nasa.gov

2. OVERVIEW OF S'COOL

The S'COOL Project is the Education and Public Outreach component of the Clouds and the Earth's Radiant Energy System (CERES; Wielicki et al., 1996), part of NASA's Earth Observing System. CERES instruments make highly accurate measurements of the energy flows in and out of the planet, on several low Earth orbit satellites:

- TRMM, the Tropical Rainfall Measuring Mission, carries one CERES instrument, which operated from late 1997 to March 2000.
- Terra carries two CERES instruments, which have been making measurements since early 2000.
- Aqua also carries two CERES instruments, which have been making measurements since mid-2004.

A key objective of the CERES project is to quantify the impact of clouds on Earth's energy budget. Thus, each CERES instrument flies with an imager (VIRS on TRMM, MODIS on Terra and Aqua), and cloud retrievals are performed as part of the CERES analysis. The S'COOL observations provide ground truth cloud reports from surface observers in locations around the world. Schools are requested to make observations within +/- 15 minutes of the CERES overpass, and report these observations to an on-line database.

As the satellite data are processed, first through the rapid FLASHFlux (Stackhouse et al., 2006) processing, then eventually through the climate-quality CERES processing, the satellite data corresponding to each S'COOL observation report are also made available through an on-line database. This enables both students and scientists to explore and study the level of agreement between the two viewpoints.

3. MODERN TOOL I: SPREADSHEETS

A relatively simple spreadsheet file was already available on the S'COOL website. This file used a small subset of S'COOL observations (3 years of data from Virginia in April) and showed three examples of data analysis that could be performed:

1) Analysis of cloud amount frequency (clear, partly cloudy, mostly cloudy, overcast) for different cloud levels (low, mid, high). See Figure 1.

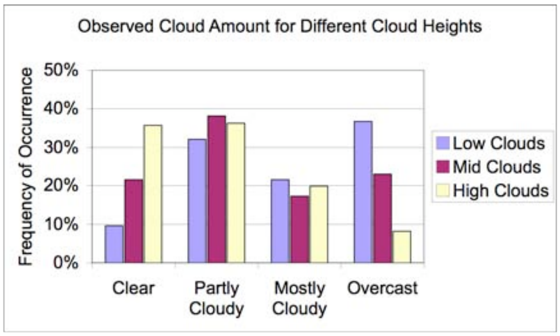

Figure 1. Sample analysis of cloud amount vs cloud height using Excel.

2) Counting the number of cloud layers observed from the ground and creating a pie chart. See Figure 2.

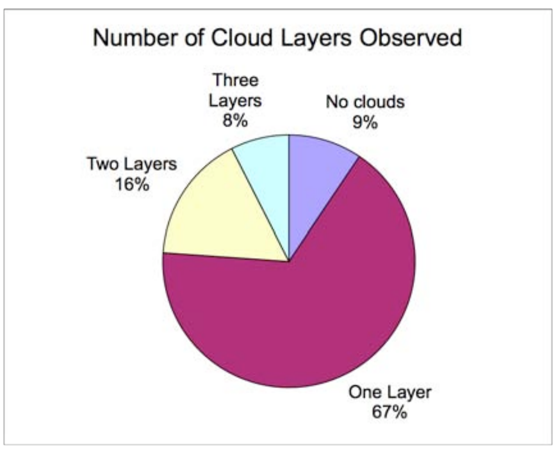

Figure 2. Pie chart showing the number of layers reported by S'COOL observers

3) Comparing the Cloud amount reported from the ground with the cloud cover obtained from satellite. Because the ground observers report cloud cover in discrete classes while the satellite can measure any value, the resulting graph is quite messy.

