CHOOSING THE MOST ACCURATE THRESHOLDS IN A CLOUD DETECTION ALGORITHM FOR MODIS IMAGERY

TRACEY A. DORIAN
National Weather Center Research Experiences for Undergraduates Program & Pennsylvania State University

MICHAEL W. DOUGLAS
NOAA's Office of Oceanic and Atmospheric Research & National Severe Storms Laboratory

ABSTRACT

We use a cloud detection algorithm that detects cloudy pixels from MODIS images by characterizing individual pixels as cloudy or non-cloudy based on the brightness values of the pixels and a predetermined threshold. The algorithm then produces mean fields of daytime cloudiness over different geographical regions. Although the cloud climatologies produced initially appeared realistic, it was found that the algorithm largely underestimated the cloud frequencies over some regions when using a threshold of *215*. Analyzing various MODIS images and recording cloudiness over different sectors served as the "ground truth" data which we compared to the algorithm output. After comparing the subjective estimates and the algorithm output for four regions of the world, we found that the algorithm underestimates cloudiness over these additional regions and that lowering the thresholds to *170-190* over oceans and *190-215* over land generally identified the thick clouds most accurately. Studying more regions or extending research on certain regions will allow us to better understand how the algorithm behaves with certain types of cloudiness and geography. Even though the thresholding technique is somewhat arbitrary, by better understanding how the algorithm behaves we can modify the algorithm to ensure that the output more accurately describes cloud climatologies around the world. If we are able to do this, then our algorithm could be used for many applications such as validating the numerical model simulations of cloud climatologies or assessing climate and potential climate change.

1. INTRODUCTION

Cloud climatologies have been developed in recent decades with increasing dependence on satellite imagery. There are many cloud detection algorithms that use different methods to produce cloud climatologies. For example, the CLAVR-1 (Cloud Advanced Very High Resolution Radiometer) algorithm classifies pixels in 4km resolution images into clear, mixed, and cloudy categories (Stowe et al., 1999). Another algorithm has been used by Ackerman et al (2003) to compare cloudiness from MODIS imagery with observations from radar and lidar products; they found that the MODIS algorithm agreed with the lidar about 85% of the time. And still there are multi-spectral algorithms that determine daytime cloud type using satellite imager data from AVHRR (The Advanced Very High Resolution Radiometer) and VIIRS (The Visible/Infrared Imager/Radiometer Suite) (Pavolonis et al., 2005).

Most cloud climatologies are available at 5-10 km resolution or greater. An example of such a cloud product is one derived from using the AIRS instrument which was launched in 2002 onboard the Aqua satellite and which has a spatial resolution of 13.5 km at nadir (Stubenrauch et al., 2010). The resolution of the MODIS visible images is considerably higher - at least 500m (two of the 3 frequencies making the color images are sampled at 500m pixel size while the red band is at 250m). With such high resolution, not only can we see the clouds over a region, but we can relate their occurrence to the underlying geography. This is

[1] *Corresponding author address:* Tracey A. Dorian, 1409 Allan Lane West Chester, PA 19380, tad240@psu.edu

very useful in determining what types of clouds form over different types of surfaces and how and why those clouds form (i.e. what kinds of meteorological processes are responsible for the formation of clouds).

We have been producing cloud climatologies at full resolution from MODIS imagery for different regions of the world. Having an algorithm that could reliably describe cloud climatologies around the world would be useful for some forecasting applications. Having an idea of cloud cover on a global scale could help climatologists predict what type of cloud cover one would expect in the future for certain regions. Ecologists needing to understand the distribution of cloud forests and other vegetation types could also use MODIS images and our algorithm (Douglas et al, 2006). Accurate cloud climatologies can also be used for climate studies.

We have used a simple algorithm to distinguish cloudy from cloud-free pixels based on the brightness of the pixel. Past work has used a threshold of *215* (0 is black, 255 is white) where anything higher than *215* brightness was considered cloudy. This threshold produced mean patterns of cloudiness that appeared very realistic. However, we found that for some regions the algorithm seemed to underestimate the cloud amount from casual inspection of the imagery and also ground reports from Peru where sites with continuous cloudiness for an entire month were analyzed by the algorithm to have 25% cloudiness. So although the MODIS images have relatively high spatial resolution, the cloud algorithm initially used to identify cloudy from clear pixels seemed to underestimate the cloud coverage over certain regions of the world, especially over oceans. The algorithm also cannot distinguish between bright land surfaces and clouds and mistakenly counts bright surfaces as clouds. With too low of a threshold, the algorithm sometimes mistakenly counts bright land surfaces as clouds and so tends to overestimate the cloud frequency. With too high of a threshold, the algorithm does not catch the thinner clouds over a region and so underestimates the cloud frequency. What we did was to try to match our climatology to the observed cloudiness seen by ground observers so that we can re-evaluate our procedure. By better understanding where the biases of the algorithm appear and making corrections to the algorithm based on those biases, we can improve the accuracy of this algorithm.

