User’s Manual For “ProcessCODData” Product 1.0

Author: Stephen M Forczyk

Oct 2,2023

*Contents*

[Figures 4](#_Toc145150184)

[1.0 Purpose 5](#_Toc145150185)

[2.0 Basic Script Capabilities 5](#_Toc145150186)

[2.1 Tool Set Description 7](#_Toc145150187)

[2.2 Hardware Requirements 7](#_Toc145150188)

[3.0 Merra 2 Data Product 7](#_Toc145150189)

[3.1 Where to get the desired Merra 2 dataset 7](#_Toc145150190)

[10](#_Toc145150191)

[3.2 Direct Download Method 10](#_Toc145150192)

[3.3 Automated download 18](#_Toc145150198)

[3.4 Subsetting the Data Files 20](#_Toc145150199)

[4.0 Setup Of ProcessCODData Scripts package 24](#_Toc145150200)

[4.2 User Setup of Dropbox Files on Local Computer 26](#_Toc145150201)

[5.0 Running the Code Package 29](#_Toc145150202)

[5.1 Movie Creation 34](#_Toc145150203)

[5.2 Required Toolboxes 34](#_Toc145150204)

[5.3 Known Code Issues 34](#_Toc145150205)

[6.0 References 35](#_Toc145150206)

## Figures

[Figure 1 Required Matlab Routines 5](#_Toc145148751)

[Figure 2 Earth Data Login Page 7](#_Toc145148752)

[Figure 3 Earth Data Landing Page For TAUHGH Data 8](#_Toc145148753)

[Figure 4 Dataset Reference Material 9](#_Toc145148754)

[Figure 5 Data Collection Calendar For M2TMNXRAD 9](#_Toc145148755)

[Figure 6 Data Access UI Item 10](#_Toc145148756)

[Figure 7 Start Point For Merra2 Data Download 10](#_Toc145148757)

[Figure 8 Select Date Range 11](#_Toc145148758)

[Figure 9 Revised Data Selection Dates 12](#_Toc145148759)

[Figure 10 Data File Download List 13](#_Toc145148760)

[Figure 11 Result Of Manua File Download 13](#_Toc145148761)

[Figure 12 Panoply Link 14](#_Toc145148762)

[Figure 13 Panoply View Of Downloaded File 15](#_Toc145148763)

[Figure 14 Creating a Plot Of TAUHGH 16](#_Toc145148764)

[Figure 15 Sample TAUHGH Plot in Panoply 17](#_Toc145148765)

[Figure 16 Download File Links 18](#_Toc145148766)

[Figure 17 Downloaded File Links 18](#_Toc145148767)

[Figure 18 How To Get a File Subset 20](#_Toc145148768)

[Figure 19 Selection a Region Subset 21](#_Toc145148769)

[Figure 20 Data Variable Selection 22](#_Toc145148770)

[Figure 21 Panoply File Check 23](#_Toc145148771)

[Figure 22 Dropbox Storage Set Up 24](#_Toc145148772)

[Figure 23 Project Folder Set Up 24](#_Toc145148773)

[Figure 24 Setting Up Local Paths 26](#_Toc145148774)

[Figure 25 Setting SAP Values 27](#_Toc145148775)

[Figure 26 File Selection Dialog 28](#_Toc145148776)

[Figure 27 Code to Produce Dialog 29](#_Toc145148777)

[Figure 28 Example TAUHGH Polar Distribution Jul 1990 30](#_Toc145148778)

[Figure 29 TAUHGH Variations 1980-2023 31](#_Toc145148779)

[Figure 30 Code To Create TAUHGHTT Table 32](#_Toc145148780)

[Figure 31 1 Year Section of TAUHGHTT Time Table 32](#_Toc145148781)

# Purpose

This document is intended to serve as a user’s manual for the release of the first version of a set of Matlab scripts to process a very specific dataset of Merra 2 data. This set of scripts will be collectively referred to as “**Merra2 COD Processing Scripts**”. This is not intended to be a general purpose script but

was designed to accomplish a very specific and limited task. (Saha et al., 2023)The software was developed using version 2022b of Matlab. It is intended for free public use but the author requests that users acknowledge the use if this code in any publications or use of any of the routines in other software packages.

The Merra 2 program stand for **M**odern **E**ra **R**etrospective **A**nalysis for **R**esearch and **A**pplications Version 2. This is a large and longstanding program designed to gather sensor measurements from a wide variety of sources and assimilate them into a common grid spanning the earth. The overarching purpose of this effort is to assist in building global models to predict future weather and climate data. Many references are available to provide detailed explanation of this complex program which has been running for more than 40 years.(Gelaro et al., 2017) The citation to the left is a good start point. The actual document can be viewed by clicking the reference which is also a link to the paper.

This specific small set of scripts was designed to extract the TAUHGH parameter from the data files in order to facilitate the study of high-level cirrus clouds.

# Basic Script Capabilities

The script set is rather small as it was designed for a specific rather than the more general purpose of reading a wide variety of netCDF formatted data than many Merra 2 products are provided to the user. The script are meant to perform the following tasks.

* An executive routine to run the overall process (ProcessCODData.m)
* A decoder routine to read the netCDF datafile (ReadCODDataset.m)
* A routine to create a timeseries for plot and display purposes (ReadCODDataset.m)
* Time series plot routine (PlotTAUHGHTable.m)
* A function to plot the individual frames of data on a map centered on the North Pole
* A screengrab routine to create movies over time if desired
* Finally some code to export the data to an Excel workbook

These will be discussed in greater detail in this section. The actual routines included are show in the following chart. The list below shows 25 items but the last 6 items below are required jpeg images and or not really code at all. Over all, the summary shows the product has about 4800 LOC (lines Of Code) but this is actually an over estimate because some routines are not active for this product at this time.



Figure Required Matlab Routines

Note that items 19 thru 25 in the list above are jpeg files. These do not have lines of code but are included in the list of required files as the scripts will use them if the PDF option is implemented at a later date . The executive script for the set of routines is item 9 or “ProcessCODData.m” .

With this information and an Earth Data account the user should be able to download a wide varieties of files from this data source. There a wide variety of data products from a number of different sources on this archive and not all share the same download steps or event the same file format (netCDF). This section should serve as a good start point.

2.1 Tool Set Description

This software package has been developed and run on Matlab release version from 2022b . The package was deliberately developed using release 2022a so as to make it more widely accessible to users who have not upgraded to the latest versions. Some tests were made on an earlier release(R2017b) that showed some issues regarding the use of timeseries. These are handy features and hopefully will be made operational in a future release, Regarding Mathworks Toolboxes the Basic Matlab and the Mapping Toolbox are required.

## 2.2 Hardware Requirements

The software is best executed with a fast Windows computer with 16 Gb of memory or more. The mapping functions can be quite computationally intensive. The author’s machine has 64 Gb or RAM and an array of Terabyte drives to store the required datasets. This software pack was modified to produce a Noth Polar plot which is less computationally intensive than more generalized map applications.

# Merra 2 Data Product

This section will describe the type of Merra 2 datafile required to include what the file contains and how to download them. Specifically, this software package is designed to deal with a single type of data output and even then for just one main variable from this file. Because of this the netCDF files are relatively small in size are can be decoded very quickly. The specific Merra 2 data product is M2IUNXASM\_5.12.4. Each product has a “landing page” which the user can think of as a start point for obtaining and working with the desired dataset.

Note that the Merra2 project contains many different types of filesets. (*Bosilovich785.Pdf*, n.d.)

## 3.1 Where to get the desired Merra 2 dataset

In order to get Merra 2 data of any kind the user must have an Earth Data Account which is free. The next graphic shows the Earth Data login page. The entry point for earth data is <http://urs.earthdata.nasa.gov/> . After pressing the get data button, the user will eventually encounter the login screen shown below. If the user does not already have an account one can be set up from this page.

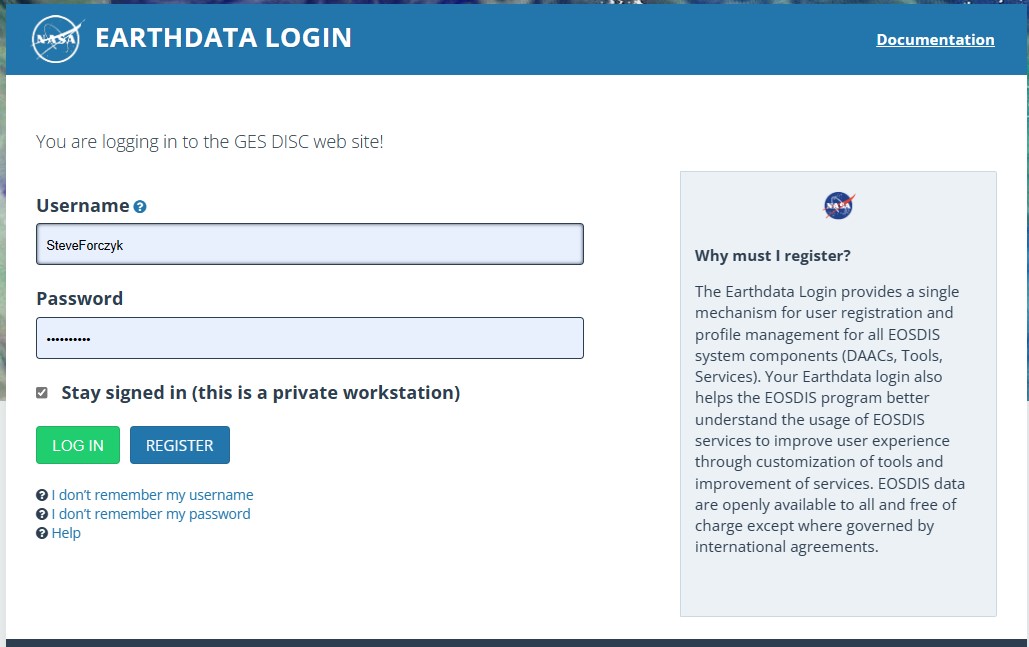


Figure Earth Data Login Page

When the user has navigated this step then the user can go to the desired landing page if this is known. If not search functions are available to assist. In this instance the landing page was stored in the Internet Favorites and can be seen below.

