

50 Points Total Available

Historical Biomass Burning in South America – DUE Friday October 20th, 2017

This lab will give you some experience with MODIS Aerosol Products including Aerosol Optical Depth and Fine Mode Fraction. You'll also practice making time series after making a regional average. Questions to be answered are labeled **Q1**, **Q2**, etc. and are highlighted in Red. More guidance will be provided as to how detailed a response I am expecting for each question or the specificity of the answer. Matlab coding on your own are highlighted in Green as **C1**, **C2**, etc.

1 DOWNLOADING & EXTRACTING MODIS L3 AOD and FMF

For this first section, similar to the cloud products lab, you'll be downloading L3 Aqua MODIS data files. For this section we'll be using data from each September on record from 2002-2016. This will give you 15 data files to download. We'll be making a time series so we need to have enough data points to say something about how things have changed over time. You search for each September by choosing the start time to be 9/1/xxxx and the end time to be 9/30/xxxx (xxxx = same year). Remember these files are large and will take time to download.

****You can download September 2002 first and work with this single file to set up the first section of code (regional averaging, standard deviation) while the other years download. Once it works for one we can automate the process to do the same thing to the other years.****

1) Download MODIS L3 Data

GET THESE L3 HDF FILES using LAADS Web: <https://ladsweb.modaps.eosdis.nasa.gov/search/>

- MODIS: Aqua → 6-MODIS Collection 6 → L3 Atmosphere Product [3] → MYD08_M3 MODIS/Aqua Aerosol Cloud Water Vapor Ozone Monthly L3 Global 1Deg CMG.
- Temporal Type: + Advanced (you just need to enter the first day of Sept to find the whole file)
 - 2002-09-01
 - 2003-09-01, Etc. (Add all of the years from 2002-2012)
- Location: World (Default if nothing else is selected)
- Files: Download each of the 11 files. This will take time since they are ~400 MB a piece!

1. September 2002 - MYD08_M3.A2002244.006.2015085143900.hdf
2. September 2003 - MYD08_M3.A2003244.006.2015085144635.hdf
3. September 2004 - MYD08_M3.A2004245.006.2015086030212.hdf
4. September 2005 - MYD08_M3.A2005244.006.2015086191822.hdf
5. September 2006 - MYD08_M3.A2006244.006.2015087144539.hdf
6. September 2007 - MYD08_M3.A2007244.006.2015089153000.hdf
7. September 2008 - MYD08_M3.A2008245.006.2015089152942.hdf
8. September 2009 - MYD08_M3.A2009244.006.2015089153027.hdf
9. September 2010 - MYD08_M3.A2010244.006.2015090164856.hdf
10. September 2011 - MYD08_M3.A2011244.006.2015090164906.hdf
11. September 2012 - MYD08_M3.A2012245.006.2015091194659.hdf
12. September 2013 - MYD08_M3.A2013244.006.2015091200253.hdf

Terra MODIS: **MOD06_L2.AYYYYDD,HHMM,CCC.YYYYDDHHMMSS.hdf**
 Aqua MODIS: **MYD06_L2.AYYYYDD,HHMM,CCC.YYYYDDHHMMSS.hdf**

Definition of highlighted text:

MOD06 = Earth Science Data Type name
L2 = Denotes a Level-2 product (or L3 for Level-3, etc.)
A = indicates following date/time information is for the acquisition (observation)
YYYYDD = acquisition year and day-of-year
HHMM = acquisition hour and minute start time
CCC = collection (e.g., '006' for Collection 6)
YYYYDDHHMMSS = production data and time
hdf = denotes HDF file format

- 13. September 2014 - MYD08_M3.A2014244.006.2015093012530.hdf
- 14. September 2015 - MYD08_M3.A2015244.006.2015280171756.hdf
- 15. September 2016 - MYD08_M3.A2016245.006.2016275201759.hdf

Q1 – (0.5 point) Looking at the file names for these 15 Septembers you’ll notice a variation in the Julian Day of acquisition. Why would some years have a start Julian day of 244 and others 245?