2. DATA AND METHODOLOGY

The “Moderate Resolution Imaging Spectro-radiometer” (MODIS) is an instrument onboard two different NASA satellites. The two satellites are called Terra and Aqua. Terra is a morning satellite which overpasses at 1030LT and Aqua is an afternoon satellite which overpasses at 1330LT. The satellites orbit the earth from pole to pole about 705km above the earth and provide global coverage every 1-2 days by sweeping 2,330km swaths. The MODIS imagery available from the website has the potential to produce very high (~250m) resolution cloud climatologies with global coverage.

The daily MODIS images used in this study were downloaded from a National Aeronautics and Space Administration (NASA) website[1]. Because of time limitations, only select regions were used for evaluating the cloudiness. These sectors included imagery from Hawaii, off the southern California coast, from northern Peru including both tropical forests and over-ocean sections, and from equatorial Africa (***Figure 1***).

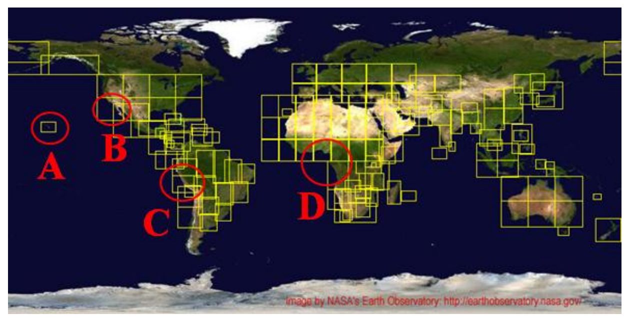


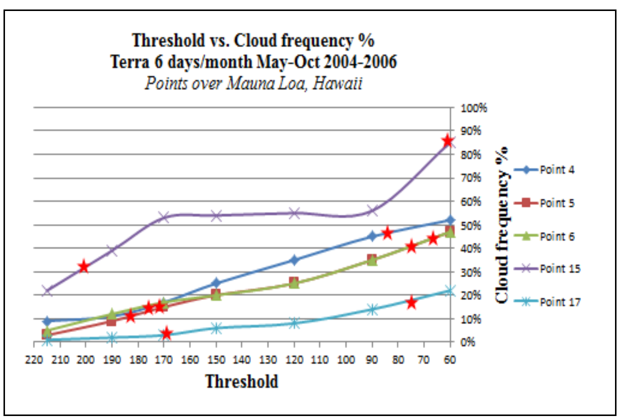
Figure 1: This figure shows the four regions, labeled A, B, C, and D, that we focused on for this research. Region A is Mauna Loa, Hawaii, Region B is off the coast of California, Region C is Northern Peru, and Region D is African sectors.

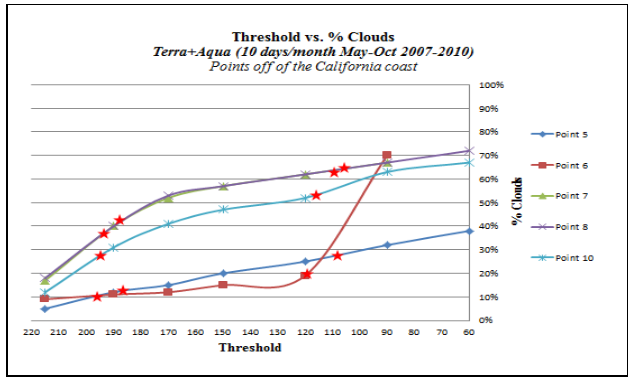
The time period that we looked at depended on the sector. The periods varied because some sectors have more years of data than other sectors (Douglas et al., 2010). After looking through hundreds of MODIS images, we characterized a point on the map as either being clear (category 1), thin clouds (category 2), or thick, bright clouds (category 3). If we looked through an entire years’ worth of data, then we would choose about 5-10 days per month because of time limitations. If, on the other hand, we looked at months May-October only, then we would go through every day of the months for however many years we were examining. After assigning a cloud category to a region for a certain number of years, the next step would be to calculate the frequency of each of those categories per year within the whole time period. Then we average over the entire interval. It was these averages throughout the entire period that we ended up comparing with the output of the cloud detection algorithm.