Figure 3 below has a wealth of useful information and is a great resource to start using this dataset.

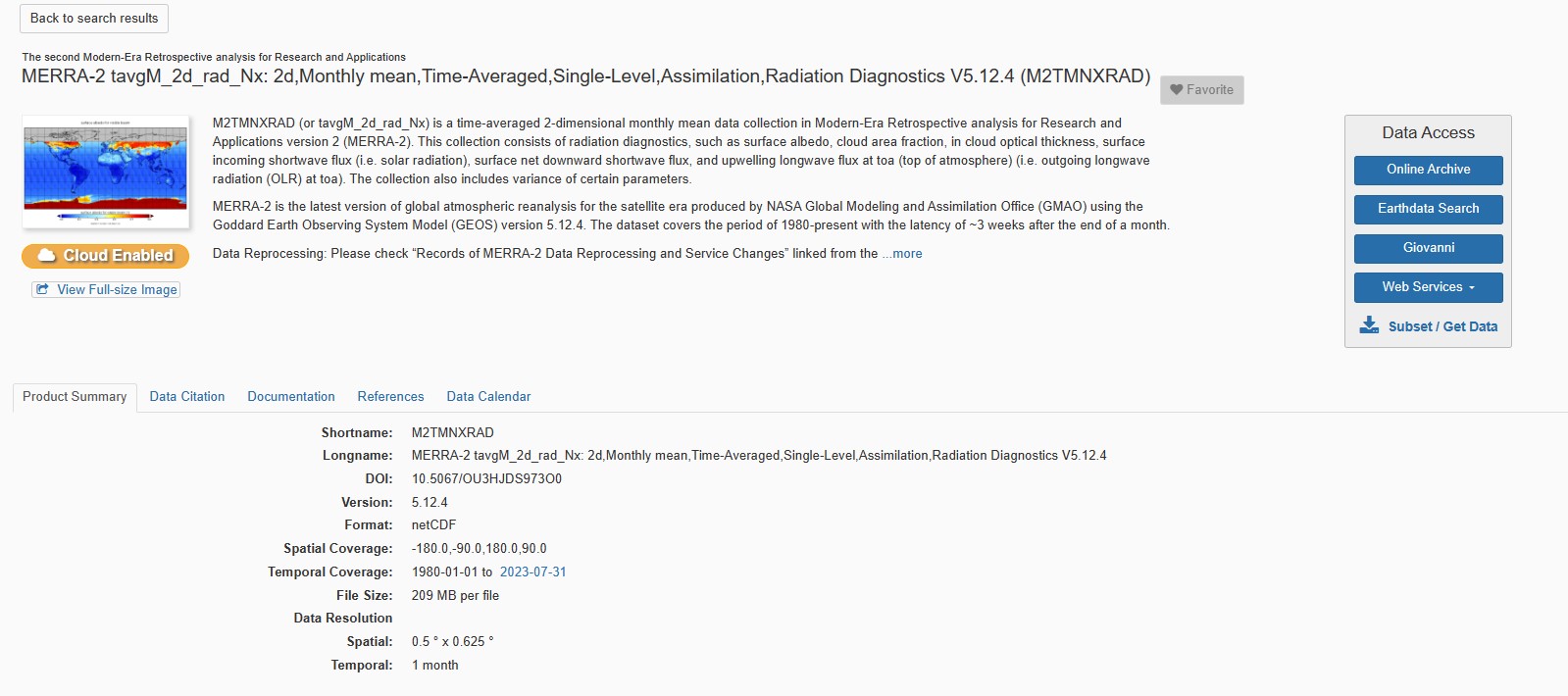


Figure Earth Data Landing Page For TAUHGH Data

The top of the page informs the reader that this dataset has the official name of “MERRA-2 tavgM\_2d\_rad\_Nx: 2d,Monthly mean, Time-Averaged ,Single-Level, Assimilation, Radiation Diagnostics V5.12.4 (M2TMNXRAD)” . The pneumonic is the closing parenthesis is the shorthand way of reference to this dataset. This description tells the user that the quantities in this dataset are 2 dimension (lat, lon) averaged monthly means. Single level assimilation references the fact that pressure level is not separately averaged or separated out. Many data sets contain data taken at up to 72 pressure levels ranging from the top of the atmosphere (0 mBar) to the ground level (1000 mBar).

Next the reader should be sure to note the wealth of reference materials available for this dataset which can be seen on the tabs at the lower left portion of the landing page. This can be seen in Figure 4 on the next page where a blue arrow highlight 5 tabs.

Tab 3 has basic data items while Tab 4 points to more specific technical memos and papers published that are related to this dataset. Another useful item is the Data Collection calendar as it relates to this dataset. The dataset calendar is available by year, through the time period that this data collection is being added to. This is displayed on Figure 5. Inspection of the chart shows that this data collection was available on every day of 2022. This is not as surprising as I may seem. First off the data represents a monthly average. Secondly ,the data is drawn from a number of sources so the lost of a single source for a day to two is not generally a big problem.

## 

Figure Dataset Reference Material

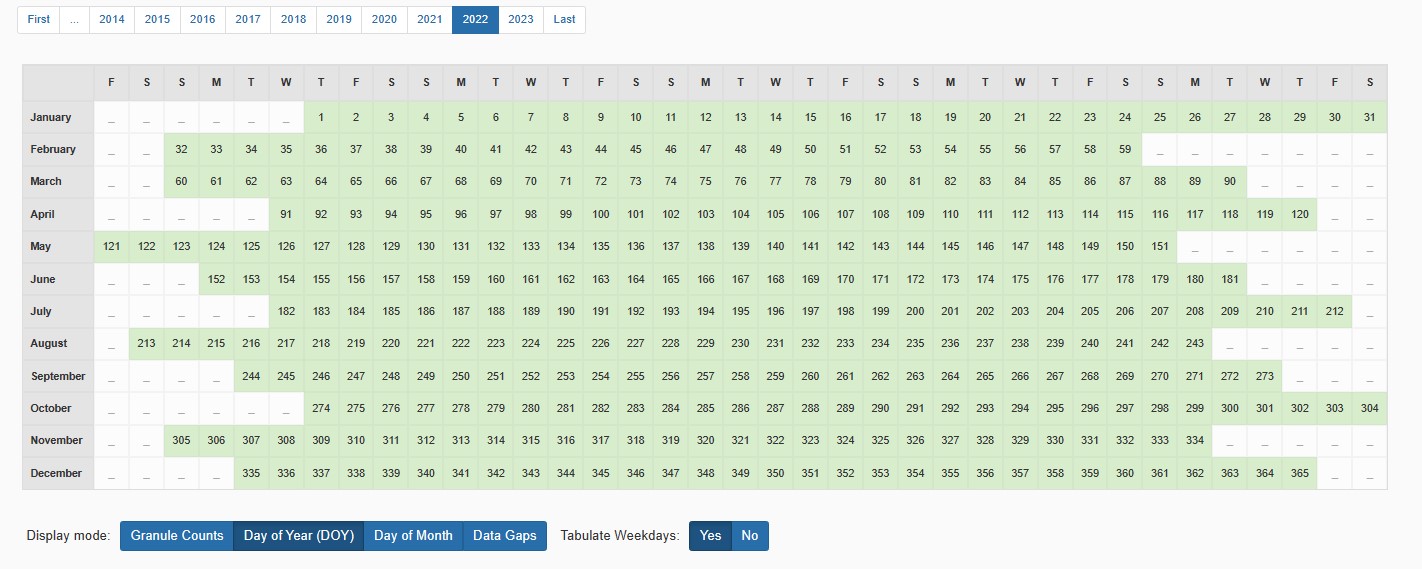


Figure Data Collection Calendar For M2TMNXRAD

Now the small box on the right side of Figure 3 is the key to downloading data from this dataset. The next graphic blows up this dialog box so the details can be seen. There is a lot of functionality covered by this user interface item. There are really 5 items on this interface as displayed in Figure 6. The first is Online Archive and the last is “Subset/Get Data”. Generally, it is easiest to work with this last item.

## 3.2 Direct Download Method

This section will highlight the direct download method from getting files from the data archives.

## 

Figure Data Access UI Item

## If the user selects the Get Data option the next charts shows the result.

## 

Figure Start Point For Merra2 Data Download

The figure above shows the result of pushing the get data button. The Get Original Files method is just fine and these files are in netCDF format (Common Data Format) is widely used. The next step is to define a date range. Do this by clicking on the Refine Date Range control and the following GUI appears.