First, the Clouds group updated this existing spreadsheet by generalizing it to accept up to 1,000 data rows from the S'COOL database (this is the maximum number of rows that can be downloaded at one time from the on-line interface). S'COOL participants can select these

		Ground report				
		No cloud	Clear	Partly Cloudy	Mostly Cloudy	Overcast
Sat report	No Cloud					
	Clear		16	7	4	0
	Partly		33	52	47	26
	Mostly		14	22	47	74
	Overcast		4	4	38	98
Total Cases with Cloud		486		Percentages		
Complete agreement		213		43.83		
Off by 1 Cloud Cover Class		221		45.47		
Off by 2 Classes		48		9.88		
Off by 3 Classes		4		0.82		

Figure 3. Matrix comparing ground vs satellite cloud cover. Note that the matrix is incomplete, allowing interested students to practice using Excel formulae.

data according to their interests (location, time period, etc) from the on-line S'COOL archive, and paste them into a prepared tab in the spreadsheet. This automatically performs the three analyses above; plus it generates a table showing a comparison of cloud cover as reported from the ground versus observed by the satellite. See Figure 3.

Secondly, the Clouds group generated a new spreadsheet file to create a visual classification of the cloud levels reported from the two viewpoints. This provides a statistical counterpart to the existing S'COOL data visualization tool (Figure 4), which is used for individual ground-to-satellite correspondences. This classification (Fig. 5) is coded to show cases with an exact match in the cloud mask (blue), an easily explainable discrepancy between the two points of view (green), a plausible discrepancy (orange), and an as yet unexplained disagreement (red). Beside each cloud mask pair the spreadsheet calculates how many of the up to 1,000 rows of the dataset being analyzed fall into that classification. Underneath (not shown) is a sorted list showing

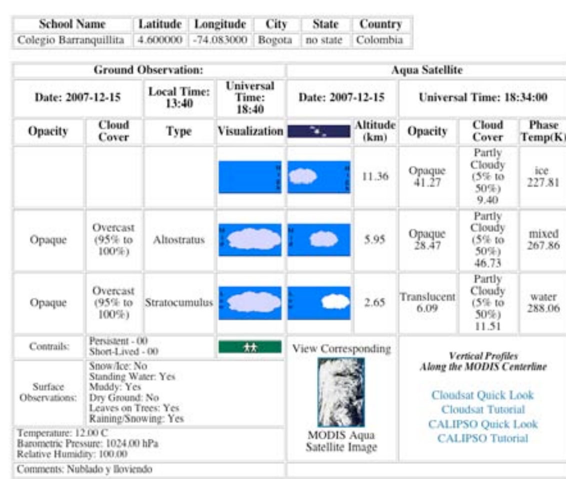

Figure 4. S'COOL on-line data visualization tool.

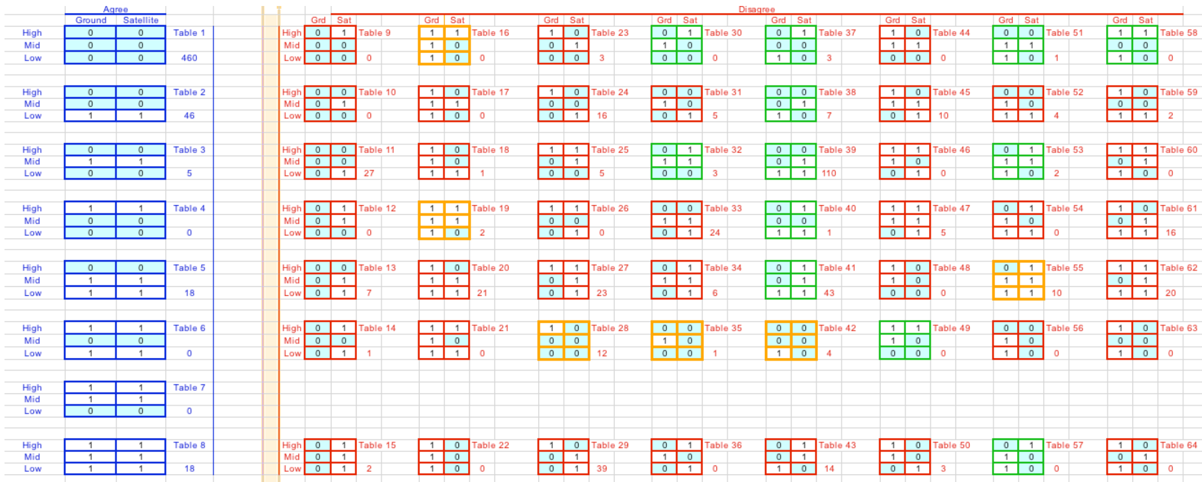

Figure 5. Visual classification of ground to satellite cloud mask pairs

which cloud mask pairs are most to least common. This analysis builds on a project performed the previous summer by another SPHERE student, to begin looking at the ground and satellite cloud mask by cloud level.

