Learning HDF5 with HDFView

This tutorial enables you to get a feel for HDF5 by using the HDFView browser. It does NOT require any programming experience.

Installation

- Download and install HDFView. It can be downloaded from the Download HDFView page.
- Obtain the storm1.txt text file, used in the tutorial.

Begin Tutorial

Once you have HDFView installed, bring it up and you are ready to begin the tutorial.

> **Unable to complete tutorial because fields are greyed out ?**
> This tutorial requires that the default HDFView File Access Mode be *Read / Write*. If fields are greyed out so that you cannot select them, then the File Access Mode is *Read Only*.
>
> To change the File Access Mode follow these steps:
>
> - Bring up HDFView
>
> - Left-mouse click on the *Tools* pull-down menu and select *User Options*.
>
> - A Preferences window pops up with the *General Settings* tab selected.
>   About half-way down you will see *Default File Access Mode*.
>   Select *Read / Write*.
>
> - Click on *Apply and Close* at the bottom of the window.
>
> - Close down HDFView.
>
> - Bring HDFView back up and try the tutorial again.
>
> PLEASE BE AWARE that selecting a File Access Mode of *Read / Write* can result in changes to the timestamp of HDF files that are viewed with HDFView. In general, a File Access Mode of *Read Only* should be used to ensure that this does not occur.

Topics Covered

Following are the topics covered in the tutorial. The first topic creates the file that is used in the subsequent topics.

1. Creating a New HDF5 file with a Contiguous Dataset
2. Displaying a Dataset as an Image
3. Creating Attributes
4. Creating a Compressed and Chunked Dataset
5. Creating an Image and a Subset
6. Creating a Table (Compound Dataset)

Topics

Creating a New HDF5 File with a Contiguous Dataset

The steps below describe how to create a file (storm.h5), group (*/Data*), and a contiguous dataset (*/Data/Storm*) using HDFView. A group is an HDF5 object that allows objects to be collected together. A dataset is an array of data values. A contiguous dataset is one that is stored as a single block in the HDF5 file.

- Select the *File* pull-down menu at the top left, and then select *New -> HDF5*.
- Specify a location and type in storm.h5 for the name of your file, and click on the *Save* button.

On the left side of the HDFView screen is the "tree" view. Once you have created storm.h5,

you will see a "5" folder icon in the tree view, followed by the name of your file, storm.h5.
At this point you have an empty HDF5 file that contains nothing but the *root* group, "/".
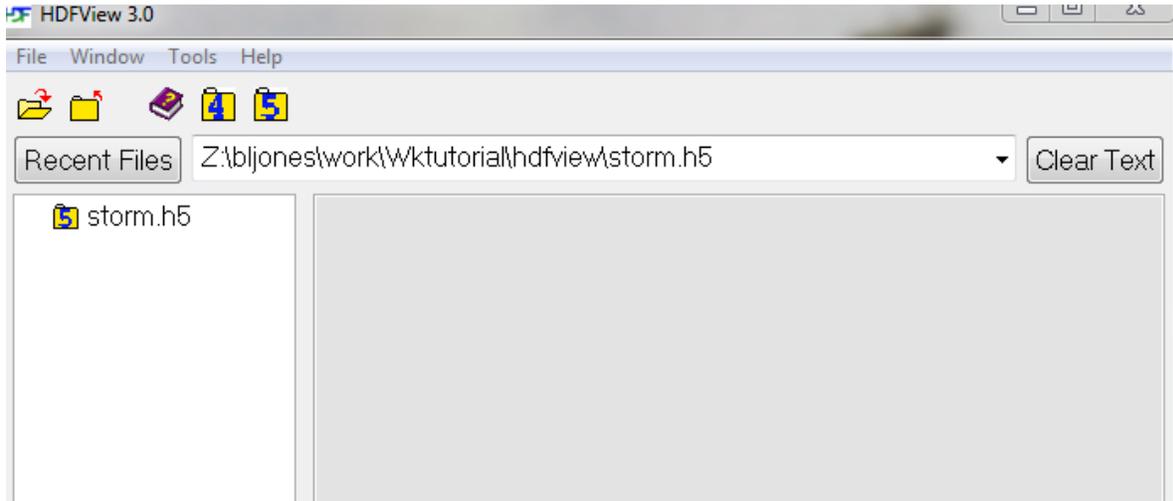
- Right click on storm.h5, and select *New -> Group* .
- Enter *Data* for the name of the group and then click the Ok button.