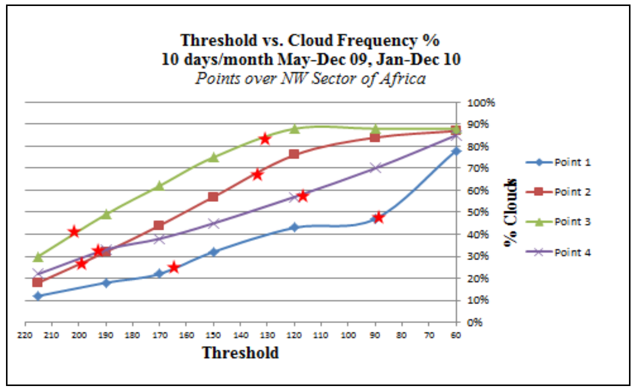
The algorithm extracts cloud data from MODIS images by characterizing pixels as cloudy or not cloudy depending on what the chosen threshold is. The way the cloud discrimination works is very similar to that by Reinke et al (2002). The novel aspect of our work is to compare the cloud frequencies generated by our algorithm with estimates obtained by visual examination of the same imagery. Subjective estimates have the advantage of more easily distinguishing clouds from thick dust, smoke or the specular reflection off a smooth sea surface. Our objective is to determine whether the output from our algorithm using different thresholds can be matched to our subjective estimates of the cloudiness using the same imagery data base. The best match is the threshold where the percentages agree with each other and we get the least "false clouds." Because running the cloud threshold algorithm took some time, only select thresholds were chosen, namely *215, 190, 170, 150 and 120.*

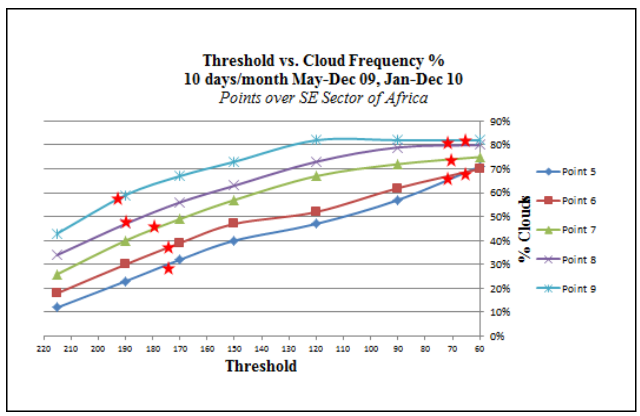
3. RESULTS/DISCUSSION

In order to find the most accurate thresholds for the four regions that we focused on, we created graphs showing cloud frequency vs. threshold for my estimates of cloud frequency. In this way we can see how the % of cloudiness changes with a changing threshold and where our subjective estimates fall compared to the algorithm cloud frequencies. ***Figure 2*** shows the graphs of cloud frequency vs. threshold for the four regions using the Terra + Aqua output.









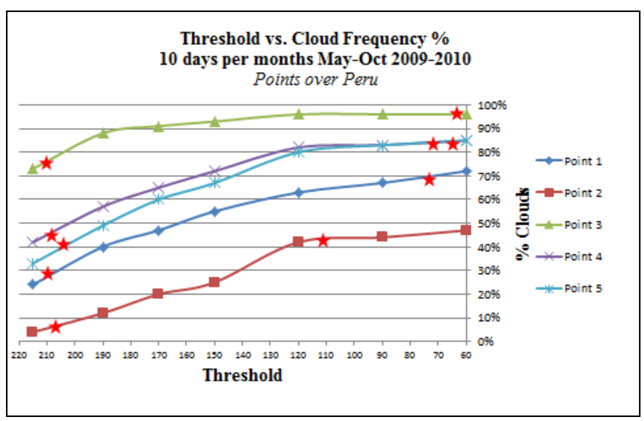


Figure 2: This figure shoes the graphs of cloud frequency vs. threshold for the Terra + Aqua algorithm output for different points from the sectors a) near Hawaii, b) off the coast of California, c) the Nigerian coastal area and the over-ocean region near Cameroon, and d) over northern Peru. The cluster of red stars closer to the left side of the graph indicates where our subjective estimates for thick clouds fell in relation to the threshold lines. The cluster of red stars closer to the right side of the graphs indicates where our subjective estimates for both thick and thin clouds fell in relation to the threshold lines.

For the Mauna Loa Sector we also performed an experiment using different sample sizes to see the changes in the cloud frequencies. We examined five points, two in the upwind "undisturbed" trade wind region, two downwind and one over the island in a region with high cloud amounts. We originally started with 6 days/month (spread out) for months May through October from 2004-2010. We then increased the sample size and added 5 more days to each month *just in the year 2004*, making it 11 days/month for the months May-October in 2004. Finally, we increased the sample size even more to include every single day of months May through October for 2004. This effort showed that changing the sample size did not produce a large change in the cloudiness values and that the maximum difference between the different sample sizes was about 5% (***Table 1).***

Just in 2004:

Cloud Frequencies for Point 5 (Terra)	%	~%
Cloud frequency for 1 with 6 days/month for May-October in 2004 (%)	64.70588235	65%
Cloud frequency for 2 with 6 days/month for May-October in 2004 (%)	23.52941176	24%
Cloud frequency for 3 with 6 days/month for May-October in 2004 (%)	11.76470588	12%
Cloud frequency for 1 with 11 days/month for May-October in 2004 (%)	68.85245902	69%
Cloud frequency for 2 with 11 days/month for May-October in 2004 (%)	24.59016393	25%
Cloud frequency for 3 with 11 days/month for May-October in 2004 (%)	6.557377049	7%
Cloud frequency for 1 with all days/month for May-October in 2004 (%)	69.04761905	70%
Cloud frequency for 2 with all days/month for May-October in 2004 (%)	20.23809524	20%
Cloud frequency for 3 with all days/month for May-October in 2004 (%)	10.71428571	11%

Table 1: The cloud frequency percentages for clear (1), thin (2) and thick (3) categories for different sample sizes for a point over the ocean near Hawaii. For category 1 there was a change from 65% to 69% to 70%, for category 2 there was a change from 24% to 25% to 20%, and for category 3 there was a change from 12% to 7% to 11%.