2) Extracting Combined “Dark Target Deep Blue” Aerosol at 550 nm for Land and Ocean and the Small Fraction of Aerosol over the Ocean (Note: Collection 6 no longer offers a “Fine Mode” over Land, ☹)

Starting with the September 2002 data make sure to check:

- Variable name – so you can call it from the HDF-EOS file
- Scaling factor – so you know if you need to scale the data while extracting it
- Offset – so you know if you need to offset the data while extracting it.
- Missing data -9999 change over to NaN
- Make sure you have a MODIS Latitude and Longitude file handy (you can use files from previous labs). If not extract and save that information again.
- For this lab set your colorbars from 0 to 0.5. CODE: `caxis ([0 0.5])`

Write your own code (starting from scratch or using pieces from previous codes) to extract both the Dark Target Deep Blue and Fine Mode over the Ocean. You can do both in the same script for the same year. We’ll also learn how to automate the code so it will step through each year and extract, scale, and offset the data automatically.

C1 – (5 points) ON YOUR OWN CODE to:

a) Extract 9/2002 Dark Target Deep Blue Aerosol Optical Depth (Land and Ocean) at 550 nm and save the data as a .csv file (Short Name = AOD_550_Dark_Target_Deep_Blue_Combined_Mean_Mean)

b) Extract 9/2002 Fine Mode Aerosol over the Ocean at 7 different wavelengths and save the data as a .csv file (Short Name = Aerosol_Optical_Depth_Small_Ocean_MeanM=_Mean).

You will need additional code that you have not seen before to deal with the Fine Mode data since has 3 dimensions 1 for the 7 wavelengths, 1 for latitude and 1 for longitude. It’s array size is: 7 x 180 x 360 ! You won’t be able to save or plot a 3D wave the same way as a 2D wave. Instead, we’ll extract JUST the wavelength that we’re interested in 550. Looking at the Metadata we see in the long name that 0.55 um (500nm) is the SECOND one in the list of wavelengths:

Retrieved optical depth for fine mode (avg solution) for 0.47 0.55 0.66 0.86 1.24 1.63 2.13 microns:
Minimum of Daily Mean

After you deal with the scaling factor, offset, and missing data you’ll need to add in this extra piece of code. After doing all of those things my variable is called “OCF_scaled_data” I need to tell it to go to the second (of 7) dimension to go to 550 nm and then take ALL of the latitude (180) and longitude (360). I can shorthand this by saying (2,:,:) since “:” means ALL the columns or rows. Next you need to “reshape” the array to make it only 180 x 360. I’ve renamed this array “OCF_final_data” but you can call it whatever you like.

```
OCF_scaled_data = OCF_scaled_data(2, :, :);
OCF_final_data = reshape(OCF_055_scaled_data, 180, 360);
```

c) Plot the global map of each to make sure that you have the correct orientation and that your -9999 removal, scaling, and offsetting worked. Add a coastline any way you choose. SAVE these two global maps to include in your lab write up.

****Hint: Make sure to load or extract the MODIS Latitude and Longitude. We will need those later to extract a small regional subset for further analysis.****

Q2 – (5 points) a) (1 pt) Based on the September 2002 map where are the ‘hotspots’ of AOD?

b) (1 pt) Based on the September 2002 map where are the ‘hotspots’ of FMF?

c) (1 pt) Where do ‘hotspots’ of AOD and FMF coincide? Based on the lecture materials speculate the aerosol type you are likely to find in these regions and provide a explanation for your choice.

e) (2 pts) What other MODIS variables could you use that we discussed in lecture that would help confirm your speculations in parts c) and d)?

3) Automating the extraction of the MODIS variables

Once you have successfully produced a map that has the correct orientation and is scaled properly we can start to modify the code to do the same thing to the other years you’ve downloaded automatically. By this time at least one other year should have downloaded. Now we can add in some key lines of code and make some simple modifications to help MatLab cycle through all the files. Make sure you put all your “September” files in the same folder, but make sure there aren’t any other MODIS data files from previous labs.