In the tree view you will see the folder (group) *Data* under the storm.h5 file.
The path to this group is */Data*.
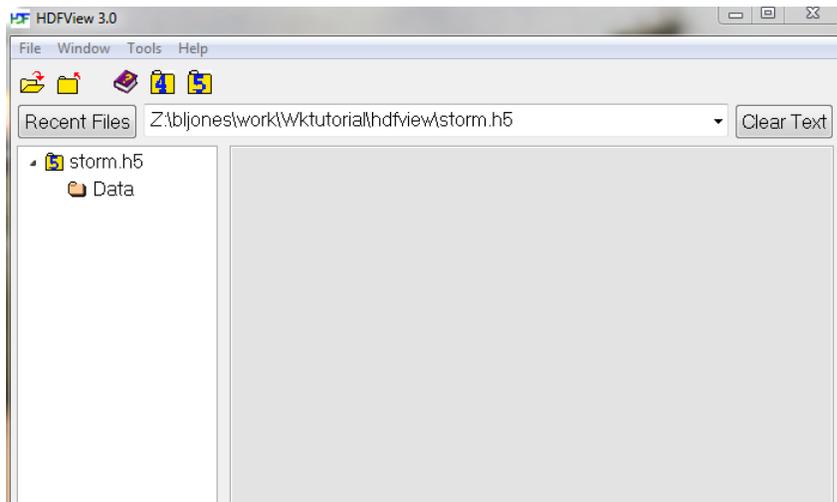
- Right click on the group *Data* and select *New -> Dataset*.
- A window pops up on the right. Fill in the information as follows, and then click Ok (leave the Datatype information as is):

Dataset Name	*Storm*
Under Dataspace, Current size	57x57
Under Dataspace, Maximum Size	57x57
Layout	*Contiguous* (default)

- Left click on the *Data* group in the tree view.

You will see *Storm* listed under the group *Data* in the tree view on the left.

The path to *Storm* is */Data/Storm*.
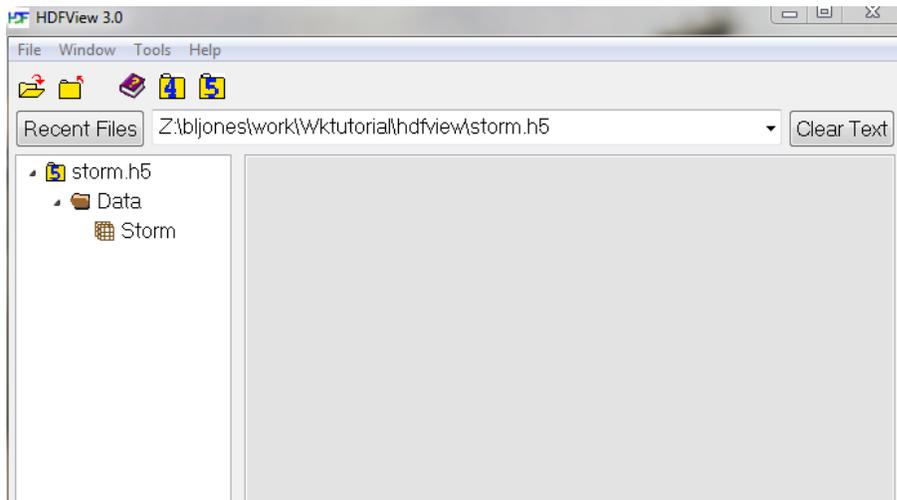
- Double left click on the *Storm* dataset in the tree view. A window with an empty spreadsheet pops open.

- Copy the data from the storm1.txt file into the dataset.

  If you downloaded storm1.txt, then right click on the *Table* menu and select *Import Data from Text File*. Specify a location, select storm1.txt and click on the *Open* button. Answer *Yes* in the dialog box that pops up (which asks if you wish to paste the selected data).