After making subjective estimates of levels of cloudiness for points 1 through 5, , we compared the algorithm's output for these same points for the thresholds of *215*, *190*, and *170.* The most accurate threshold range for over land was around *190-215* whereas the most accurate threshold range for over the oceans was around *170-190.* An image of the Mauna Loa Sector over Hawaii is shown with the cloud frequencies and most accurate thresholds for the five points in ***Figure 3.***

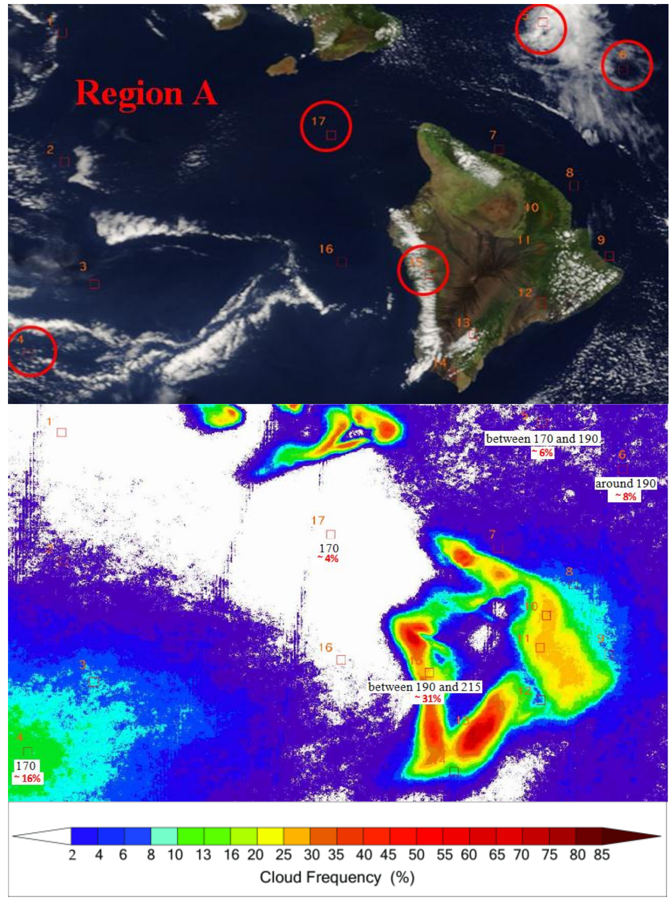


Figure 3: *Top:* The five points that we looked over for the Hawaii sector (Region A). *Bottom:* The cloud frequencies (in red) calculated for the *Terra* run 6 days/month May-Oct 2004-2006. In black are the results of the most accurate thresholds over Mauna Loa, Hawaii with the *215* threshold cloud frequency graph. The background is the *215* threshold color scheme. Thresholds of about *170-190* seemed like sufficient thresholds for over the oceans, whereas higher thresholds of about *190-215* seemed like sufficient thresholds over land. This goes along with the fact that oceans have low albedo whereas land surfaces have more reflective surfaces. The land seemed to have much higher cloud frequencies than over the oceans for all thresholds over Hawaii. This is most likely because the land surface over Hawaii is not constant in elevation, and so there is forced lifting and convection over the island during the daytime as the trade wind

easterlies interact with the island topography and flow around the island (Leopold 1949 pg. 319).

One important limitation of these ideal thresholds is that we are comparing the cloud frequencies of category 3's to the algorithm output, and so we are really just finding the most accurate threshold to capture the thick, bright clouds. This way we can say that this threshold can *at least* accurately pick up the thicker clouds. In order to catch the thinner clouds, the thresholds would need to be lower. The reality is that the algorithm may never catch all of the clouds that are over a region, because even the lowest thresholds may not pick up very thin cirrus or stratus clouds and may instead pick up brighter land surfaces.

