ATMO 611 – Satellite Data Applications Lab Assignment #1

50 Points Total Available

Introduction to MODIS, Matlab and File Types – DUE Friday Jan 23rd, 2015

This lab is meant to introduce you to the MODIS file format system and simple functions in Matlab to open, extract and manipulate data. You will be required to include responses to questions, figures and Matlab code with your completed lab write up. Questions to be answered are labeled **Q1**, **Q2**, etc. and are highlighted in Red.

1 MODIS FILE FORMAT BACKGROUND

The data are in Hierarchical Data Format - Earth Observing System (HDF-EOS) format.

HDF-EOS is the standard data format for all EOS data products. HDF-EOS is a multi-object file format developed by The HDF Group. For more information about the HDF-EOS format, tools for extracting binary and ASCII objects from HDF, and a list of other HDF-EOS resources, see the NSIDC HDF-EOS Web pages. Visit and read through the information on the following pages:

- HDF Group Info : http://www.hdfgroup.org/why_hdf/
- NSIDC webpage http://nsidc.org/data/hdfeos/intro.html

Q1 – (2.5 points) Summarize the HDF file format system. MODIS data is stored in HDF-EOS files. How do these files differ than regular HDF files? Describe the benefits of the HDF format for earth, ocean, land and atmospheric data.

2 DOWNLOADING DATA:

1) Downloading MODIS Atmosphere Data
- Visit the main MODIS Atmosphere Data Site: http://modis-atmos.gsfc.nasa.gov/
- Click on "Monthly" from the Menu on the Right Hand Side of the page (near the photos)
  READ through this section to get a better understanding.

Q2 – (2.5 points) MODIS uses an "Equal-Angle" Grid for Level 3 Data. Define what "equal-angle grid" and "equal-area grid" mean. How are they different? What is one drawback of using and equal-angle grid vs. the equal-area grid structure?

- Go to LAADS - https://ladsweb.modaps.eosdis.nasa.gov/search/order/1
- Select the following using the tabs.
  - a) Sensor: MODIS: Aqua
  - b) Collection: 6 - MODIS Collection 6, Level 1, Atmosphere, Land
  - c) Products: L3 Atmosphere Product [3]
  - d) Products: MYDO8_M3
  - d) Temporal Selection: August 1-31st
  - e) Location: World (default)

f) Files: There should only be one. You can right click to down load it on the right. It's large (393 MB) so it will take some time to download.

What are MODIS Data "Collections"? A MODIS data "Collection" is basically a MODIS data version. When new & improved science algorithms are developed, the entire MODIS dataset (from launch) is reprocessed and then tagged & distributed as a new "Collection". During the processing of a Collection, an attempt is made to use the same version of the Science Algorithms or Program Executables (PGEs). However, sometimes a bug is found in one or more of the PGEs in the middle of Collection processing; and if the bug is not serious, processing will complete with the new corrected PGE. These anomalies and problems in processing are noted on the Known Problems page. One can always identify the Collection number for a particular HDF file as it's always included (as a 3 digit number) as part of the HDF filename. There have been six MODIS data Collections (or Versions) processed since MODIS/Terra was launched in early 2000. The Collection versions created thus far are 001, 003, 004, 005, 051, and 006. It should be noted that Collection 051 only contained updates for some MODIS Data Products, which is why it was tagged with a 051 (a surrogate for version 5.1), since it was considered only a partial update of MODIS Data Products. Finally, changes for the current Collection (006) are completed. Processing and availability of Collection 6.1 data is progressing.

MYD08_M3.A2017213.006.2017249165809.hdf

Q3 – (2.5 points) MODIS File names are extremely helpful in telling you what data is stored inside. Deconstruct the file name you just downloaded explain what each part means. You may need to Google and search the MODIS data sites to determine the naming conventions.