  Alternately, you can copy/paste directly. Select and copy the data in a separate window. Position your cursor at (0,0) in your table, and select *Paste* from the *Table* menu.

Data values will be added to the spreadsheet for the dataset.
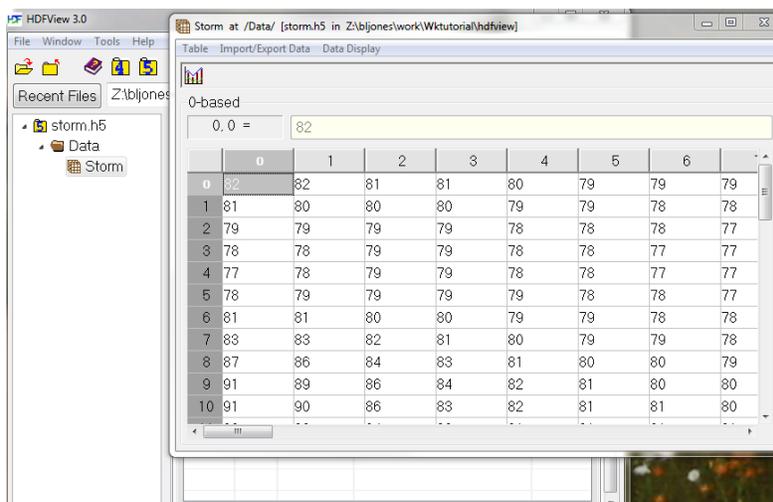
- Close the dataset, and save the data.

Displaying a Dataset as an Image

Any dataset can be viewed as an image in HDFView. Below are the steps that demonstrate this.

- Right click on *Storm* in the tree view, and select *Open As*.
- Select the *Image* button under *Display As* (near the top) in the Dataset Selection window that pops up. Then click OK at the bottom of the window.
  The image will be displayed. At the top of the *Image* window, you will see both a + and - magnifying glass button. Make the image larger and smaller with these buttons.

- The rainbow icon brings you to the Image Palette window. Click on that to play with the palette (GrayWave probably is the best choice). Close.

Creating Attributes

Additional information to describe an object can be stored in attributes. An attribute can be added to a group or dataset with HDFView.

The following illustrates how to add an attribute to the group /Data:

- Click on the /Data folder in the tree view. You will see two tabs, Object Attribute Info and General Object Info, in the pane on the right site of the HDFView window.

In the Object Attribute Info pane you will see that there are no attributes attached

to the group. There is an Add Attribute and Delete Attribute button for adding and

removing attributes.
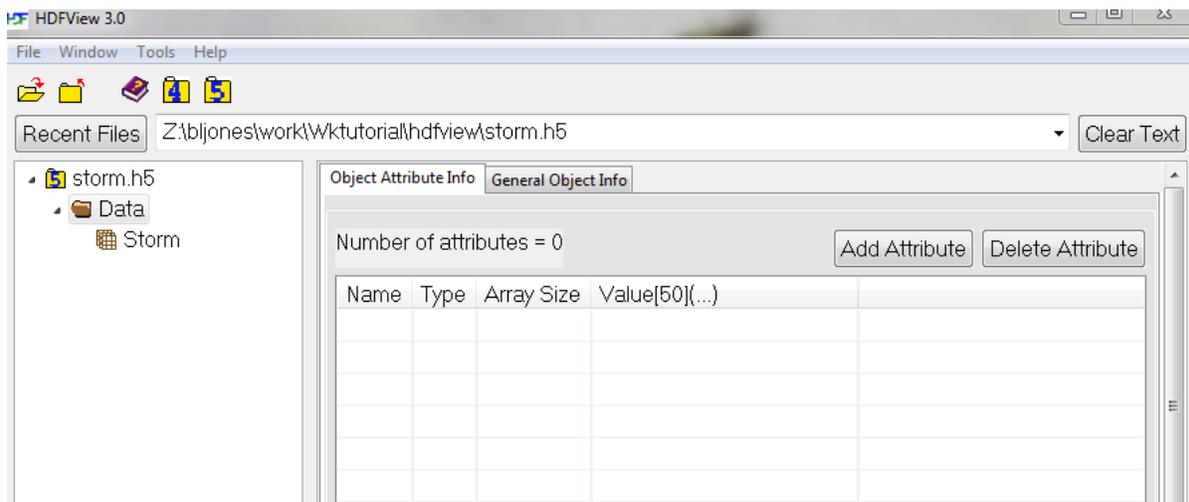
- With the left mouse button, select the Add Attribute button.
- Select the Add Attribute button to add an attribute with these values:

Name	BatchID
Type	INTEGER
Size (bits)	32
Value	3343

- Select the Ok button. The attribute will show up in the Properties window.
- Close the Properties window.