If we’re going to automate this process we need to know how many files we’re going to be working on. An easy way to do this is to see how many files in the directory start with a pre-determined string of characters. In our case all the files start with “MYD08_M3.A....” So, we can use the “**dir**” function to find out how many files start with this string by using the “*” to denote the variation in the file names due to the different months and years. We can then turn that structure (which stores a list of the filenames) into a number by determining the “**length**” of that structure. This gives us the “number of months” of data we’re going to cycle through. If you’re going to be using a lot of data files it’s sometimes handy to use a “**waitbar**” that let’s you see the progress of your data calculations. This is based on the number of files. You can add a message if you like.

```
%~~~~~  
% Finding the number of files in the directory  
MODISdir = dir('MYD08_M3.A*.hdf'); % Finds the needed files  
num_months = length(MODISdir); % Turns that information into a number  
MODIS_waiting = waitbar(0,'Please wait, working on a lot of MODIS files.');
```

Next we want to start the loop. We’ll need a few things including the number of files (num_months) and a counter, in this case “n” (though it can really be anything you want). We’ll also use the “**getfield**” function to find the full name of the file so we can extract information from it (specifically the year). See below

```

n = 1; % The counter index
while n <= length(MODISdir) % Start of the loop, will work while n <= the number of files
    name = getfield(MODISdir(n), 'name'); % Gets the name of the file
    disp(['working on: ', name])% Prints the name of the file you're working on

```

Next we want to automate saving our files. This is the right time to extract information from the files names (year) to help us differentiate the files we create. We can use the “name” that we derived above and identify the location in that text string where we find the year number (in our first example 2002). In the case of MODIS files it is characters 11-14. We can then create a text string that we want each file to have (say “MODIS_DTDB_” and “MODIS_OCF_” for MODIS Dark Target Deep Blue and MODIS Ocean Fine. Once we have a text string and the year information we can concatenate them into one filename using the “strcat” function.

```

%-----
% Getting out the year information from the filename
% File name format: MYD08_M3.A2002244.051.2008321123937
%
year = name(11:14);% Takes out the year 'YYYY' found 11-14 spaces from the first M
%
MODIS_DTDB = 'MODIS_DTDB_'; % Making a Text String
MODIS_OCF = 'MODIS_OCF_'; % Making a Text String
DTDB_filename_save_string = strcat(MODIS_DTDB, year); % stringing Name and Year
OCF_filename_save_string = strcat(MODIS_OCF, year); %stringing Name and Year

```

Once you have your file name string you can open the file (one by one). Instead of typing out the individual file names we can use the “name” variable that we created earlier. Instead of writing out the whole long MODIS file name you simply put “name” in it’s place. This let’s Matlab open each file without you having to type them in manually. See below:

```

DTDB_data = hdfread(name, ' AOD_550_Dark_Target_Deep_Blue_Combined_Mean_Mean');

```

Now that you can open the file you can copy and paste your previous code for double, scaling, offset, -9999 removal. After you insert your working code from before you’ll need to save the data using your pre-determined filename save string. Using “dlmwrite” we use slightly different syntax to instruct Matlab make a new file string each time through the loop (pulling out a new year each time). See example below.

```

dlmwrite([DTDB_filename_save_string, '.csv'], DTDB_scaled_data, 'delimiter', ',', 'precision', 15);
dlmwrite([OCF_filename_save_string, '.csv'], OCF_final_data, 'delimiter', ',', 'precision', 15);

```

* Remember you had to do one more step for the Ocean Fine Mode data, that’s why my data name is different than the DTDB data.

Finally, you’ll need to set the waitbar based on the number of files so it displays the fraction of files that have been completed. Then, you can set your counter to go to the next part of the loop by setting n (your counter) to n+1. That will tell Matlab to go on to the next file. Once you’ve closed the loop by adding “end” you’re done and you can close the waitbar.

```

waitbar(n/length(MODISdir));
    n = n+1;
end
close (MODIS_waiting)

```

This process will result in 15 files for DTDB and 15 files for OCF. You don't need to make plots of these data so don't include the figure creation code or just comment it out if you were adding to your original code. You should have 30 files now!