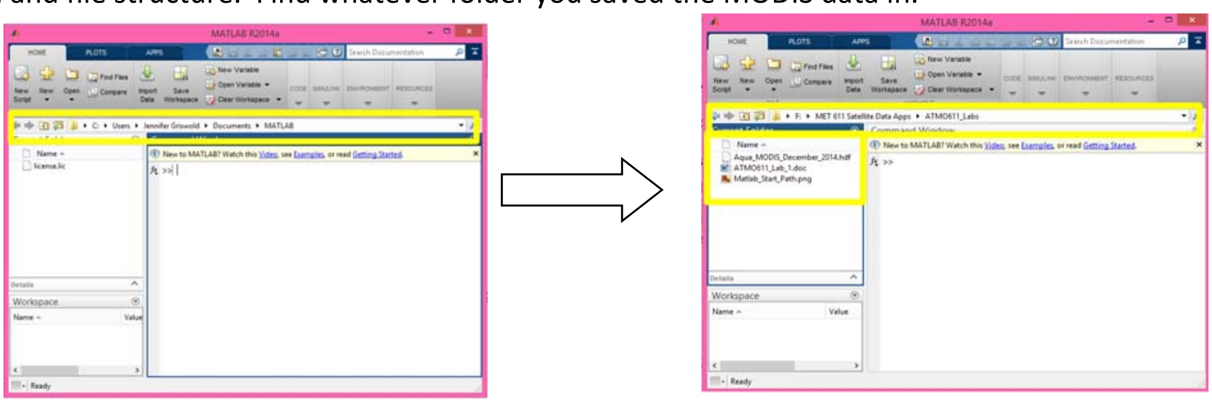
k) Rename this file: Aqua_MODIS_August_2017.hdf

3 BROWSING THE HDF FILE

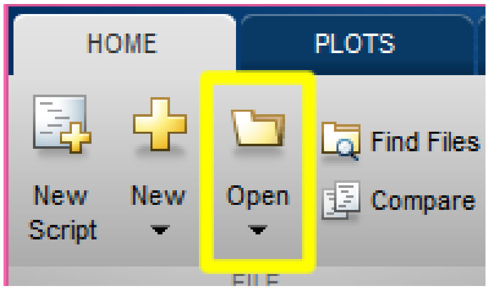
This next step allows you to "peak" into the HDF file to see what is stored inside. You will need to open Matlab at this time.

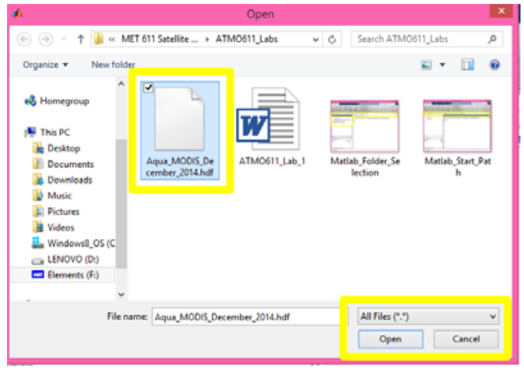
1) Start Matlab on your computer (either the lab machine or your own person laptop or desktop)

2) Navigate to where you stored your MODIS hdf file using your "Path" This is the bar that has your drive, and file structure. Find whatever folder you saved the MODIS data in.



3) Accessing HDF data requires knowledge of how the data is stored (how many dimensions), the dates and/or times associated with each file, and the variables available. Before you can write a script to extract data from your newly downloaded HDF file you need to "view and browse" through the HDF file and take notes. To open and browse through the HDF4 or HDF-EOS file, select **Open** from the **Home** tab. You can open multiple files in the HDF Import Tool by selecting Open from the Home tab. This will allow you to select your file.





4) Once you Click "Open" the "HDF Import Tool" Will open a new window.