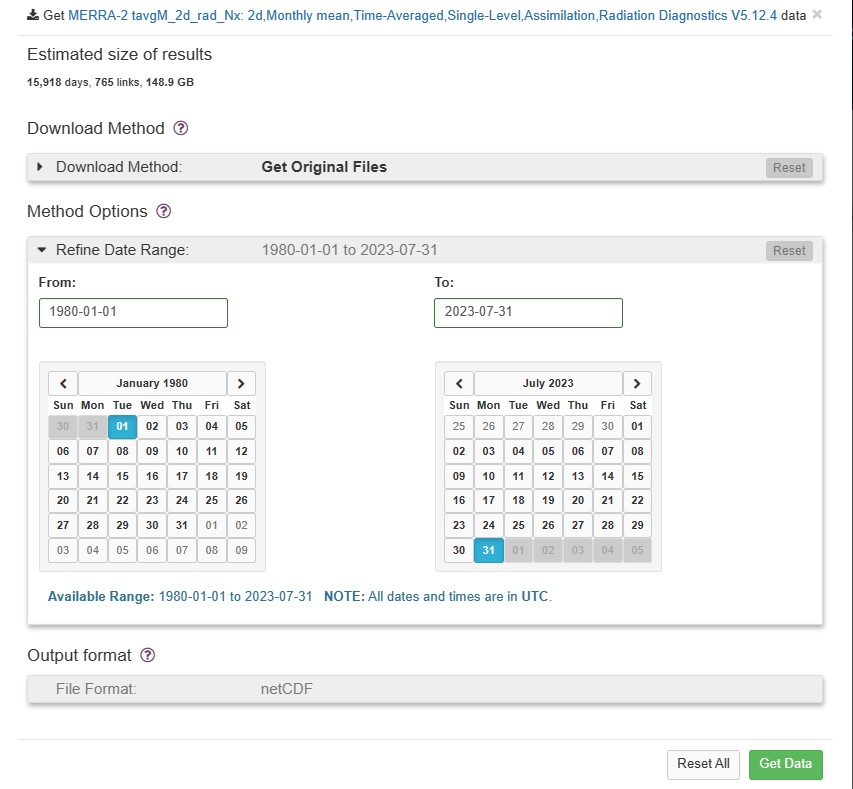


Figure Select Date Range

Note that the default date range spans the entirety of the data collection period for this product. This would not be a good option to choose in one step! Instead let’s choose a smaller demonstration set. This is illustrated in the next graphic Figure 9 . Inspection of this figure shows that the user has asked the web site to all the files for just 1990. Since the files are created at a monthly interval this means that the request is for 12 files. Note that the download method chosen to download the original files. This means that the complete files with all variables are being demanded. Other options are available which will be discussed shortly.

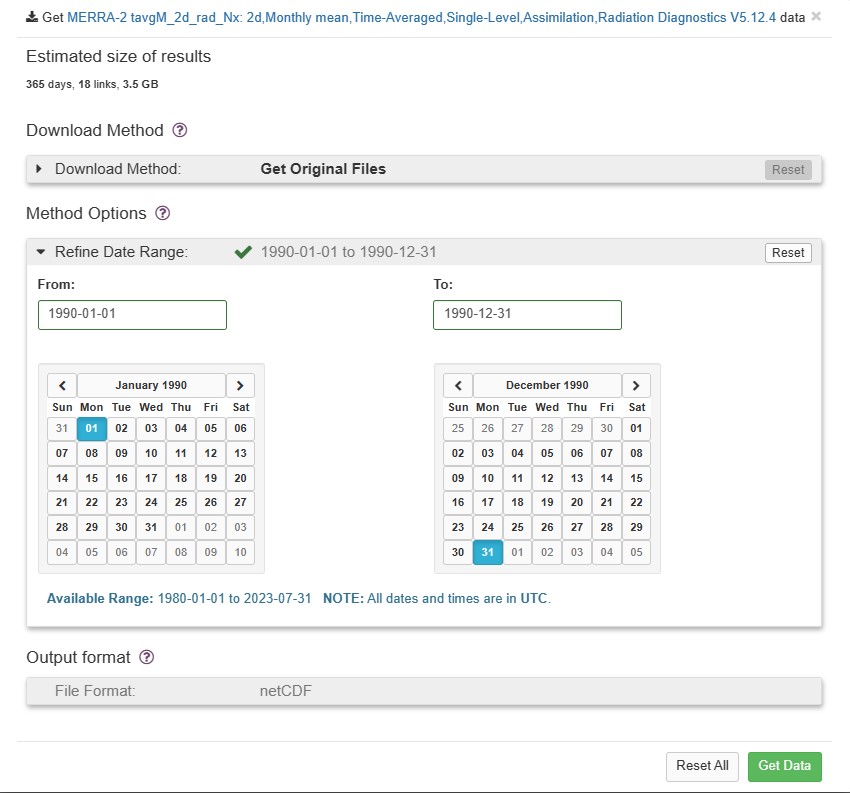


Figure Revised Data Selection Dates

If the user now presses the Get Data button the following UI will be presented to the user Figure 10.

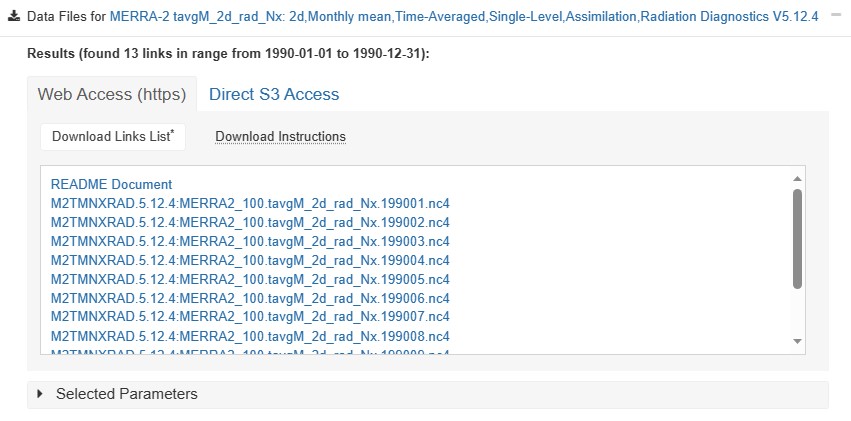


Figure Data File Download List

This is a scrollable list. To download these files the user should manually click on each file one at a time after waiting a few seconds after each click. Once this is done the data download will begin. The user should not try to download more than about 5 files at a time. This means click the first few files on the list and when the download is complete click the next batch of files. This can be a bit time consuming but it is the simplest way to proceed.

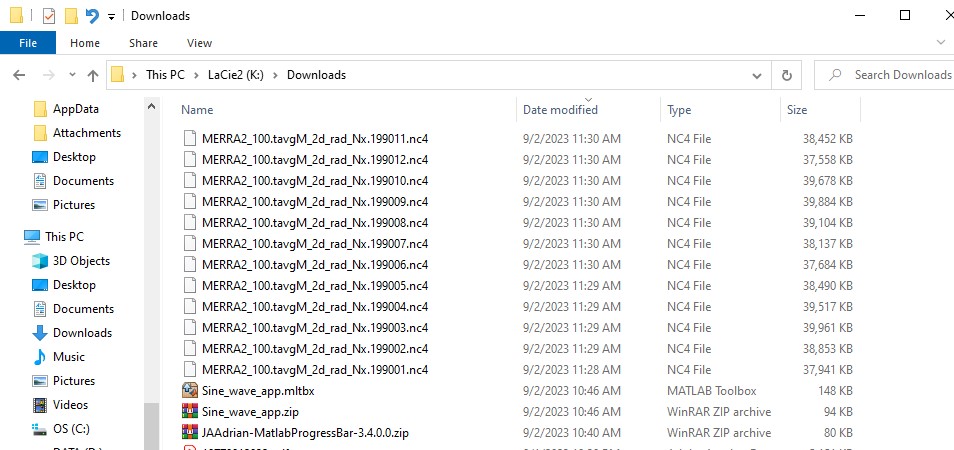


Figure Result Of Manua File Download

The chart above shows the download folder after all 12 files requested have been downloaded. Each of the netCDF files ends in a ‘.nc4’ file extension. The decoder can work with anything with a ‘.nc” extension. Each file is about 37 Mbytes long.

Some tips on this process is to set the default download folder location to somewhere other than the C drive. Experience has shown that downloading large datasets onto the C drive will greatly fragment the disk which can slow up all program operations. A better choice it to use another internal or external drive to avoid this issue.

So, what does one of these files contain? A very good way to see this, and to quickly visualize selected variable the “go to” choice is the use **Panoply** software which is freely available from NASA. Click this link [NASA GISS: Panoply Data Viewer: Downloads](https://www.giss.nasa.gov/tools/panoply/download/). Doing this will bring up the following web page.

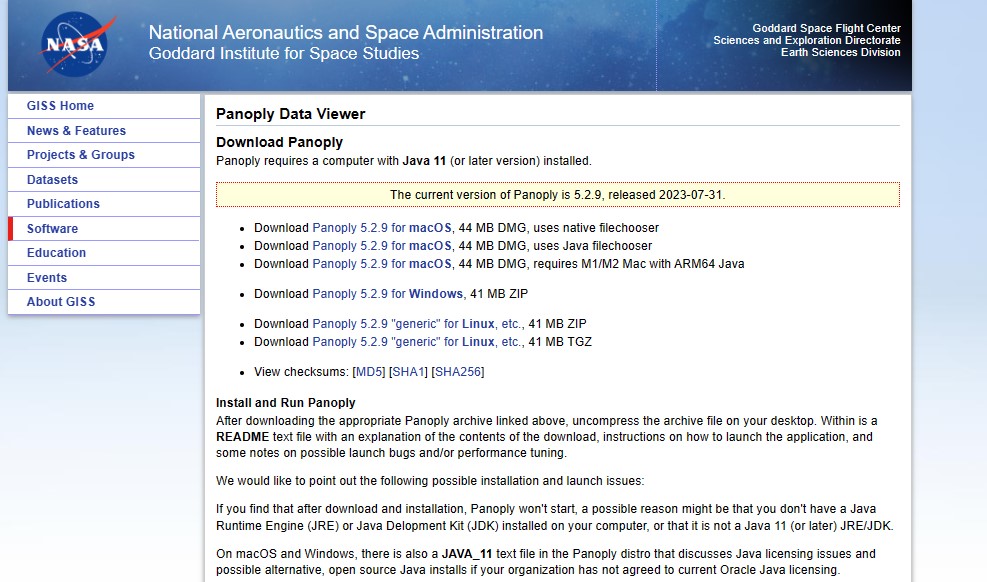


Figure Panoply Link

## The user should follow the instructions to download this important piece of software.

## Now the user clicks on the Panoply software link (Java coffee cup on the start bar) and when the app opens up selected the desired file. The next figure shows what happens when one of the just downloaded files is selected.