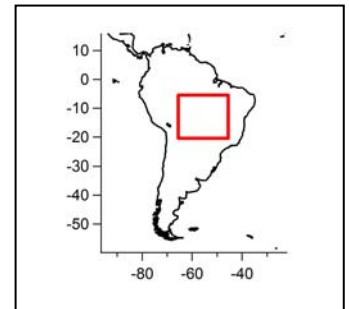
2 REGIONAL AVERAGE AND TIME SERIES

Just like in the previous step we want to test out our new techniques (regional average and setting up a file to save time series information) on a single file first. You have two files (for September) for each year, one for DTDB and one for OCF. We'll set up a method to isolate a region of interest in our data file and then calculate an average of that chosen box. We then want to store that data in a new file so we can plot the average as a function of year. You'll be using a similar "while loop" to cycle through all the files to calculate the regional average for each file and place them in the appropriate place in your time series file.

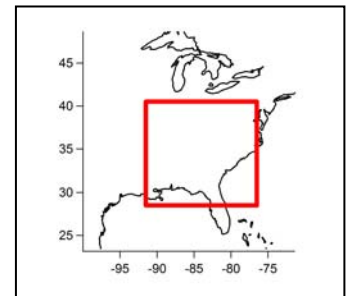
Let's get started by loading the necessary data files. You'll need your latitude and longitude files and the DTDB and the OCF data from September 2002. First we'll set up code for DTDB and then you'll add in new code to handle the OCF on your own.

1) First you need to know what 2 regions we'll be zooming in on.

First, we'll zoom in on Brazil. We'll be using a latitude range of 5.5°S to 20°S, and a longitude range of 65.5°W to 45.5°W. This corresponds to **-5.5° to -20° Latitude, -65.5° to -45.5° Longitude** in the MODIS latitude and longitude files which we'll need to search for.



Second, we'll zoom in on the Southeast US. The latitude and longitude for the Southeast US will be 28.5°N to 40.5°N, and a longitude range of 91.5°W to 76.5°W. This corresponds to **28.5° to 40.5° Latitude, -91.5° to -76.5° Longitude**



2) Regional and Time Series Code

We'll start his code same as always buy loading the latitude and longitude (use previous code you've made before) and then we'll include code to specify the boundaries for our two boxes. See below.

```

%-----
% Zoomed in Boxes of Interest: (Choose your own boundaries)
%
% To clarify the "Top" of the box will be the value that would be larger
% or more positive in the northern hemisphere (or less negative in
% the southern hemisphere). And the "Bottom" of the box will be the smaller,
% less positive value. For example...top = -0.5 bottom = -20.5 or top = 40 and
% bottom = 20.
%
% Boundaries for Zoomed in Box in South America
% finds the index location in the original lat file for the right latitudes
% finds the index location in the original lon file for the right longitudes

```

```

BRZ_box_zoom_top_center_lat = find(MODIS_Lat(:,1) == -5.5);
BRZ_box_zoom_bottom_center_lat = find(MODIS_Lat(:,1) == -20.5);
BRZ_box_zoom_left_center_lon = find(MODIS_Lon(:,1) == -65.5);
BRZ_box_zoom_right_center_lon = find(MODIS_Lon(:,1) == -45.5);
%
% Boundaries for Zoomed in Box in Southeast US
% finds the index location in the original lat file for the right latitudes
% finds the index location in the original lon file for the right longitudes
USA_box_zoom_top_center_lat = find(MODIS_Lat(:,1) == 40.5);
USA_box_zoom_bottom_center_lat = find(MODIS_Lat(:,1) == 28.5);
USA_box_zoom_left_center_lon = find(MODIS_Lon(:,1) == -91.5);
USA_box_zoom_right_center_lon = find(MODIS_Lon(:,1) == -76.5);
%
%~~~~~