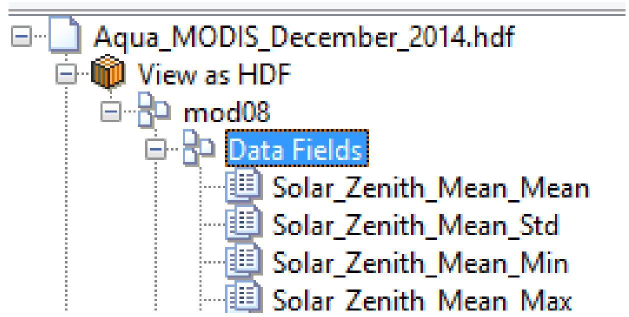
5) Click on the "+" next to "View as HDF"

6) Click on the "+" next to "mod08"

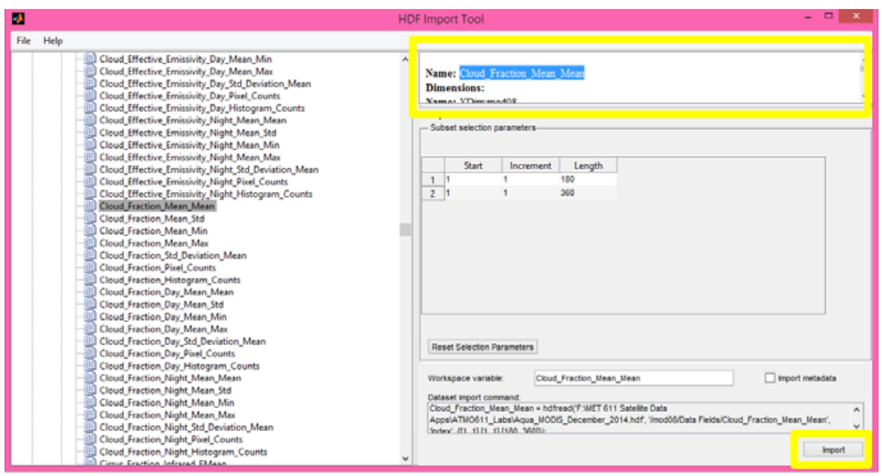
7) Click on the "+" next to "Data Fields"

This brings a drop down of ALL the Variables stored in the hdf file.

8) Browse through to find the Variable named "**Cloud_Fraction_Mean_Mean**"



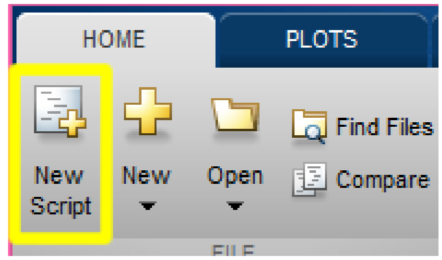
9) This will bring up information on this particular variable. You can learn very important information about whatever variable you have chosen.



Q4 – (5 points) Using the HDF Tool summarize key information about the variable you have chosen above. What are the grid dimensions, what do they represent in terms of latitude and longitude? What are the units? What is the scale factor? What is the offset? What is the fill value for no data? What is the long name, what does it tell you about the variable?

10) I know you're tempted to Click "Import" to load that single variable into Matlab. But we're not going to do that using the HDF Import tool. We're going to use the "hdfread" function. Make sure to jot down that the variable we want is called "Cloud_Fraction_Mean_Mean"

4 CREATING A NEW MATLAB SCRIPT (.m)



1) Open a new “m” file or Matlab script. We’ll be adding lines of code progressively through this portion of the lab and then you will add your own (modified) code to look at three other variables later on in the lab. To open a new “m” file go to the “Home” tab and select “New Script. ” This will open up a new window. This is where you will start typing the commands and what will become your practice script/code.

2) Click on the Tab titled “Editor – Untitled” now go back up to your Home tabe and save your new “m” file as “MODIS_Practice_Lab_1_Your_Last_Name”

3) Go back to the “Editor” Tab and add a script header using the % sign. In order to type “comments” in Matlab you simply need to include a **%** symbol in front of the text. It can be at the start of the line or after some working code.

