

MultiSpec Tutorial: Handling HDF Formatted Image Files

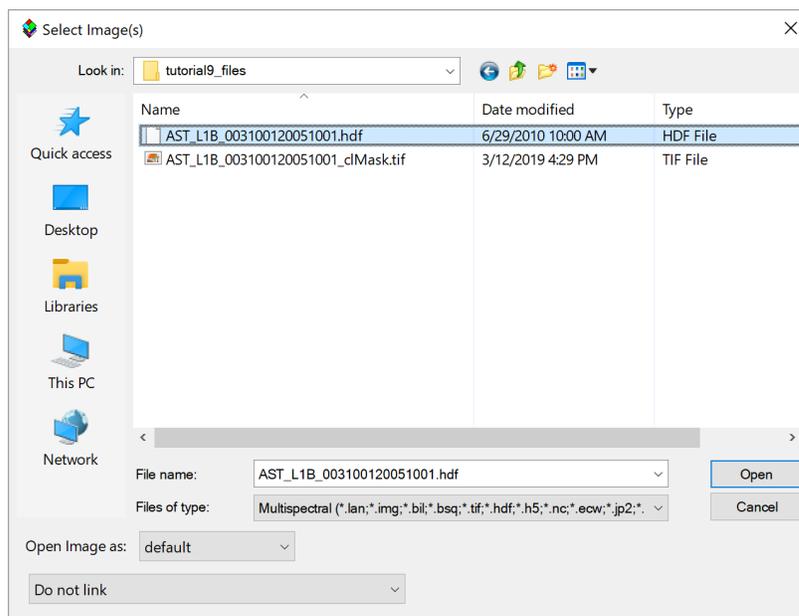
Requirements: MultiSpec application and AST_L1B_003100120051001.hdf.

Hierarchical Data Format (HDF) is a data file format designed by the National Center for Supercomputing Applications (NCSA) to assist users in the storage and manipulation of scientific data across diverse operating systems and machines. HDF supports a variety of data types: scientific data arrays, tables, and text annotations, as well as several types of raster images and their associated color palettes. There are two different varieties of HDF, one known as HDF4 and the newer one called HDF5. HDF5 can handle very large file sizes among other improvements (<https://www.earthdata.nasa.gov/about/esdis/esco/standards-practices/hdf5>). NetCDF (<https://www.unidata.ucar.edu/software/netcdf/>) is another hierarchical data format. MultiSpec can handle most HDF4, HDF5, and NetCDF formatted files which contain image data.

A HDF formatted file can contain from one to many data sets. The data sets can be any combination of tables, arrays and raster images. Each of the raster images can contain multiple channels and represent different spatial resolutions and data types. MultiSpec allows one to select an image data set (or group of image data sets) in the HDF file to be used like any other raster image file. This tutorial illustrates how to select HDF data sets to treat as any other image file.

Some common image files that are stored as HDF formatted files are Terra & Aqua MODIS and Terra ASTER data. The image used for this tutorial was acquired by the ASTER sensor on the Terra satellite on October 1, 2005 over an area around West Lafayette, Indiana.

9.1 From the **File** menu, select **Open Image....** A dialog box will open to allow one to select the image data file one wishes to use.



- 9.2 Select AST_L1B_003100120051001.hdf in the Tutorial9 folder and then select “*Open*”, or simply double-click on AST_L1B_003100120051001.hdf.
- 9.3 The “Set Multispectral File Format Specifications” dialog box will appear if there is more than one data set in the file.

Set Multispectral File Format Specifications

For: AST_L1B_003100120051001.hdf

Size Specifications

Number of Lines	4200	Preline Bytes	0
Number of Columns	4980	Postline Bytes	0
Number of Channels	3	Prechannel Bytes	0
Start Line Number	1	Postchannel Bytes	0
Start Column Number	1		