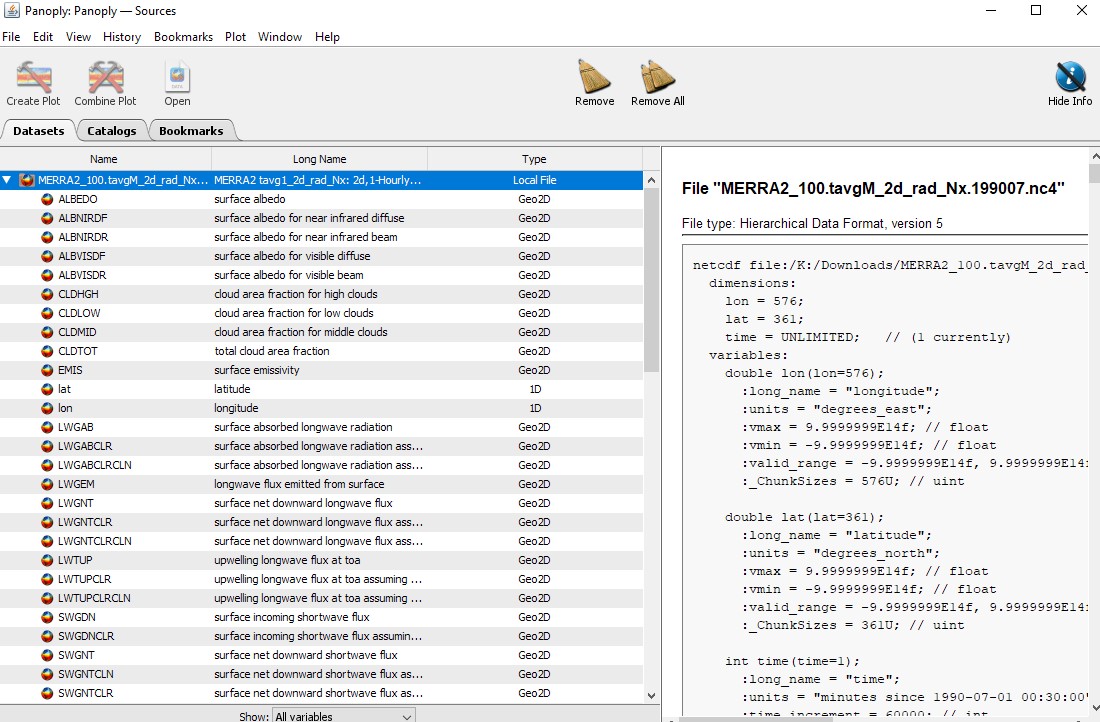


Figure Panoply View Of Downloaded File

This figure shows a typical view presented to the user once a selected netCDF has been opening. The left panel shows an opened view of the file where each variable is shown in a list. This list is scrollable so only some of the variables are shown. The right panel provides key data for the file or a variable that is selected. This information was used in writing the Matlab decoder routine “ReadCODDataset.m” .Note that the full file contains over 50 variables. The decoder provided in this code package only decodes 3 variables, namely TAUHGH, lat and lon. This greatly reduced the size of the files to be stored. In addition, the geographic area of interest was subset of the whole Merra grid which is 576 x 361 . By limiting the latitude to 55 to 90 deg the new subgrid is 576 x 51 . In fact, the final file size is about 103 Kbytes vs 38 Mbytes for the full up data collection.

If the user presses a single variable a content menu comes up which allows the user to see a 2 D plot on a map of the desired variable assuming it is a 2 D georeferenced array. To start the process, refer to the next chart Figure 14.

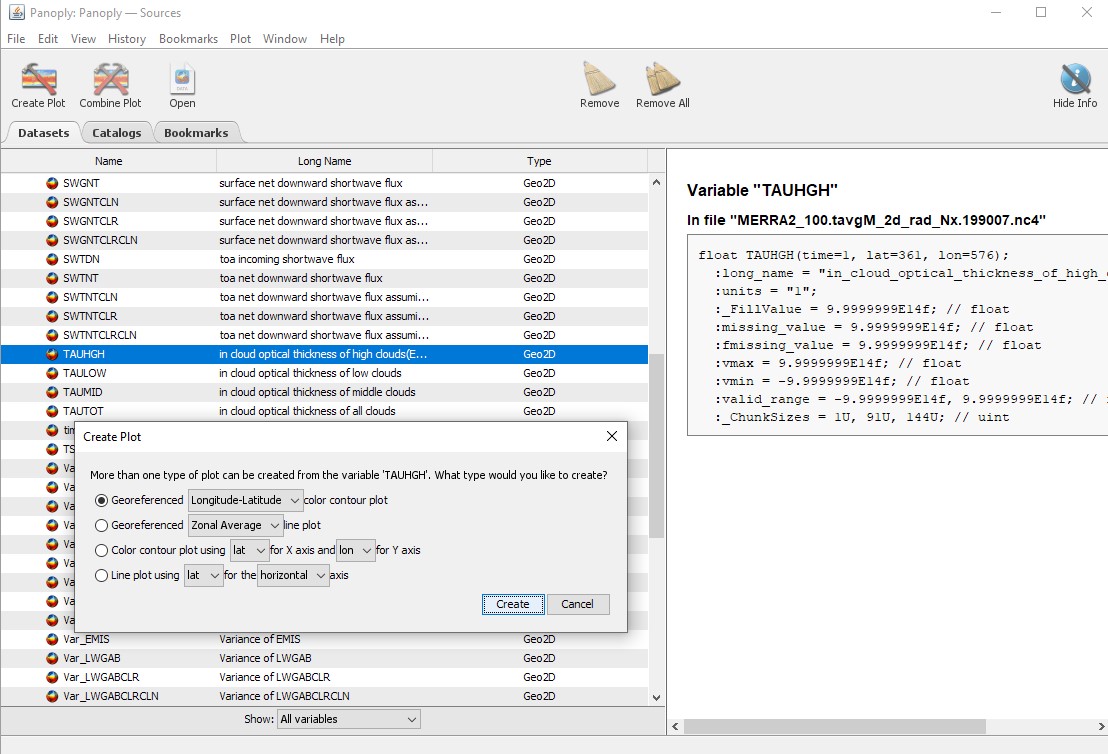


Figure Creating a Plot Of TAUHGH

TAUHGH was chosen as the plot value because this was the desired variable for study for the project that Matlab software package was created for. This variable is georeferenced as all the Merra2 variables for this data product were assimilated to a 576 x 361 grid. There are 576 grid point along the longitude direction at an interval of 0.625 deg while there are 361 latitude points at a .5 deg interval. The longitude grid goes from -180 Deg to 180 deg while latitude spans -90 to 90 deg.

After the user selects the georeferenced plot Panoply creates the following display Figure 15.

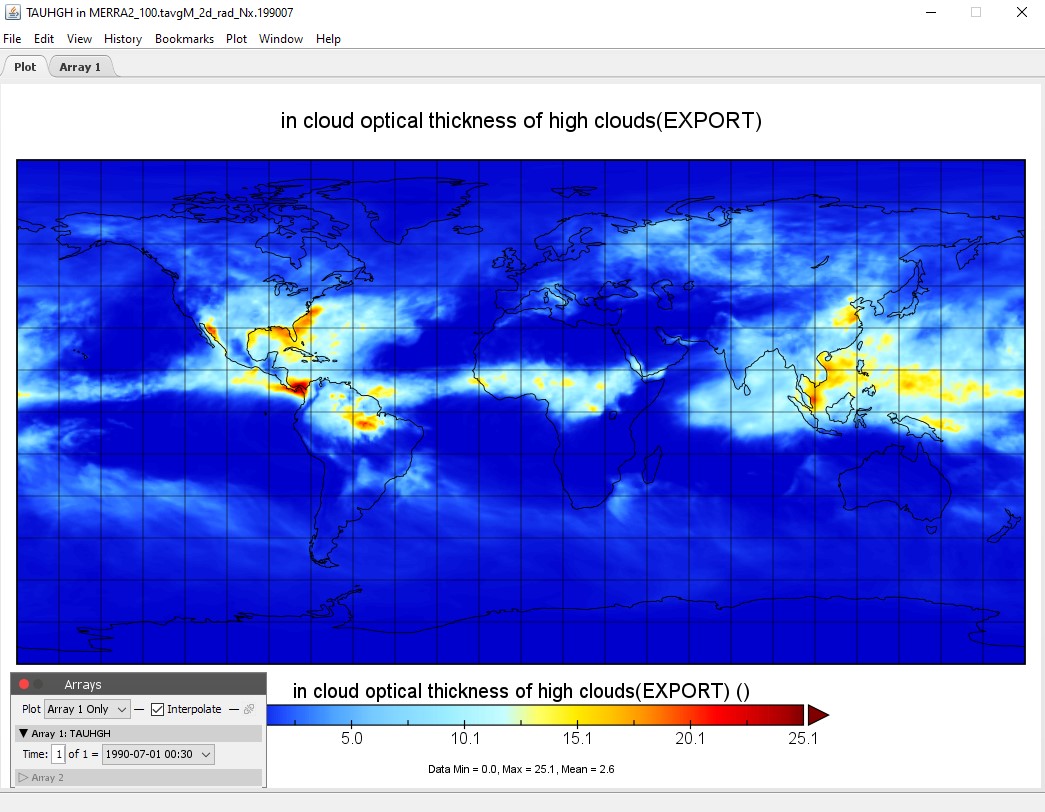


Figure Sample TAUHGH Plot in Panoply

In order to enhance the instructional content, the inset box was moved to the screen capture area. This has data on the array being plotted. Some files have the variables over various time intervals over 24 hours. Popular time spacings are 1, 3 and 6 hour time intervals. The data product that is the subject of this Matlab code is a monthly average so the concept of specific hour intervals has no real meaning.