```

Next we want to create empty arrays to store our regionally averaged data so we can make a time series. First we find the number of files using the “dir” function, this time searching for your DTDB file that end in .csv and only vary based on the year (use the * to do this). After we use “dir” we use the “length” function to get the number of files. Then, use the “zeros” function and the “num_months” value that I automatically determined using the “length” function to set the number of rows equal to the number of files (will be 15) and then set the number of columns so I’ll have one for the year, one for the mean, and one for the standard deviation (we’ll have 3 columns total). Once we have created the empty arrays we can create easy file name strings to use for later when we actually save the final time series file.

```

%~~~~~
% Finding the number of files in the directory
MODISdir = dir('MODIS_DTDB*.csv');
num_months = length(MODISdir);
%
% Making an empty array to store the year, mean, and standard deviation for a total of
three columns
%
BRZ_DTDB_timeseries_file = zeros(num_months, 3);           % for MODIS DTDB
USA_DTDB_timeseries_file = zeros(num_months, 3);           % for MODIS DTDB
%
% FILE NAME STRING
MODIS_DTDB_BRZ_Save_String = 'All_Years_MODIS_DTDB_BRZ';   % Save String
MODIS_DTDB_USA_Save_String = 'All_Years_MODIS_DTDB_USA';   % Save String
%
%~~~~~

```

Now we get to start our loop and, like before, we want to use a counter (n = 1), the number of files in the directory and use the “getfield” function so we can take the “year” out of the file name to use later when we save data for each year into our time series file.

```

n = 1;
while n <= length(MODISdir)
    name = getfield(MODISdir(n), 'name');
    disp(['working on: ', name])
    %~~~~~
    % Getting out the year and month information from the filename
    % File name format: MODIS_DTDB_2002
    % (If your file name is different you have to “count in” to get to your YEAR
    % information).
    %
    yearyear=name(12:15); % Takes out the year 'YYYY'
    %
    year=str2num(yearyear); % turns the string (text) year into a number value
    %~~~~~

```

Once we start the loop we can open the file and start taking out our “boxes” for Brazil and the USA. I’ll show an example below for Brazil. You’ll need to add code to do the USA calculations for C2. We start by reading in the data. Since we’ve automated the loop we won’t have to type in the individual names. We can simply use the “name” variable. But, because we’re using .csv files you’ll need to identify the delimiter (the purple ‘,’) and then tell the code to start at row 1 and column 1 (which is denoted as 0,0 in the dlmread function)

```
% Reading the data
MODIS_DTDB_data = dlmread(name, ',', 0, 0);
%~~~~~
```

Once the data file has been opened we can find our data boxes. The most difficult part of taking excerpts of data out of a larger file is knowing that you have actually obtained the correct box. If you want to take out a subset of data you need to specify the rows and columns (Rows, Columns). Matlab likes to see it in the format:

EXCERPT = Original_Data(**Row_A: Row_Z, Col_A : Col_Z**)

Where the ‘:’ means from some row number to another. In our case we’ll be looking at something like this:

EXCERPT = Original Data(Top_Lat_Index : Bottom_Lat_Index, Left_Lon_Index : Right_Lon_Index)

```
%~~~~~
% Extracting Boxes and Taking Means for the BRZ Box
%
% BRZ Excerpts for MODIS Dark Target Deep Blue Optical Depth
MODIS_DTDB_BRZ_excerpt =
MODIS_DTDB_data(box_BRZ_top_center_lat:box_BRZ_bottom_center_lat,
box_BRZ_left_center_lon:box_BRZ_right_center_lon);
%
% Finding the non-zero data that you want to use to calculate the mean and standard dev.
MODIS_DTDB_BRZ_excerpt_pos = find(MODIS_DTDB_BRZ_excerpt > 0);
% using a subset (the positive values) of the subset (zoomed in to Brazil)
% Here is where you use “mean” and “std” to calculate the mean and standard dev. Notice
that you’re
MODIS_DTDB_BRZ_avg = mean(MODIS_DTDB_BRZ_excerpt(MODIS_DTDB_BRZ_excerpt_pos));
MODIS_DTDB_BRZ_std = std(MODIS_DTDB_BRZ_excerpt(MODIS_DTDB_BRZ_excerpt_pos));
%
% Saving the Variables
% You can save specific values based on year(row “n” that changes per loop) and
specifying the column (column 2 = mean, column 3 = standard deviation).
MODIS_DTDB_BRZ_timeseries_file(n,1) = year;
MODIS_DTDB_BRZ_timeseries_file(n,2) = MODIS_DTDB_BRZ_avg;
MODIS_DTDB_BRZ_timeseries_file(n,3) = MODIS_DTDB_BRZ_std;
%~~~~~
n = n+1;
end
```

```
dlmwrite([MODIS_DTDB_BRZ_Save_String, '.csv'], MODIS_DTDB_BRZ_timeseries_file, 'delimiter',
',', 'precision', 15);
```