> *NOTE: Including a program header and comments throughout your script are critical to developing useful and usable scripts. I’ve included a sample header below. Important info to include are the script title, purpose, author, and date of most current revision. From personal experience this can save the day when you need to go back to an old script to re-process or re-plot data for your thesis.*

```
% TITLE: MODIS_Practice_Lab_1_Griswold.m
%
% PURPOSE: To practice manipulating Level 3 Global Monthly gridded Aqua
% MODIS data. This is to be used for the ATMO 611 course during the Fall
% 2017 semester at UH Manoa.
%
% DATE Version 1.0: 9-13-2017
% Author: Dr. Jennifer Griswold
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

5 LOADING THE VARIABLE USING hdfread

Instead of using the hdftool to import the variable we’ll use the hdfread function which has the format:

data = hdfread(‘filename.hdf’, ‘datasetname’);

Where “data” is the name that you choose to use for the variable and the file name is in single quotes with the full name and extention. I chose Aug_2017_CF to stand for August, 2017 Cloud Fraction. You enter in the exact variable name that you wrote down from the hdftool in the ‘datasetname’ space. It should look something like this:

```
% Load file using hdfread
Aug_2017_CF = hdfread('Aqua_MODIS_August_2017.hdf','Cloud_Fraction_Mean_Mean');
```

Note the “;” at the end means it won’t print your command, or the array, to the screen. You should include this after each line to make the code run faster and cleaner.

Either type the name of the code at the command line “>>” or click “RUN”

6 SIMPLE MANIPULATIONS TO AND MAPS OF THE IMPORTED VARIBLE

After you have imported the variable it should now show up in your Workspace. If you click on the variable it will open in a "Variable" window so you can view the values in each cell. Since you have loaded "Cloud_Fraction_Mean_Mean" you may noticed that the numbers seem high for a "Fractional" value. In Q4 you identified the "Scale Factor" and "Offset" that need to be applied in order to use the data properly.

1) Converting the array to a double precision array. I prefer to convert my data into the "double" format right at the beginning for higher precision calculations later on. Add the following, using your own variable name, to convert the data values to double precision.

```
% Making it a double precision number
Aug_2017_CF = double(Aug_2017_CF); % making it "double"
```

2) Now, let's do some easy manipulations to the data. Remember that the raw cloud fraction file has numbers in the muliple-thousands. You should have noted that you needed to apply a scaling factor to the data of 0.0001 and offset of 0. You can also rename the variable so it is easy to identify as the scaled array. You can do this is one line of code as shown below:

```
% Correcting cloud fraction to account for scaling factor and offset
Aug_2017_CF_Scaled = (Aug_2017_CF .* 0.0001) + 0 ; % No offset since it is equal to zero
```

> *NOTE: In order to multiply all the cells of a matrix/array you need ot use ".*" instead of simply "*" You can read about the matrix and array functions (+, -, *, and /) and their conventions in the help files.*

3) Now that you have a usable data set we will plot it in a gridded format that so we can add a global map outline. To make a realistic map we also need to define the latitude and longitude files. You can create these arrays in Matlab itself or import them from a file. We'll create it here in the code so you can see how that is done. Remember the MODIS Level 3 Data is on a 1° x 1° grid. That means 180 degress of latitude and 360 degress of longitude. To make an array of numbers you use square brackets "[]" and then need to transpose or switch the array from (1 row and XX columns) to (1 column and XX rows) by adding the apostropy at the end.

new_numerical_array = [start value : increment : end value]';

So, to make my new MODIS Latitude and Longitude arrays it would look like this:

```
% Creating Latitude and Longitude Arrays
MODIS_Lon = [-179.5:1:179.5]'; % This is in the correct alignment
MODIS_Lat_updown = [-89.5:1:89.5]'; % This is upside down (S on top of N)
MODIS_Lat = flipud(MODIS_Lat_updown); % flipud "flips" the array upside down
```

> *Note: that you have to start with the <u>smaller value -89.5</u>, in the case of Latitiude, even though it creates an array in which the Southern Hemisphere (the negative values) is on "top". In this case we use the "flipud" function to flip the array so that the Northern Hemisphere values are at the top of the array in the correct orientation. <u>If you don't flip the array your map will be upside down!</u>*

7) Now we can make a figure with the cloud fraction data and our latitude and longitude arrays. We do this by using the "figure" and "pcolor" functions. You'll also see in the example below how to interpolate, remove the grid, add a title and a colorbar. You will also be able to specify the axes to zoom in certain regions. You will be able to use this example code in the rest of the lab.

```
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
figure;                      % Prepares a new figure window
pcolor(MODIS_Lon, MODIS_Lat, Aug_2017_CF_Scaled);     % pseudocolor plot function (rainbow is default)
shading interp               % smooths data so it doesn't look like grid-boxes
hold on                      % tells matlab something else will be ploted
grid off                     % turning the grid off
axis([-179.75 179.75 -89.5 89.5])       % specifying the exact latitude and longitude box you want
title({'August 2017 Mean Cloud Fraction'},'FontSize',15) % Title
xlabel('Longitude','FontSize',12)   % x axis label
ylabel('Latitude','FontSize',12)    % y axis label
% add a standard colorbar.
h=colorbar('h','Location','EastOutside');
set(get(h,'ylabel'),'string','MODIS Aqua Cloud Fraction (unitless)');
%
```

Note: Use Matlab help to learn moreabout the figure, pcolor, shading, hold on, grid off, axis, title, xlabel, ylable, colorbar, and set functions. The Help files are extremely detailed.

Either type the name of th ecode at the command line ">>" or click "RUN" A figure window will pop up!

8) Adding the Outlines of the Continents

This looks fine, but it's hard to see the continents without an outline. Go to the class website and download the file "**world_coastline.csv**" and save it into the same folder as your data and script. It was also emailed along with the assignment and is on Laulima. It is a "comma separated file".

There are two columns in this file. Column 1 is the latitude and column 2 is the longitude.

Use the function **dlmread** to load the .csv file into your code before your figure. It will look something like this:

coastline_variable_name = dlmread('filename.csv', 'delimiter type', row start, column start);

- The file name will be world_coastline.csv
- The delimiter type in this case is comma, so you'd need to put ','
- The row start is 1 (since there is a NaN in the first row and we want to skip that)
- The column start is 0 (since we want to start at the first column)

To take out the latitude and longitude you can simply make a new variable and select the correct column. It will look something like this for latitude. Where the ":" tells Matlab you want ALL the rows and the FIRST Column.

lat = coastline_variable_name(:,1);

It will look something like this for longitude. Where the ":" tells Matlab you want ALL the rows and the SECOND Column.

lon = coastline_variable_name(:,2);

You essentially need to include three lines of code: one to load the coastline file, and two to take the out latitude and longitue columns.

8) Now we need to incorporate the world coastline into the figure code. To do this we simply need to add in one line of code after the “hold on” in the current example. “hold on” allows you to layer information in a single figure. Here we have added

plot(lat, lon, ‘k’);

“plot” adds the new data, “lat” is the y variable name, lon is the x variable name and ‘k’ tells Matlab what color to make the data points. Black is the easiest to see when using the standard colorbar (rainbow).

> *Note: Use the Matlab help to learn the different symbols used for the various colors by searching for “LineSpec”*

```
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
figure;                      % Prepares a new figure window
pcolor(MODIS_Lon, MODIS_Lat, Aug_2017_CF_Scaled);     % pseudocolor plot function
(rainbow is default)
shading interp               % smooths data so it doesn't look like grid-boxes
hold on                      % tells matlab something else will be plotted
plot(lat,lon,'k');           % adding the coastline, 'k' = black color line
grid off                     % turning the grid off
axis([-179.75 179.75 -89.5 89.5])      % specifying the exact latitude and longitude box
you want
title({'August 2017 Mean Cloud Fraction'},'FontSize',15) % Title
xlabel('Longitude','FontSize',12)   % x axis label
ylabel('Latitude','FontSize',12)    % y axis label
% add a standard colorbar.
h=colorbar('h','Location','EastOutside');
set(get(h,'ylabel'),'string','MODIS Aqua Cloud Fraction (unitless)');
%
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

Either type the name of th ecode at the command line “>>” or click “RUN” A figure window will pop up!