## 3.3 Automated download

There is another technique that can be used to download files for the archives. This involves the **wget** command used in the command window. This is a very handy tool especially to handle large or length downloads. It is fairly simple but not as intuitive if the user has never used **wget**. This technique starts out just like the manual technique of the previous section until the user gets to the step shown in Figure 10. At this point there is another path that can be taken. The next figure shows what happens if the user takes the following action.

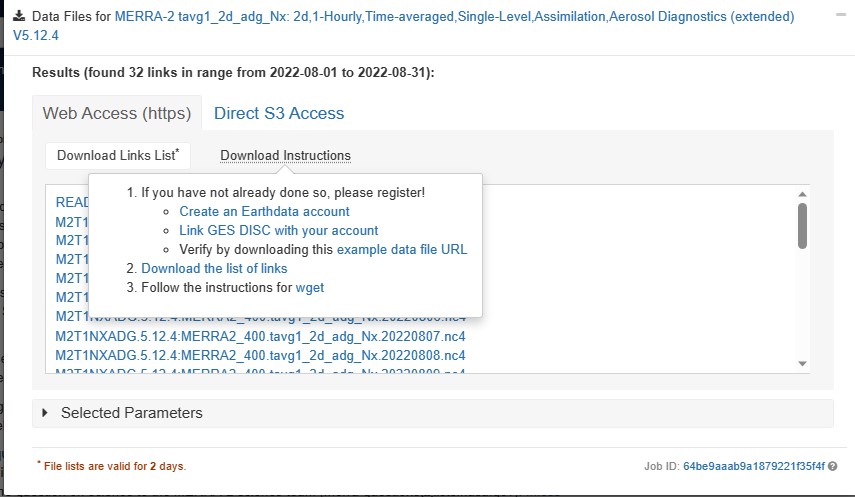


Figure Download File Links

What is different in this example is that the user hovers the cursor of the download instructions. Assuming the user has an Earthdata account look at step 2 which will download a list of file links-not the files themselves.

This will be sent to a text file which will look something like this next chart.

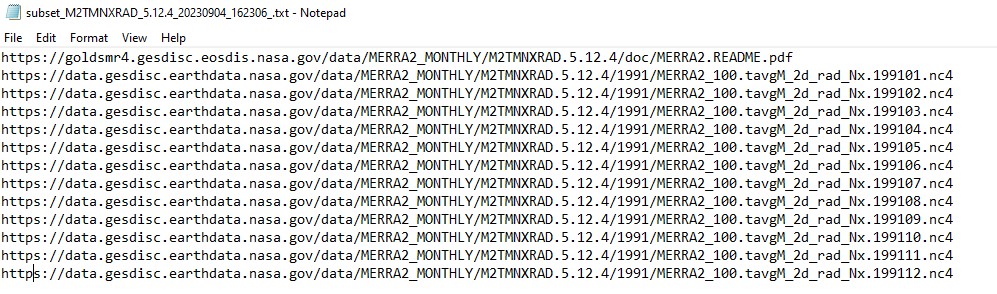


Figure Downloaded File Links

In this example case the are links to 12 monthly files for the year 1991 and a readme file. The readme file is not really needed so the first item on this list can be deleted. The user should save the remaining items into a simple text file say “**M2TMNXRAD\_Data1991.txt**” .

Next the user needs to open up command window in the folder where he would like the downloaded files to appear. Save the text file to this same folder. Then only one line needs to be typed into the command window at the prompt. This single line of code is directly below.

wget --http-user=SteveForczyk --http-password=XXXXXXX -i M2TMNXRAD\_Data1991.txt

The username must appear in the command as well as the account password which is hidden in this example. Execute the line of code and the download will begin. Note the entries in the text file are the addresses where the desired data has been stored for the user. This is a kind of holding area which will only be maintained for 2 days after creation so the download should be initiated without delay. Once the download has been completed the script will terminate.

Note the name assigned to the downloaded file may have many extraneous text add on to the file name but the user should still be able to recognize the date and time period the file name refers to. After the download is complete the user will likely have to write a short script to rename the files to something easier to work with. The choice I made was to name the file as YYYYMM.nc4 . Included in the source code package is a script that was used to perform such a name switch. The file is “RenameCODFiles.m” .Note that a one size fit approach to this process is not recommended. This is because the actual name of the downloaded file is highly dependent on any user selection made that modifies the initial file name.

## 3.4 Subsetting the Data Files

Regardless of the method used to download the files, there is another step that can be performed to reduce the content of the file to be downloaded. For this project it was desired to reduce the disk space storage by limiting the geographic cover of the file and to only download the TAUHGH variable. This section shows how it can be done.

This next figure refers back to an earlier graphic Figure 9 and shows the effect of a different choice in how the file is to be downloaded. Comparing Figure 9 to Figure 18 the big difference is the different file selection technique. In this example, the file download method has been changed from “Get Original Files” to the next radio button which is “Get File Subsets using OPeNDAP”. This choice now allows the user to specify a geographic area (instead of the entire Merra2 grid which spans the globe) and/or select specific file variable(s) for download. Selecting this option opens up the new options.

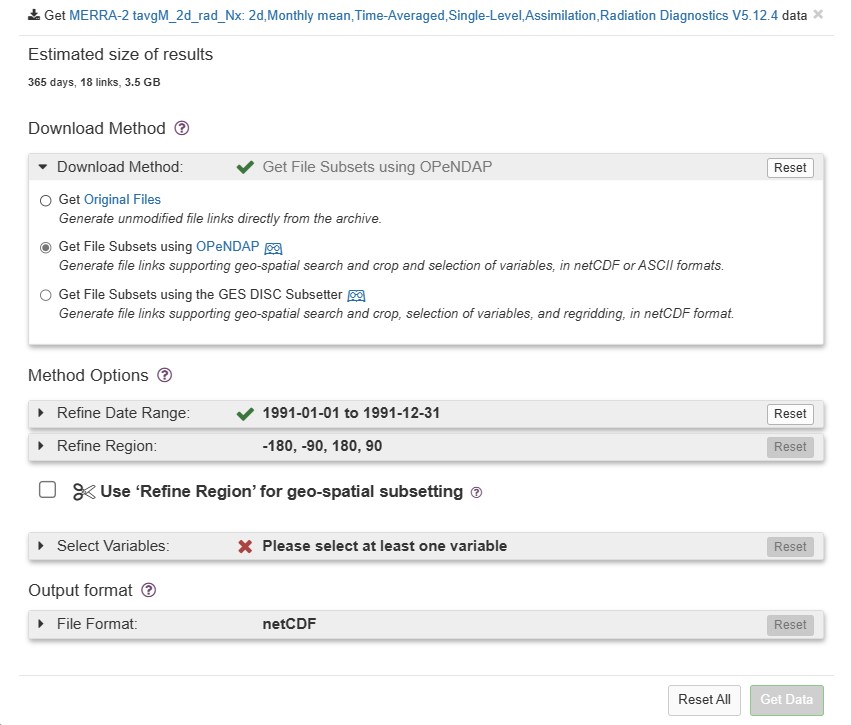


Figure How To Get a File Subset

Moving on to select the Refine Region option we can click the box and see a new GUI.



Figure Selection a Region Subset

In this chart the longitude still ranges from -180 deg to 180 deg but the latitude extent was reduced from +55 to +90 deg to concentrate on the polar regions. The small map inset highlights this change. Because we have chosen to subset the data we must select which variables to download. The next chart shows what happens when the user clicks on the variable selection line.

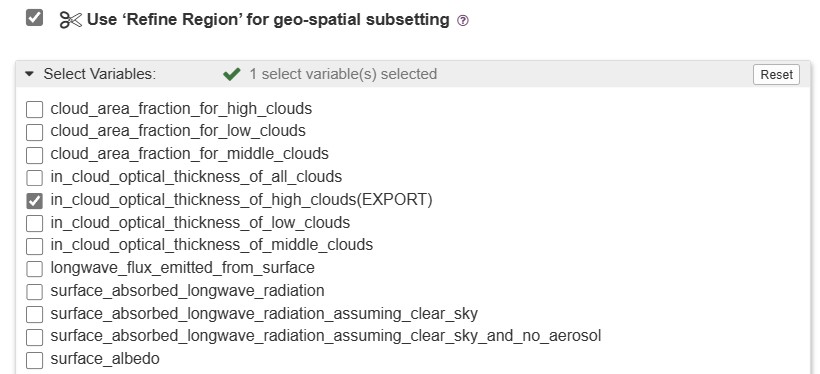


Figure Data Variable Selection

The entire dialog box that opens is scrollable and the more variables on the list, the bigger the UI element is. For this example, only a portion of the dialog is shown. As intended , a single variable in in cloud optical thickness or (TAUHGH) was selected for download. The user is free to pick and choose as he pleases. The more variables selected the bigger the file will be. Since the geographic area was small and only a single georeferenced variable is selected the file size shrinks to well under 1 Mbyte each whereas a full up file would be over 200 Mbytes.

Panoply provides a simple means to see if the final file is what was desired. Here is a screenshot of Panoply opening one these files. Inspection of Figure 21 on the next page demonstrates this fact. There is only 1 Georeferenced file (Geo2D) saved which is TAUHGH. The other 3 files are all 1 D variables that provide time and grid information which can not be removed by the user. The right panel shows the dimensions of TAUHGH have been reduced from 576 x 361 to 576 x 51 points. The 51 points are the latitude values on 0.5 deg centers running from +55 Deg to +90 deg. So, the subsetting operation has produced the desired files.