This new tool has many uses, including the ability to pinpoint cases of disagreement for further examination and consideration.

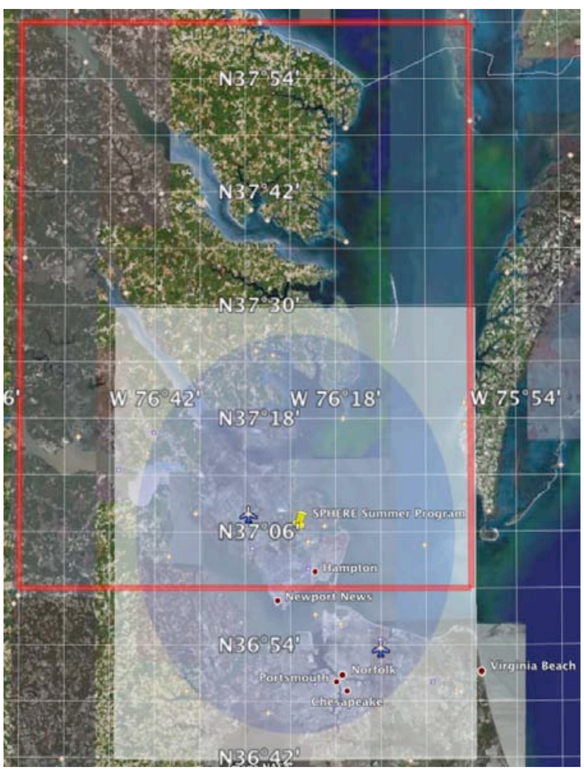

Figure 6. Google™ Earth visualization for the location of the SPHERE summer program, including the two satellite region overlays.

4. MODERN TOOL II: Google™ Earth

Google™ Earth has emerged in recent years as a powerful tool for visualizing locations on the surface of the Earth. The Clouds group was challenged to develop a kml file that would accomplish several things for the S'COOL Project:

1) Identify the specific location of a participating school, based on the latitude/longitude information they provided when registering for the project.

2) Overlay a rectangle around the S'COOL whose corners are located at the nearest integer latitude/longitude. This outlines the region for which satellite cloud data are used for the standard S'COOL comparison.

3) Overlay a circle with a radius of 0.4 degrees of latitude/longitude centered upon the school. This outlines the region for which satellite cloud data are used for the newer, "footprint" S'COOL comparison.

The Clouds group successfully developed a set of placemarks and ground overlays to accomplish these goals (Fig. 6). Their approach has been automated and made interactive by the S'COOL database expert (Rogerson). The file also includes the school name, the number of observations submitted, and their preferred language. It is now available to participating schools via the S'COOL website and allows them to refine the latitude/longitude coordinates of their school by looking for landmarks in the Google™ Earth imagery. Two other options are also available:

1) View all registered S'COOL participants (placemark only).