9) Save the figure as a .jpg so you can include it in your lab report.

Q5 – (5 points) Define Cloud Fraction. What bands are used to derive this variable? Using this test figure what global patterns can you identify in the August 2017 cloud fraction data? Where are cloud fractions highest in August? Where are they lowest? What global cloud features are related to these areas of high and low cloud fraction?

7 MODIFY THE SCRIPT TO LOAD THREE ADDITIONAL VARIABLES USING hdfread

1) Go back to the HDF Import to find information on the three variables listed below:

- Cloud_Optical_Thickness_Combined_Mean_Mean
- Cloud_Effective_Radius_Combined_Mean_Mean
- Cloud_Top_Pressure_Mean_Mean

2) Make sure to check the scale factors and offsets by using the HDF Import tool again. Once it’s open you can leave it open and search through multiple variables without closing the window.

ON YOUR OWN: You will also need to account for "bad" or "blank" data represented by -9999 in the data files. It is advised that you replace these -9999 with NaN. Use Matlab Help to see how you can search ("find" function) for the -9999 or > 0 values and replace them with NaNs.

3) Once you've added the necessary code save the figure files in .jpg format so you can include these three images, with appropriate titles and color bars into the final lab report.

8 ON YOUR OWN – ZOOMED IN FIGURES MAKING A FOUR-PANEL PLOT USING SUB-PLOT

1) Using the axis line of the example map code chang ethe latitude and longitude values to zoom in on a region of your own choce. This can be anywhere you notice an interesting patter in one of the variables.

```
axis([-179.75 179.75 -89.5 89.5])        % specifying the exact latitude and longitude box
```

2) One of the most important skills for using Matlab is to be able to search through the Help files and find what you need. This part of the assignment is meant to give you the change to use the Help files to greate a four panel plot including all four of the variables using your new Zoomed in Region. It doesn't matter which are on top o bottom, just that all four are in the same figure panel. Where you specify the number of tile rows (m) and coluns (n) and the plaement of a particular figure tile (p).

subplot(m,n,p)

9 QUESTIONS & DEFINITIONS: Use any resources you need (MODIS website, research articles, google)

Q6 – (5 points) Define Cloud Optical Thickness. What band(s) are used to derive this variable? Where do we see the thickest and thinnest clouds? What would you change about this figure to make it easier to identify regions with thick and thin clouds? (Excluding Antarctica)

Q7 – (5 points) Define Effective Radius. What band(s) are used to derive this variable? Where do we see the lowest cloud effective radius observations? Where do we see the largest droplet radii? (Excluding Antarctica) How can you improve on this figure? What would you change?

Q8 – (5 points) Define Cloud Top Pressure. What band(s) are used to derive this variable? Where to we see the lowest cloud pressures? The highest? What relationship can you infer from the cloud fraction and cloud top pressure data sets? What type of cloud can you infer is found of the western coasts of continents based on your figures?

10 LAB REPORT WRITE UP

This is not a typical lab and will consist of a somewhat less formal write-up than others later in the semester. For this assignment provide the following in digital or paper format by next Friday, September 22nd, 2017:

1) (2.5 points) A brief summary of the labs main goal
2) (32.5 points) Answers to questions Q1 through Q8 with References
3) (6 points) Four individual figures with full global lat-lon properly labeld with titles, axes labels, and colobars
4) (4 points) A four panel figure of the four variables zoomed into your region of choice in one figure
5) (5 points) A copy of your code