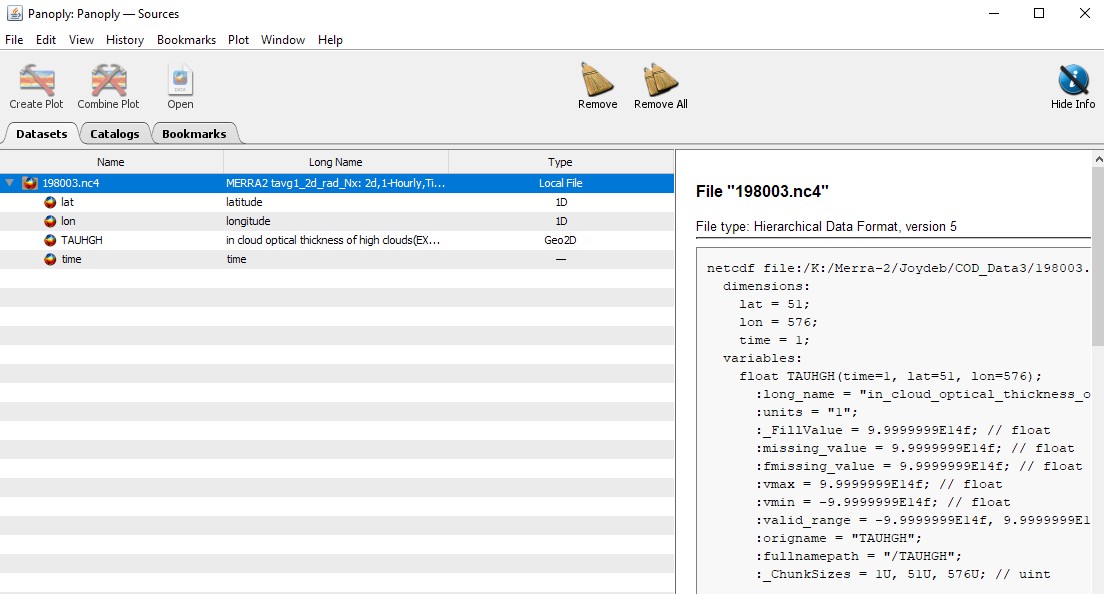


Figure Panoply File Check

# 4.0 Setup Of ProcessCODData Scripts package

The software developed for this task is stored on Dropbox because the size exceeds what can be directly stored on Matlab Central. Overall size at the package is about 1.3 GBytes most of which is taken up with geographic data. Links are provide so that any user with the link can access any data in the project folder on a Read Only basis.

On the next page the graphic Figure 22 depicts the top level set up of the storage used in Dropxbox. The Project folder for these scripts is called “Merra2CODProject. The links provided are to this folder. <https://www.dropbox.com/scl/fo/1yrqt3wnrt5rw0li8p1qk/h?rlkey=02yol6t1g5ddahvo3ufysgv4m&dl=0>

Note that the data repository has been mostly stored on GitHub but this site has a 25 Mb file size limit for most users. The code (m files) are not large but some of the mapping files exceed this limit so the user is better served downloading this data from the Dropbox link above.

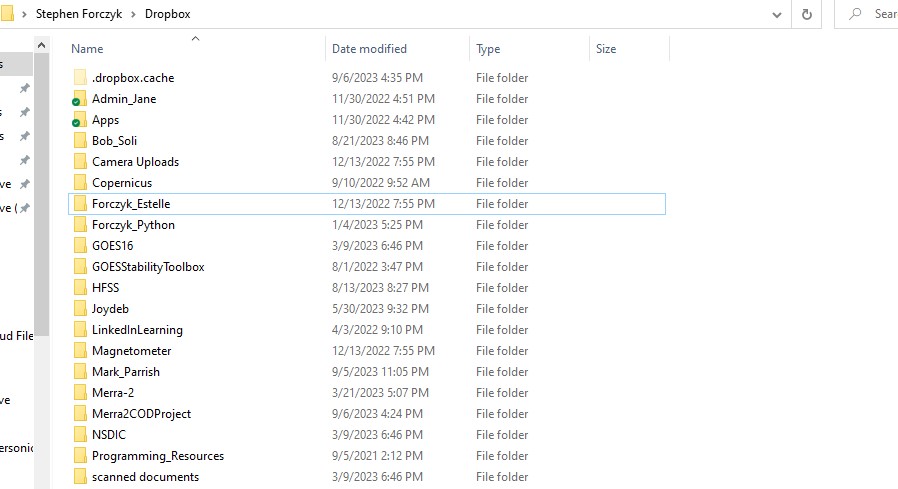


Figure Dropbox Storage Set Up

Now let’s go inside the project folder as this is the general way the project should be set up on another computer. There are a plethora of file types and the author’s preference is to group files in a folders based on their type. This saves a lot of time looking for input or output files.

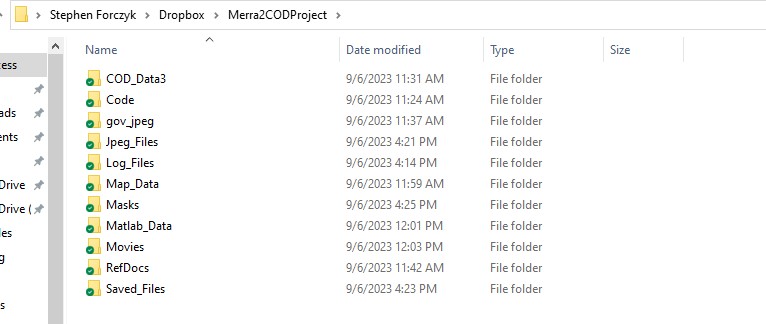


Figure Project Folder Set Up

The basic contents of these folders are shown in the Table Below

**Table 1 Dropbox Project Folder Configuration**

|  |  |  |
| --- | --- | --- |
| **Folder** | **Files** | **Size Of Folder (Mbytes)** |
| COD\_Data3 | The actual netCDF data files | 67 Mbytes |
| Code | Matlab Scripts & functions | .43 Mbytes |
| Gov\_jpeg | Jpeg files from a variety of sources | 4.9 Mbytes |
| Jpeg\_Files | Matlab Generated jpeg files | Variable but could be hundreds of files |
| Log\_Files | Log File Generated By Run | Variable but << 1Mbyte each |
| Map\_Data | Border Files in matlab format | ~ 1 Gbyte |
| Masks | Mask Files | 14 Mbytes |
| Movies | Matlab Created Movie File | 22 Mbytes or less |
| RefDocs | Document References mostly pdf | 257 Mbytes |
| Saved\_Files | Matlab Generated Files to be saved | <2 Mbytes |

Below is a link to the Dropbox folder for this project. The user can view and download these files but not edit them that location

<https://www.dropbox.com/scl/fo/1yrqt3wnrt5rw0li8p1qk/h?rlkey=cz84bq5rpxjdgunj79rkaipos&dl=0>

## 4.2 User Setup of Dropbox Files on Local Computer

The easiest way to set up this project on a new user’s local computer is to click the link above and then download these files one folder at a time at the desired location. The overall content is about 1.3 GBytes shown the chosen location should have at least this much free space. If the user has multiple hard drives the best choice by far is to avoid using the C drive because constant usage of this script will tend to fragment the drive and slow down all operations.

The key script is the “**ProcessCODData.m**” file . This file is where the user will make choices of what files to process and what flag settings will be set to. Prior to running the script, it is important to define the key paths to these folders on the user’s local drive. This is relatively simple and once set will probably not changed again.

Shown below ,Figure 24 is a code snippet from the main script **ProcessCODData.m** covering lines 83 thru 113 of this routine. The user must set the actual paths in use on his own computer. For example, on LOC 84 the user must specify where the Merra 2 files are located. In this case these files ,were located on an external drive labelled K in a Merra2 folder. This must be done for each of the paths indicated. The paths are all global variables so they are accessible for all routines and functions. Not shown is this code section is the addpath statements that will then add them to the matlab search paths. The user must ensure that these folders exist before such an approach can work.



Figure Setting Up Local Paths

The next step is to set up the system adjustable parameters. These must be set in the code itself prior to running the script. The location in the **ProcessCODData.m** can be seen from the next code sectiondisplayed on the next page.

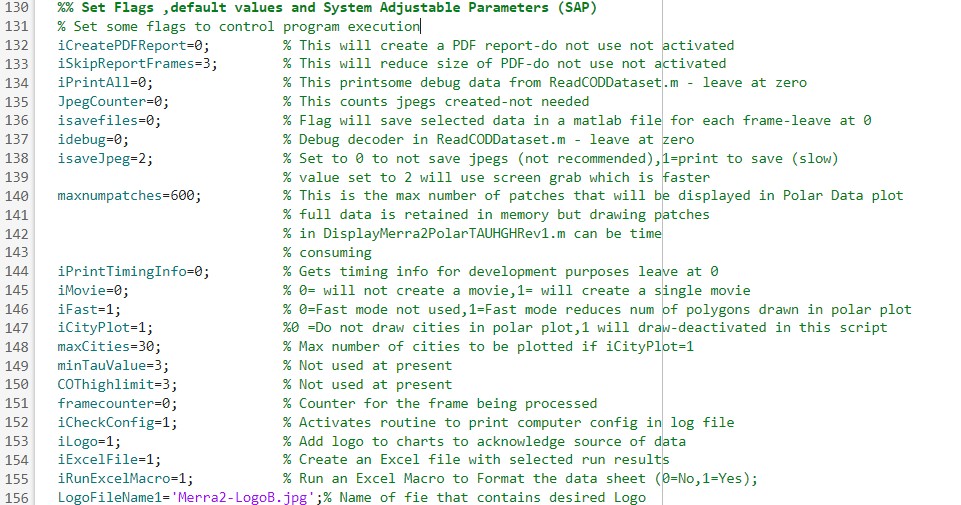


Figure Setting SAP Values

Examining the section of code which spans lines 130 thru 156 of the code shows the values of the user defined flags and give a quick summary of the meanings. These must be set by the prior to the run to incorporate any changes from these “defaults” . Not every flag is really exercised for every run. For example ,the iCityPlot is not used because the north pole plots have two few cities in the Region Of Interest (ROI) to be useful. The flag to Run Excel macros at the moment is deactivated in the script because of issues with macro permissions.