Band Interleave Format: Band Non-sequential

Data Value Type: 8-bit Unsigned Integer Swap Bytes

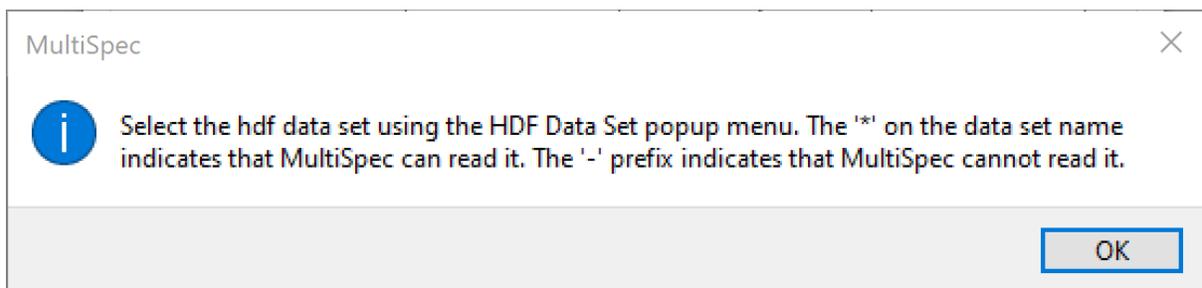
Treat Lines from Bottom to Top

Data Fill Value:

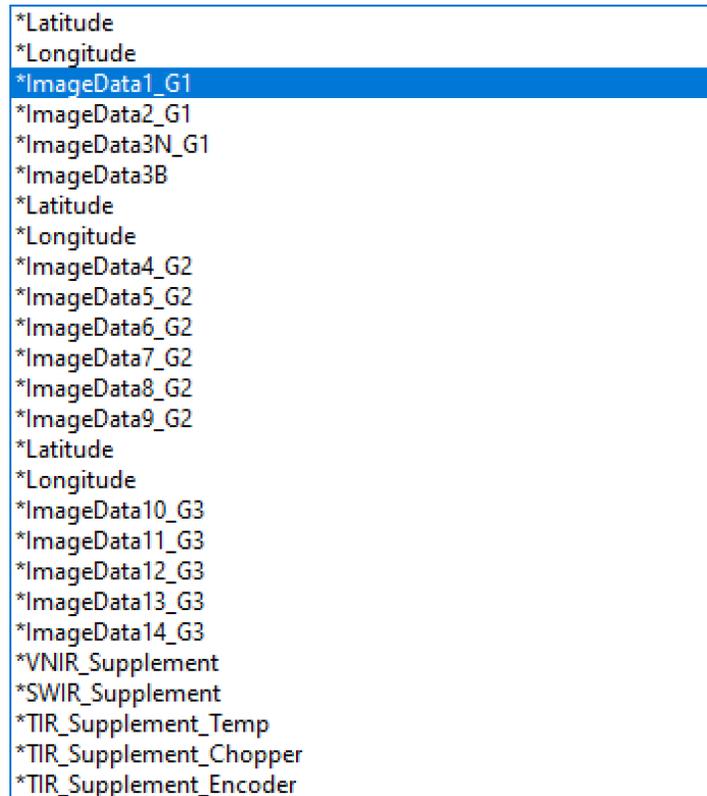
HDF Data Set: *ImageData1_G1

OK Cancel

This dialog box allows one to see what data sets are available in the file and to select the data set(s) to be used in further analysis. One can select the '?' button to obtain a description of what to do with the HDF Data Set popup menu and what the symbols at the beginning of the data set lines mean.



9.4 Select (mouse down) on the HDF Data Sets popup menu to get the popup menu shown below.



The popup menu lists all of the data sets in the HDF formatted file. The ‘*’ in front of the data set name indicates that the data set is a raster image that MultiSpec can handle. A ‘_’ in front of the name indicates that the data set is not a raster image or is a raster image that MultiSpec cannot handle such as one with complex data values.

The _G1 labels at the end of three of the data set names indicate that these three data sets will be grouped (linked) together automatically if selected. The data sets have the same number of lines & columns and the data type is the same for all three. One can select any of the three sets (ImageData1, ImageData2 or ImageData3N) to then be able to work with an image with three channels. The same is true for those data sets with the _G2 label (6 channels) and for those with the _G3 label (5 channels). For the ASTER data in this tutorial, the G1 data sets are the VNIR channels (15 meters spatial resolution), the G2 data sets are the SWIR channels (30 meter spatial resolution) and the G3 data sets are the TIR channels (90 meter spatial resolution).

9.5 Select the ImageData1 data set. You will notice that the “Set Multispectral File Format Specifications” dialog box now displays the specification for the selected data set illustrated below.