The next area where cloud frequencies were evaluated with Aqua and Terra imagery was over the California coast for the period May-October 2007-2010. The subjective estimates of cloudiness indicated maximum cloudiness over the ocean points usually occurred in the *morning hours,* or the Terra overpass. In other words, cloud coverage tended to decrease from morning to afternoon over the Pacific Ocean for my sample points. This may be a result of daytime heating of the cloud layer which actually "breaks" tends to dissipate the low stratus clouds over the Pacific Ocean by causing the water molecules within the cloud droplets to evaporate. The stratus decks are often seen on the west coasts of continents because of warm air moving over colder waters (which were caused by upwelling). For the one land point though, the same amount of cloud coverage was more or less present between the morning hours and the afternoon hours, with bright thick clouds 10% of the time in the morning and about 9% in the afternoon. The cloud frequencies that we found from examining the MODIS images and the thresholds that were found to be most accurate for the Aqua + Terra images are summarized in ***Figure 4.***

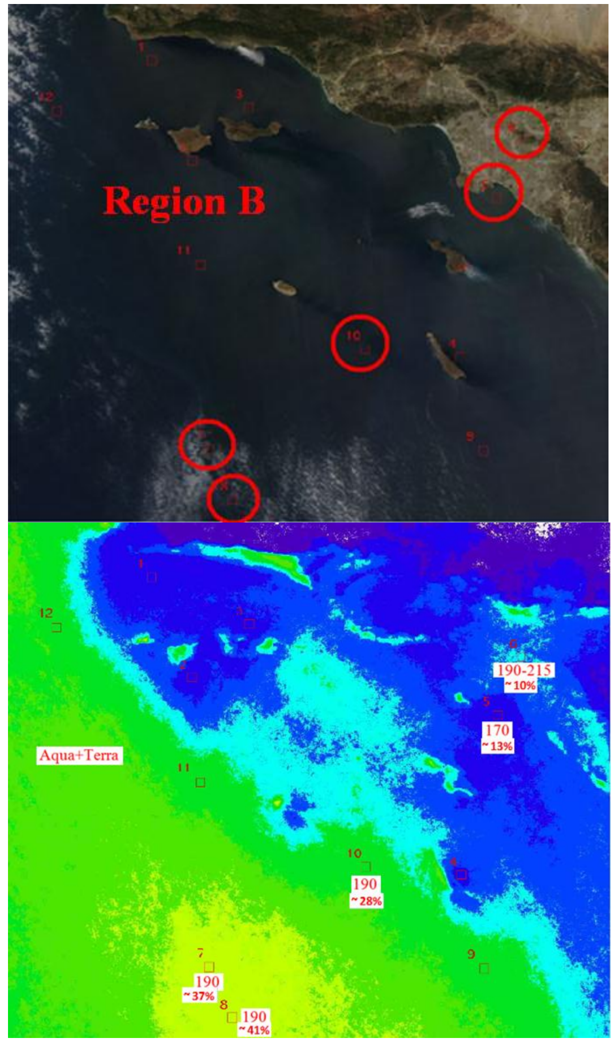


Figure 4: *Top:* The five points that I looked at over California coast and part of the eastern Pacific Ocean (Region B). *Bottom:* The subjective cloud frequency estimates and the most accurate thresholds for each of the five points with a *215* threshold background.

The overall best threshold for ocean points was between *170* and *190.* For the land point over California (Point 6), it looked like a good threshold would be between *190* and *215*. These thresholds are very similar to the thresholds that we calculated for the Hawaiian points, where a threshold of about *170* or *180* seemed ideal for ocean points, and somewhere around *200* seemed like an ideal threshold for land points.

The next region that we focused on was over northern Peru including some tropical forests. We focused on five points over this sector and looked at 10 days/month for months January through December in the years 2009 and 2010. ***Figure 5*** shows the average of the Terra and Aqua cloud frequencies that we estimated for the five points, and so it essentially shows the cloud climatologies around noontime. The figure also shows a map of cloud frequency and most accurate thresholds overlaying the *215* threshold algorithm output. The ideal thresholds over Peru ranged between *190* and *215*. What was interesting about this particular region was the extreme variation between the cloud frequency over Point 3 and the cloud frequency over Point 4.

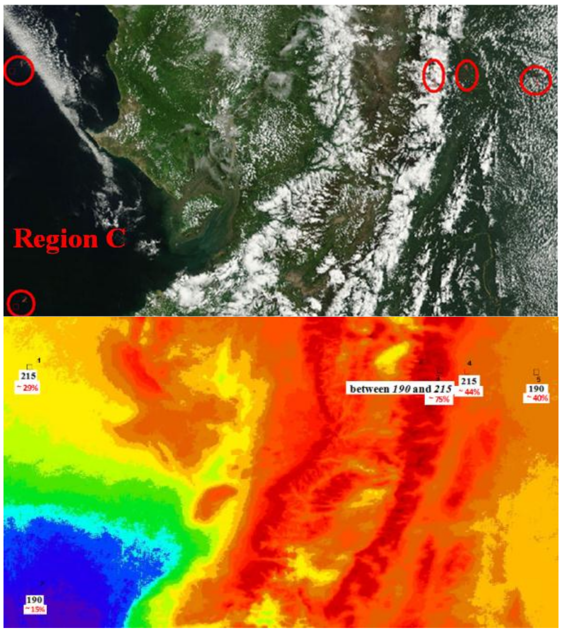


Figure 5: *Top*: The five points that I looked at over the Peru sector (Region C). *Bottom:* This figure shows the most accurate thresholds and the cloud frequency percentages for the five points over Peru for months January through December 2009-2010 (Aqua + Terra). The background is the *215* threshold output.