It would have been possible to bring up a GUI dialog box to set these values dynamically but at this stage of the project it does not really appear necessary to do so. One the paths have been set, and the SAP values entered in the code the user is ready to run the code. There is only 1 dialog box that appears during the code execution and this asks the user to select which files will be processed. The user must select at least 2 files or all the files. These files each cover a 1 month period and the naming convention should ensure they will be processed in consecutive time order. As the code functions it will generate a series of polar based maps spanning the 55N to 90N latitude section of the globe. If the user choses a movie file, an ” .mp4” movie file will be created. Of particular use to the user, a log file will be created which records the software and hardware configuration existing at the time of the run as well a flag settings and SAP values. The log file will also print out selected values of the run. Each logfile has a unique name based on the time of run initiation. [SampleLogFile](RefDocs/SampleLogFile.txt) .

# 5.0 Running the Code Package

After reading the previous sections. The user is now ready to run the Matlab script set. The user should bring up the **ProcessCODData.m** code and hit run. If all the paths have been set up correctly. The only user interaction required is the selection of required files. This dialog box is shown in the screengrab below.

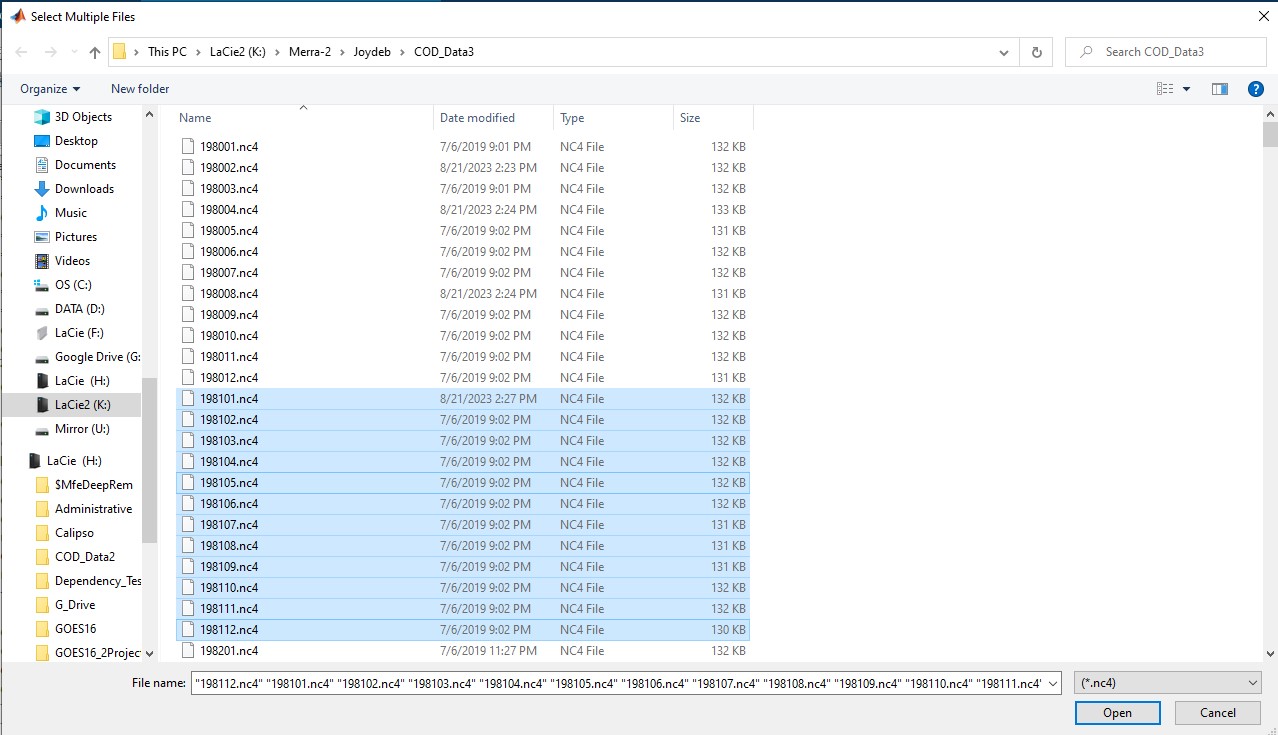


Figure File Selection Dialog

In the example shown above, the user choose to process 1 year set of files , namely the 12 months that make up 1981. After selecting these files hit open and processing will start. If the user hits cancel program execution will stop.

The code that produced this dialog is shown below. It is very straightforward and easy to understand its functionality. There is a requirement to select at least two files for a couple of reasons. The first concerns how the dialog box produces lists and the second is the time series functions. It is not logical to have a time series of just 1 point.

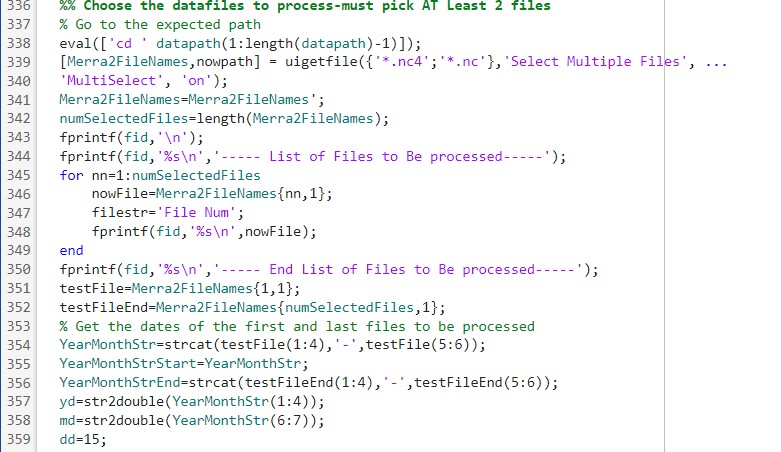


Figure Code to Produce Dialog

After the files are selected the main script will loop through the files one at a time. When processing the first file, the video object will be initialized if a movie is requested. For each file in the loop the routine “**ReadCODDataset.m**” will be called to decode the dataset. This routine does some special actions on the first pass thru in order to initialize arrays and to set up a mapping reference called **RPix**. The area in sq km of each subgrid is calculated. All this needs be done only once as the grid is fixed.

On every pass through the list of files, the TAUHGH variable is read in the decoder section of this routine and then the area of the region of interest that that exceeds a value f 1,2,3,4 or 5 is calculated. In this case the ROI is just the polar regions where cloud optical depth (COD) is less than found in lower latitudes. These values are stored to create time series plots later.

Once in each loop the result is plotted using the routine ” **DisplayMerra2PolarTAUHGHRev1.m** “ .Typical output from this routine is shown in the next chart. Inspection of Figure 28 below shows the COD/TAUGHGH value plotted on a polar region plot that shows country borders. In order show the map features as clearly as possible ,the values of TAUHGH were plotted as patch objects which are colored according to their value. It is assumed that no values exceed 7. The colorbar on the right side of the figure, unfortunately doesn’t seem to run through the same color scale. Nevertheless ,it is possible to see where the high cloud optical thicknesses are. This image from July 1990 was selected because it shows a relatively high level of TAUHGH values in the north polar region.

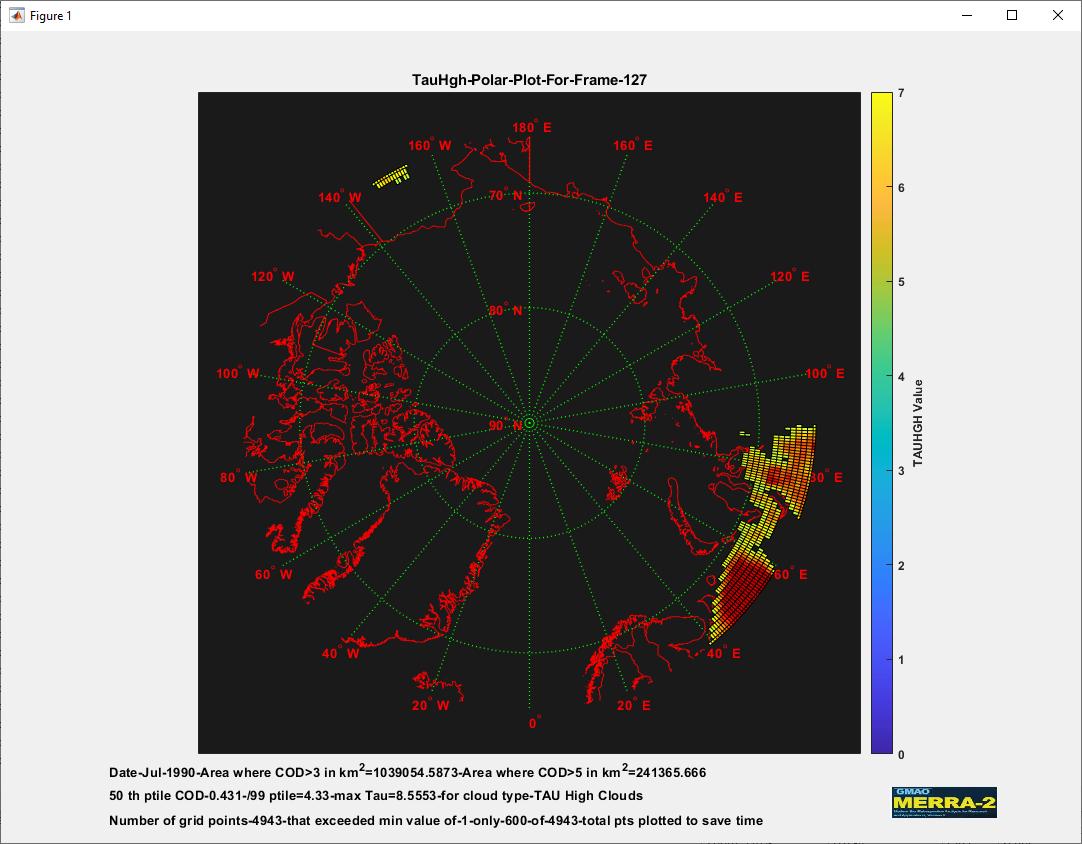


Figure Example TAUHGH Polar Distribution Jul 1990

At the bottom of the figure the text gives the date as month and year ,along with the area where COD exceeds 3 and 5.Line 2 gives a few basic statistics about the TAUHGH values. The bottom line of text provides the total grid points where TAUHGH exceeds 1. In order to save plot time, the iFast flag is set to 1. In this event no more than 600 patches will be plotted. The patches will be sorted by TAUHGH value in descending order so the thickest clouds are always plotted. If iFast were set to 0 all patches would be plotted but run times can get quite large.