Set Multispectral File Format Specifications

For: AST_L1B_003100120051001.hdf

Size Specifications

Number of Lines	4200	Preline Bytes	0
Number of Columns	4980	Postline Bytes	0
Number of Channels	3	Prechannel Bytes	0
Start Line Number	1	Postchannel Bytes	0
Start Column Number	1		

Band Interleave Format: Band Non-sequential

Data Value Type: 8-bit Unsigned Integer Swap Bytes

Treat Lines from Bottom to Top

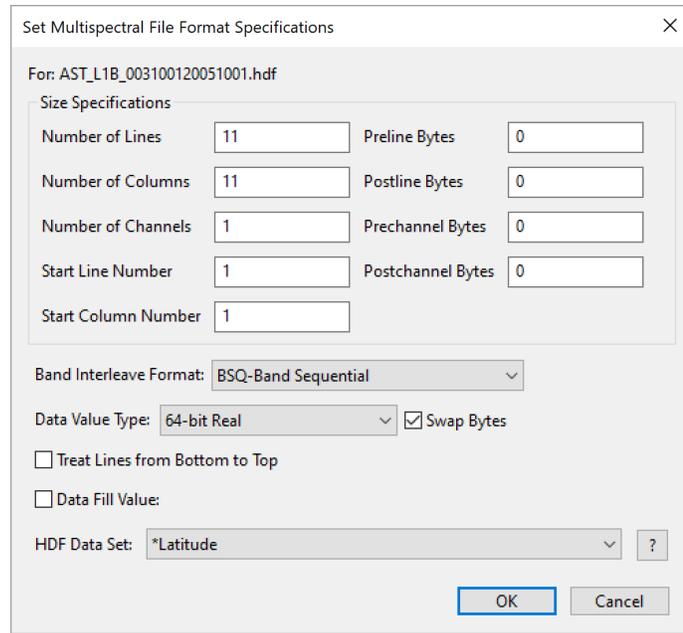
Data Fill Value:

HDF Data Set: *ImageData1_G1

OK Cancel

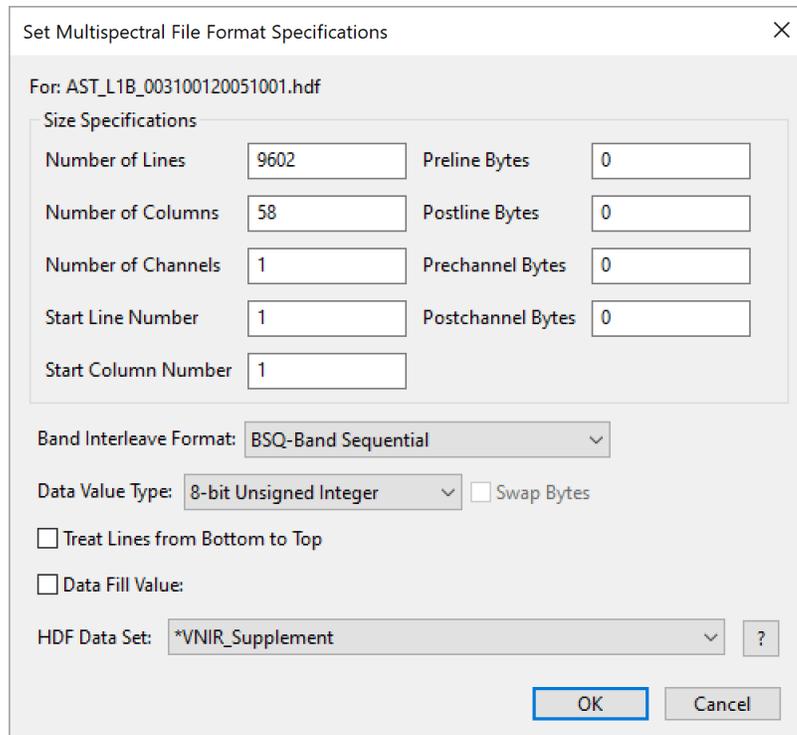
There are 4200 lines and 4980 columns in the selected ‘grouped’ data sets. The number of channels imply that 3 data sets are grouped together (although note that a single data set can have multiple channels). The data value type is 8-bit unsigned integer.

If one now mouses down on the HDF Data Sets popup menu again and then selects the first Latitude data set, one will obtain the information for this data set as shown below.