Both Points 3 and 4 are very close to each other, and yet Point 3 had around 75% cloud coverage while Point 4 had around 44% cloud coverage. Taking a closer look in "Google Earth" showed that Point 3 lies on a steep slope of the Andes while point 4 is on flat land. The daytime ascent of air along the slope likely explains the higher cloud amounts at Point 3**.** Points 4 and 5 are next to each other as well but have very similar cloud frequencies.

After Peru we looked at both the Northern sector and Southern sector of Africa. First we looked at four points off of the coast of Western Africa in the Northern sector. The time period we looked at was May-December 2009 and January-December 2010, and we looked at 10 days per month. Between the Terra and the Aqua cloudiness means, the cloud frequency increases between the Terra and Aqua overpass times were small, 0-6% for the four sites. The cloudiest point seemed to be Point 3, which was a point over a city, with about 40% of the time having bright, thick clouds in the time period.

After looking through the output of the different thresholds, we found that a threshold of about *190* seemed to be most accurate for all the points except for Point 3, which needed a threshold of *between 190 and 215* ***(Figure 6).***

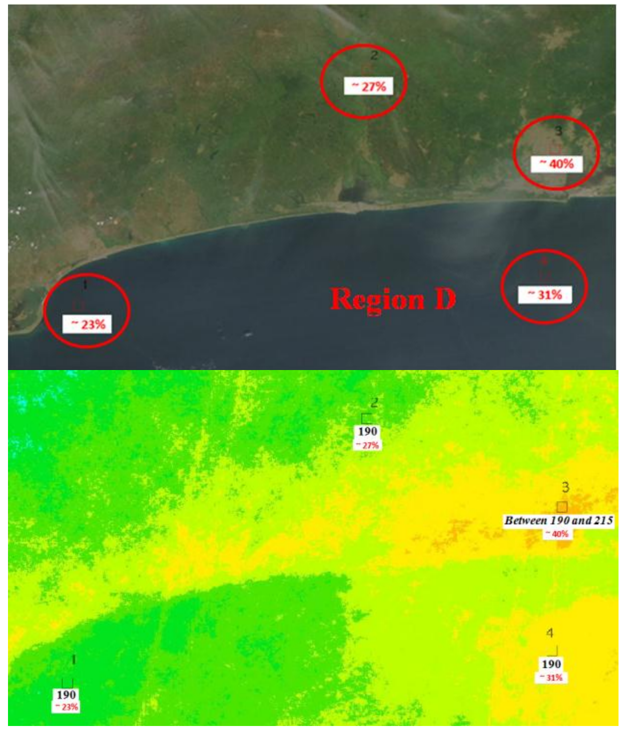


Figure 6: *Top:* The four points that we looked at for the NW Sector of Africa. *Bottom:* The most accurate thresholds and the subjectively determined cloud frequency percentages for the four points over the NW Sector of Africa for Aqua + Terra 10 days/month May-Dec 2009 and 10 days/month Jan-Dec 2010. The background is the *215* threshold output. Beneath this image is the color bar for the *215* threshold that we used to read the cloud percentages from the algorithm output.

For the Southern sector, we looked at 5 more points off the African coast. Of all the points we found that Point 9 had the largest cloud frequency of about 58% of category 3's within the time period. The most accurate thresholds for Aqua + Terra that I found were *170* for Points 5, 6, and 7, and *190* for Points 8 and 9 ***(Figure 7).*** A figure shows the four regions we looked at and both the cloud frequencies and the most accurate thresholds we found for those four regions.