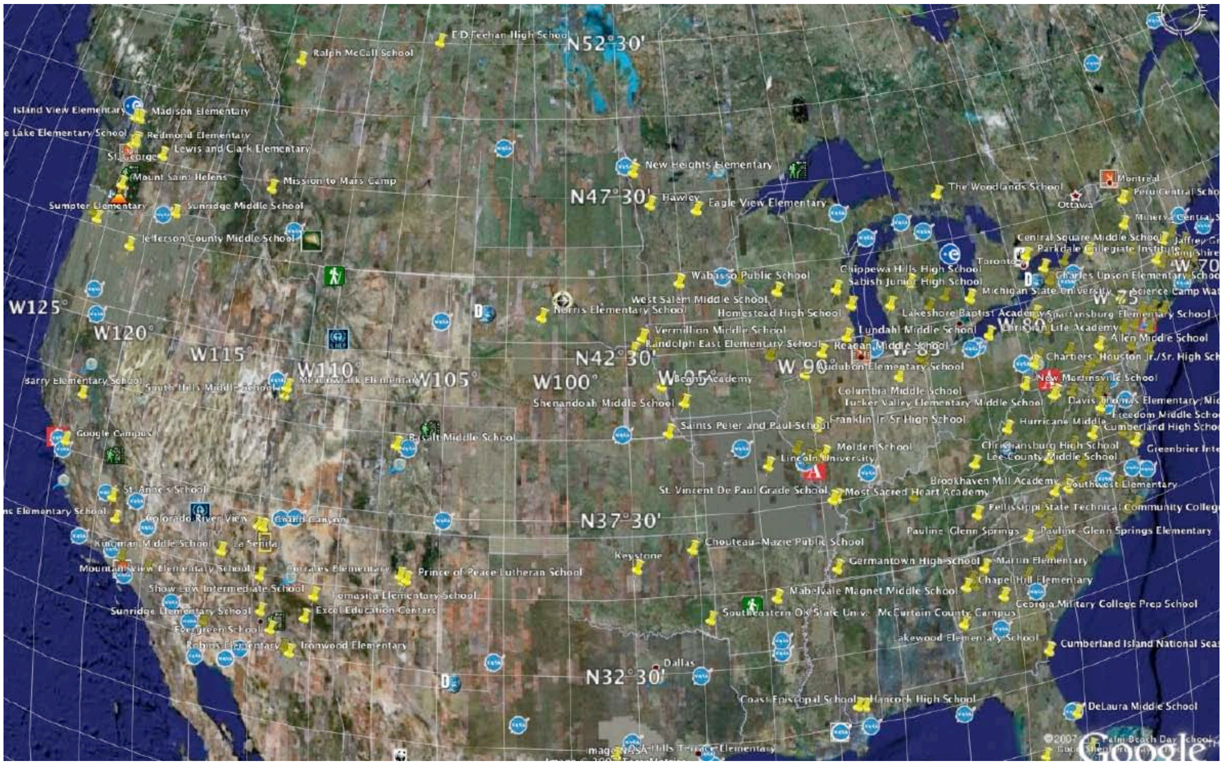

Figure 7. Active S'COOL participants in North America, as visualized through Google™ Earth.

2) View all schools that are currently observing. This is defined as any school that has submitted a report within the last two years. See Fig. 7.
Both of these files also give the school's name and the number of observations they have submitted to S'COOL.

5. CONCLUSION

Thanks to the efforts of the Clouds group during the summer of 2007, several new and user-friendly data analysis tools are now posted on the S'COOL website for use by the project participants in schools around the US and the world. Unlike previous data analysis tools for the project, these use modern software which is more accessible to the K-12 audience for S'COOL. In many states, in fact, the ability to use spreadsheet software is a required learning standard around the middle school grade level. Thus, the S'COOL project can offer educators yet another enriching way to achieve their educational goals.

6. ACKNOWLEDGMENTS

D. F. Young of NASA Langley Research Center created the initial spreadsheet analysis file for the S'COOL Project.

Sarah Niemeier of Texas began the cloud mask study during the Summer 2006 SPHERE program.

NASA's Science Mission Directorate provided funding for the SPHERE program through grant NNG06GH31G.

7. REFERENCES

Chambers, L. H., D. F. Young, P. K. Costulis, P. T. Detweiler, D. B. Stoddard, R. Sepulveda, Joyce D. Watkins, and A. Falcone, 2003: The CERES S'COOL Project, *Bull. Amer. Meteor. Soc.*, **84**, pp. 759-765.

Stackhouse, P. W. , D. P. Kratz, G. R. McGarragh, S. K. Gupta and E. B. Geier, ``Fast longwave and shortwave (FLASH) flux models.'' 12th Conference on Atmospheric Radiation, Madison, Wisconsin, July, 2006.

Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee, G. L. Smith, J. E. Cooper, 1996: Clouds and the Earth's radiant energy system (CERES): An earth observing system experiment, *Bull. Amer. Meteor. Soc.*, **77**, pp. 853-868.