When all the files have been processed a new plot can be created from the saved monthly area values of TAUHGH. The chart, Figure 29 on the next page shows just such a plot. As each date frame (1 month estimate of TAUHGH data) is processed 5 different areas are calculated. These areas are the north polar region size in km^2 where the COD is greater than 1,2,3,4 or 5 .

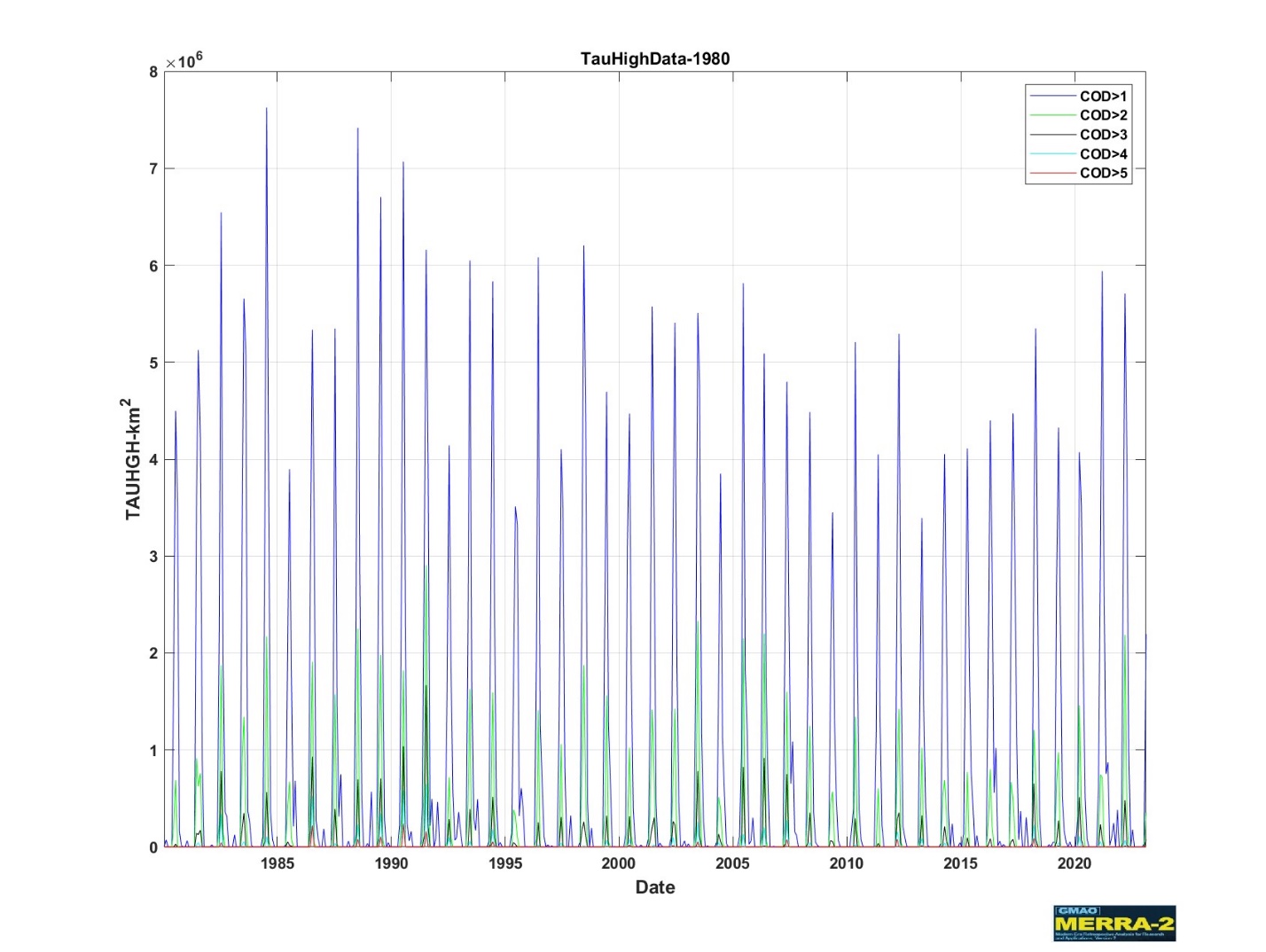


Figure TAUHGH Variations 1980-2023

This time period covered by this timeseries plot is large enough to show the entire Merra2 data collection values but at this scale is his hard to read the actual months. If one looks at the time series table TAUHGHTT the details are clearly visible. First the user should examine how this table was created. To see this look at a small section of code from the **ProcessCODData.m** routine reproduced on the next page on Figure 30.

With only about 10 lines of code the user create two tables. The first table ,TAUHGHTable, has the area at one month intervals where the area in the ROI was greater can COD >1 ,2,3,4 and 5. This is a table variable but not a prime timetable as it does not have a properly formatted date associated with each row of data. This lack is remedied by create a true TimeTable with the addition of a Time Column that has a monthly time step starting on Jan 1980. This timeseries can then be plotted to create Figure 29.

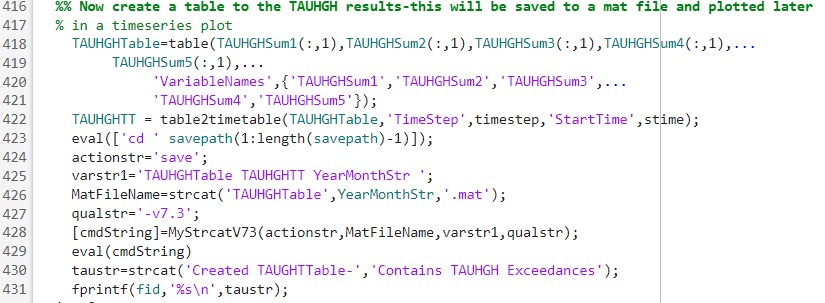


Figure Code To Create TAUHGHTT Table

Because the of the length of the table it is not convenient to show all the rows of data in a single graphic. The next chart just shows the entries for 1980. The time variable can now be clearly seen in the first column, The next 5 columns contain the area in the north polar ROI where the TAUHGH value exceed 1,2,3,4 and 5 and up. It is immediately apparent that the 3 summer months have the highest optical cloud depths and this pattern repeats of the 40+ years of data collection

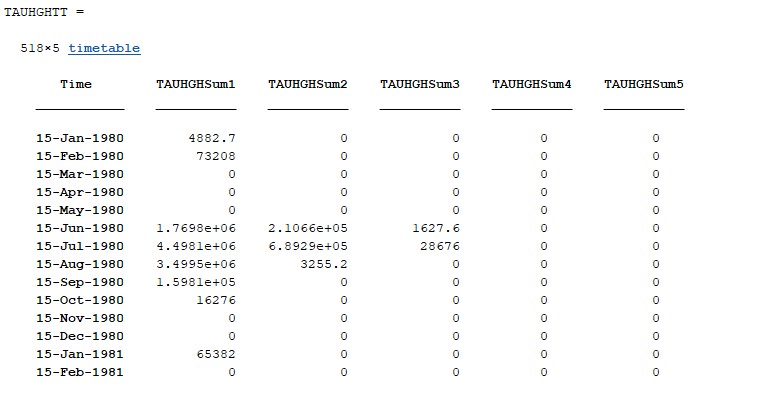


Figure 1 Year Section of TAUHGHTT Time Table

## 5.1 Movie Creation

As noted earlier ,it is possible to create movies of the individual monthly datasets us Matlab. If iMovie is set to 1 a movie will be created. The user does not assign a name to the movie as the name is based on the dataset start and end dates. Any movies created will be stored in the Movie folder. A sample movie could be attached here but these files can be larger and playback would be slow. The best course of action is to go to the movie folder an open the .mp4 movie file in the user’s player of choice.

## 5.2 Required Toolboxes

Every attempt was made to keep the footprint of this software package small. For this reason ,only the basic Matlab (R022B and above) and the Mapping toolbox are required. Generally, the Matlab Report Generator is optionally used to create a report after a run but this did not appear to be required in this instance. It might added in any updated software package as an option.

## 5.3 Known Code Issues

There are a few known code issues which will be noted here.

* The polar plot color scale needs to be adjusted to the actual TAUHGH values but this is only cosmetic
* The attempt to run Excel Macros directly from Matlab has encountered problems due to a recent Windows Software update of 9/4/23 . Workaround is available
* Time Series Tables not working on release R2017b or earlier due to difficulty in created time column-a workaround should be available at a later time

## 6.0 References

A list of reference materials is included below. Only some of these are referred to in this guide but they were deemed to be a useful reading list

Gelaro et al., “The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2).”

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Silva, A. M. da, Gu, W., … Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Saha, J., Price, C., Plotnik, T., & Guha, A. (2023). Are thunderstorms linked to the rapid Sea ice loss in the Arctic? *Atmospheric Research*, *294*, 106988. https://doi.org/10.1016/j.atmosres.2023.106988