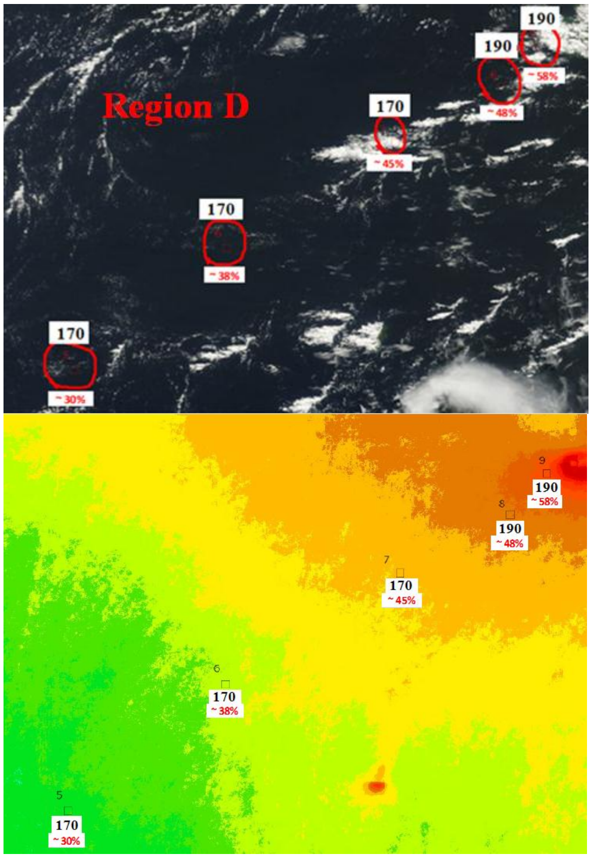
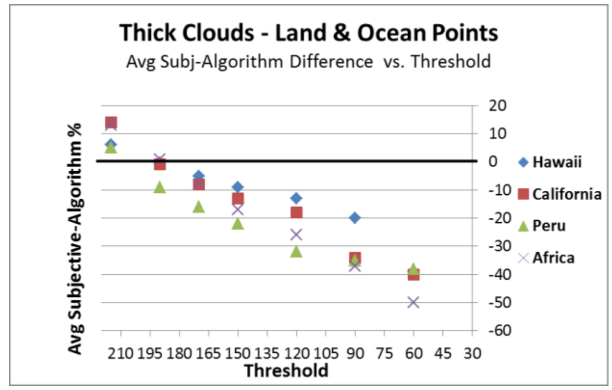
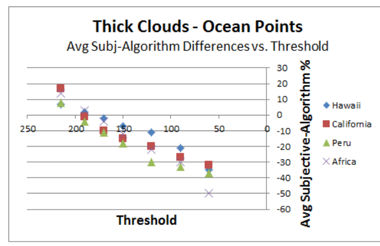


Figure 7: *Top:* The five points that we looked over for the SE sector of Africa. *Bottom:* The most accurate thresholds and the cloud frequency percentages for the five points over the SE Sector of Africa for Aqua +Terra 10 days/month May-December 2009 and January-December 2010. The background is the *215* threshold output.

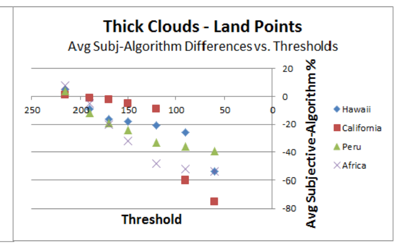
One last thing that we did was to look at which thresholds were most accurate for all four of the regions combined. We created graphs of the difference between our subjective estimates and the algorithm percentages versus each individual threshold. We looked at thick clouds over land and ocean points, over only ocean points, and over only land points. Additionally we looked at both thick and thin clouds over land and ocean points, over only ocean points, and over only land points (***Figure 8***).



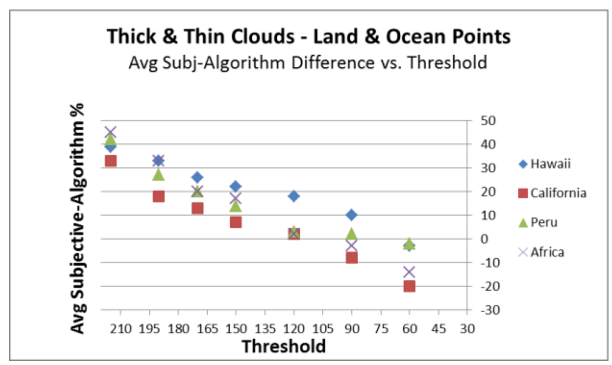
a)



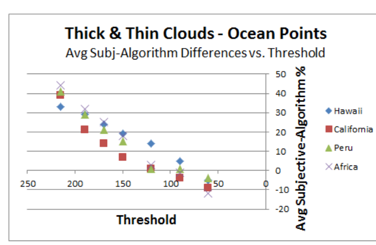
b)



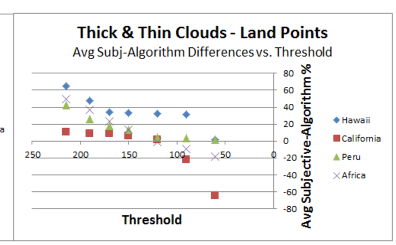
c)



a)



b)



c)

Figure 8: These figures show the average differences between our subjective estimates and the algorithm cloud frequency percentages for the individual

thresholds. a) A graph for thick clouds over both ocean and land points, b) A graph for thick clouds over only ocean points, c) A graph for thick clouds over only land points, d) A graph for thick and thin clouds over both land and ocean points, e) A graph for thick and thin clouds over only ocean points, f) A graph for thick and thin clouds over only land points. The most accurate thresholds for these four regions are the areas on the graphs where the differences are closest to zero.