You can test the code with just the Brazil data and then add in the USA. Make sure you end up with a table of numbers with 15 rows and 3 columns. Make sure that column 1 is the year, increasing from 2002-2016.

C2 – (20 points) ON YOUR OWN CODE:

a) Add in code to extract the Southeast USA data box and calculate the mean and standard deviations for 2002-2016. NOTE: you cannot make a timeseries of OCF for Brazil since there is NO land data for it, so you don't have to do the OCF for the USA either.

b) Make a second code (or add to your current code) to produce figures using all the data files you have produced thus far.

i) You will produce one figure at this stage with THREE panels. The TOP panel will show Brazil DTDB (while adding errorbars, using the "errorbar" function in Matlab)) the MIDDLE panel will show the Southeast USA DTDB data with error bars. You may need to use the help files to get this right, the BOTTOM panel will show the Southeast USA OCF Data with errorbars.

To make a Multi-Panel Figure you use the syntax below. You start the same with "figure" and then you use the function "subplot" specifying how many panels (3), columns (1), rows (1, 2, 3). You copy your figure code into the space that is noted by "plot(x,y1)" etc.

```
figure
subplot(3,1,1);
plot(x,y1)
```

```
subplot(3,1,2);
plot(x,y2)
```

```
subplot(3,1,3);
plot(x,y1)
```

c) Use your 3 data files (BRZ DTDB, USA DTDB, and USA OCF) to make a large table that includes 4 columns (first one for the year).

i) Arrange the columns showing the mean +/- std for BRZ DTDB, USA DTDB, and USA OCF so it is easy to compare region to region.

ii) Highlight/Bold the MAXIMUM Value for each of the three data sets

Q3 – (10 points) Evaluation of Your Time Series Figures

a) (1 pt) Looking at the BRZ AOD time series (and the table) which year has the highest overall AOD? Which year has the lowest?

b) (1 pt) For Brazil DTDB, how as AOD changed over time, if you were to draw a trendline would AOD be increasing, decreasing or staying the same between 2002-2014?

c) (2 pts) Speculate what may cause variations in the amount of AOD in the Brazil region. (Open ended... I just want to see what explanations you can develop for this trend).

d) (2 pts) Look at the Southeastern USA DTDB time series. Does that maximum DTDB happen in the same year as in Brazil? How does the maximum value in the SE USA compare to BRZ (compute a ratio of USA to Brazil)?

e) (2 pts) Does the SE USA display the same trend as BRZ over time? Speculate why these regions may or may not have different trends in DTDB (Open ended, I want to see what you're thinking).

f) (2 pts) What type of aerosol do you think is present in Brazil? In the Southeastern US (Open Ended)?

Q4 – (3 points) Aerosol Type and MODIS Data Indications

a) (1 pt) What MODIS variables (and expected values) would I use to indicate the presence of SMOKE?

b) (1 pt) What MODIS variables (and expected values) would I use to indicate the presence of POLLUTION?

c) (1 pt) What MODIS variables (and expected values) would I use to indicate the presence of DUST?

7 LAB REPORT WRITE UP

Provide the following in digital or paper format by Friday, March 6th, 2015:

- 1) (4.5 points) A brief summary/conclusion of the labs main goal and breif discussion about the dark target alorithm that is used for this assignment. Also, describe the different types of aerosols that you are likely to encounter in these chosen regions and how seasonality plays a role in global AOD values.
- 2) (25 points) Coding question C1 and C2, Figures and Table
- 3) (18.5 points) – Science/Interpretation Question Q1, Q2, Q3, and Q4
- 4) (2 points) – emailing code