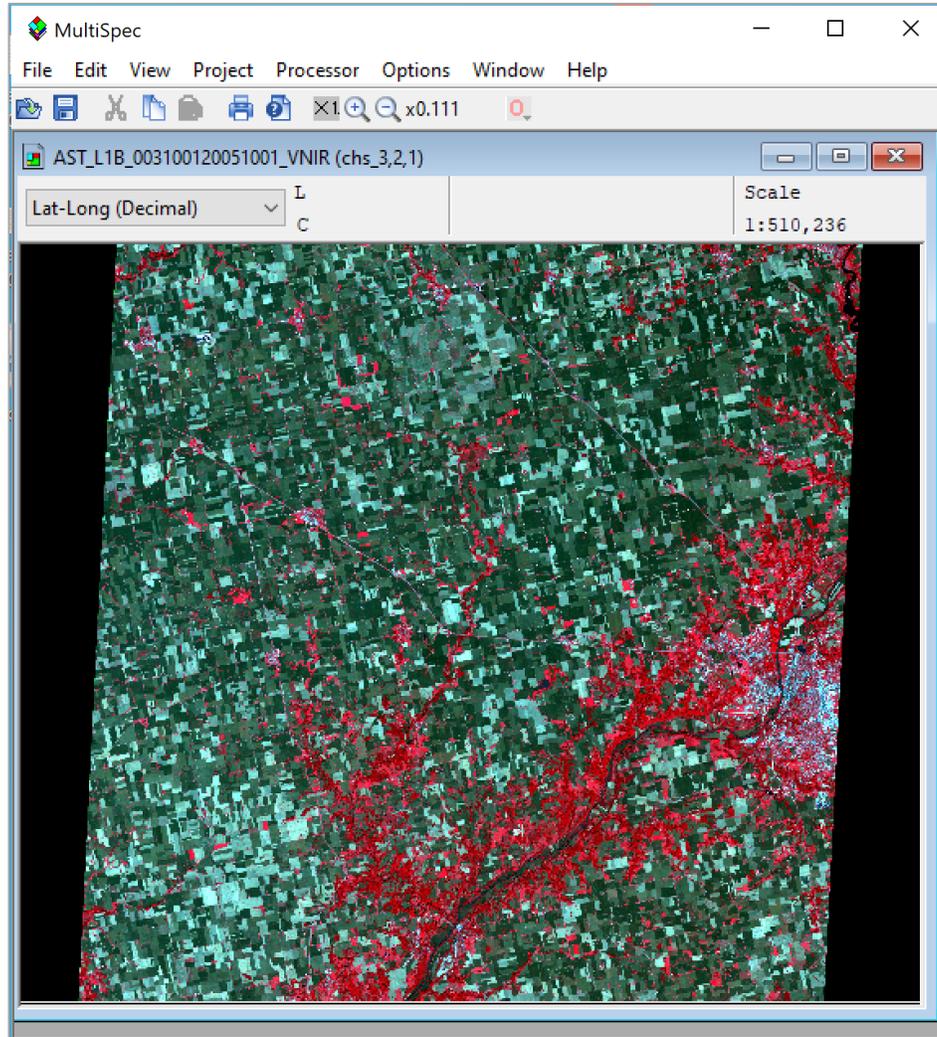
There are 11 lines, 11 columns and 1 channel in this Latitude data set. The data value type is 64-bit real. The “Swap Bytes” box may or may not be checked depending on the processor in the computer being used.

If one now mouses down on the HDF Data Sets popup menu again and then selects the VNIR_Supplement data set, one will obtain the information as shown below.



There are 9602 lines, 58 columns and 1 channel in this data set. The data value type is 8-bit unsigned integer. The VNIR_Supplement data set contains calibration information.

- 9.6 Now reselect the “ImageData1” data set and then select the “OK” button. The ‘Multispectral Display Specifications’ dialog box will now appear to allow one to display the image just like that from any other image file. (See the Display & Inspection of Image Data tutorial. The default channel combination display of this image will appear as below.



Moreover, one can use the data set like any other image data file such as creating an image file in a different format (Processor->Reformat->Change Image File Format) or creating an unsupervised classification (Processor->Cluster), etc.

- 9.7 One can change the data set being used if needed by selecting the “Edit->Image Description...” menu item. This will open the “Set Multispectral File Format Specifications” dialog box as described in section 9.3.

Notes: MultiSpec has been programmed so that if one is working with several similar HDF formatted data sets (such as several ASTER data files), the default selected data set will be the same one that was selected in the most recently opened HDF file.

Concluding Remarks

There are many other operations that one can do with MultiSpec. See other tutorials or the MultiSpec Introduction at the MultiSpec web site for more information. The MultiSpec web site is: <https://engineering.purdue.edu/~biehl/MultiSpec/>. Or contact Larry Biehl at biehl@purdue.edu with questions.