In order to find the most accurate threshold range, wherever the smallest differences between our subjective estimates and the algorithm cloud frequencies were was where the most accurate threshold was. Looking at only thick clouds over these four regions, the most accurate threshold range for just over ocean points was 170-190, over land point was 190-210, and over both land and ocean points was about 190-210. On the other hand, looking at *both* thick and thin clouds over these four regions, the most accurate threshold range for just over ocean points was 90-120, over land was 60-120, and over both land and ocean points was about 75-120.

4. CONCLUSION

After looking at these four different regions and viewing how our subjective estimates compare to the algorithm output, it is apparent that in order to get the most accurate results you need lower thresholds between 60 and 120 to capture thick and thin clouds over these four regions, and you need higher thresholds between 170 and 210 to capture just thick clouds over these four regions. It is useful to see the most accurate thresholds for capturing thin clouds and thick clouds separately because it is unrealistic to assume that the algorithm can ever catch all clouds over a region. By dividing thick and thin clouds into separate categories, we can better see how the algorithm behaves with different types of cloudiness. Future areas of research could include finding the most accurate thresholds for additional regions and viewing more points within the sectors.

There are limitations of our work. One such limitation was that the algorithm only identifies bright thick clouds. In order to find out the true percentage of clouds, both thick and thin, one should use two different thresholds. Another limitation of this work was our subjective estimates of cloudiness over particular regions. Obviously subjective estimates require distinguishing thick from thin clouds. Sometimes it was difficult to decide whether the clouds should be considered thin or thick clouds. We generally followed the same judgment when going through points to distinguish between what was considered “thin” and what was considered thick. Problems also arose with reading the algorithm output. The algorithm output was judged using a color scale, but as the cloud frequency percentage increased, it got harder and harder to tell what percentage matched with the colors. Finally, we only looked at specific thresholds, namely *215, 190, 170, 150, 120, 90, and 60.* For certain regions, it is possible that the most accurate threshold may have very well landed between those chosen thresholds.

Clouds play a very important role in the amount of solar energy that reaches the earth’s surface, and also the amount of infrared radiation that is released or not released from our atmosphere. Cloud climatologies produced using 5km resolution, 250-500m resolution, or any other resolution can help us determine small-scale weather changes over different terrain like sea-land breeze. Using our new data set of MODIS images from the NASA website can help provide a new perspective on global cloud climatology, and climatologists everywhere can view our results and compare our results to previously produced climatologies. Perfect replications of cloud climatologies involves lot of teamwork from groups of scientists eager to produce accurate cloud fields, and also involves a lot of time and dedication to produce the best possible cloud climatology products that can be used by anyone who is curious and interested.

5. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. AGS-1062932. Special thanks also to the individuals responsible for maintaining the MODIS images used for this project. Also we extend our thanks to Abdul Dominguez, Robert Nelson, Matt Jay, and Rahama Beida for their help and support with this research.

6. BIBLIOGRAPHY/REFERENCES

Douglas, M., R. Beida, and A. Dominguez, 2010: Developing High Spatial Resolution Daytime Cloud Climatologies for Africa. Extended Abstracts. *29th Conference on Hurricanes and Tropical Meteorology (May 9, 2010.)* Tucson, AZ, USA. American Meteorological Society, P2.23.

Douglas, M., T. Killeen, and J.F. Mejia, 2006: Use of MODIS and GOES imagery to help delineate the distribution of cloud forests along the eastern Andean slopes. *14th Conference on Satellite Meteorology and Oceanography* P3.18. Feb 1 2006.

Luna, L.B, 1949: The Interaction of Trade Wind and Sea Breeze. *Journal of Meteorology.* Vol. 6, Issue 5. pp. 312-320.

Pavolonis M.J., A.K. Heidinger, and T. Uttal, 2005: Daytime Global Cloud Typing from AVHRR and VIIRS: Algorithm Description, Validation, and Comparisons. *Journal of Applied Meteorology*. Volume 44, Issue 6, pp. 804-826

Reinke, D.L., C.L. Combs, S.Q. Kidder, and T. H. Vonder Haar, 1992: Satellite Cloud Composite Climatologies: A New High Resolution Tool in Atmospheric Research and Forecasting. *Bulletin of the American Meteorological Society*. Vol. 73, Issue 3, pp. 278-285

Short, D.A. and W.M. John, 1980: Satellite-Infrared Morning-to-Evening Cloudiness Changes. *Monthly Weather Review.* Vol. 108, Issue 8. pp. 1160-1169

Stowe, L.L, P.A. Davis, and E.P. McClain, 1999: Scientific Basis and Initial Evaluation of the CLAVR-1 Global Clear/Cloud Classification Algorithm for the Advanced Very High Resolution Radiometer. *Journal of Atmospheric and Oceanic Technology.* Vol 16, Issue 6. pp. 656–681.

Stubenrauch C.J., S. Cros, A. Guignard, and N. Lamquin., 2010: A 6-year global cloud climatology from the Atmospheric Infrared Sounder AIRS and a statistical analysis in synergy with CALIPSO and CloudSat. *Atmospheric Chemistry and Physics (ACP).* pp. 8247-8296.