Adding an attribute to a dataset is very similar to adding an attribute to a group. For example, the following adds an attribute to the */Storm* dataset:

- Left mouse click on the */Storm* dataset in the tree view. You will see the *Object Attribute Info* and *General Object Info* tabs on the right.
- In the *Object Attribute Info* pane select the Add Attribute button and enter an attribute with these values. (Be sure to add a *String Length* or the string will bet truncated to one character!):

Name	*Units*
Type	STRING
String Length	3
Value	*m/s*

- Select the Ok button. The attribute will be displayed in the window.
- Close the Properties window.

Notice that a scarlet letter A is attached to both the group and dataset in the tree view on the left.

This indicates that there is at least one attribute attached to the object.
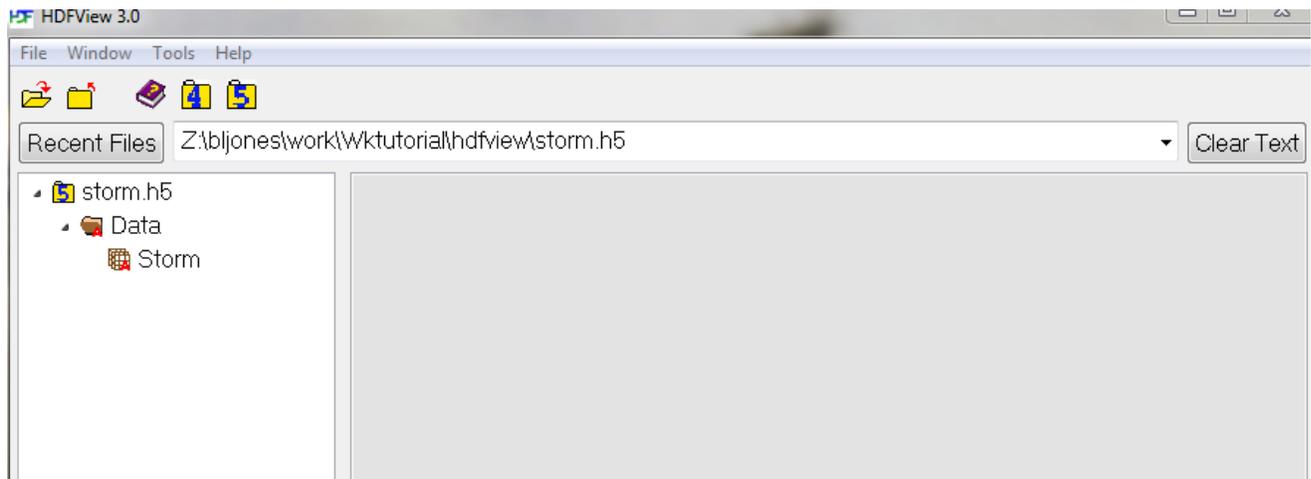

Creating a Compressed and Chunked Dataset

A chunked and compressed dataset can be created using HDFView. A compressed dataset is a dataset whose size has been compressed to take up less space. In order to compress an HDF5 dataset, the dataset must be stored with a chunked dataset layout (as multiple *chunks* that are stored separately in the file).

Please note that the chunk sizes used in this topic are for demonstration purposes only. For information on chunking and specifying an appropriate chunk size, see the Chunking in HDF5 documentation.

Also see the HDF5 Tutorial topic on Creating a Compressed Dataset.

- Right click on storm.h5. Select *New -> Group*.
- Enter *Image* for the name of the group, and click the OK button.

A new group (folder) entitled *Image* will be displayed in the tree view

on the left, under the *storm.h5* file. Its path is: */Image*
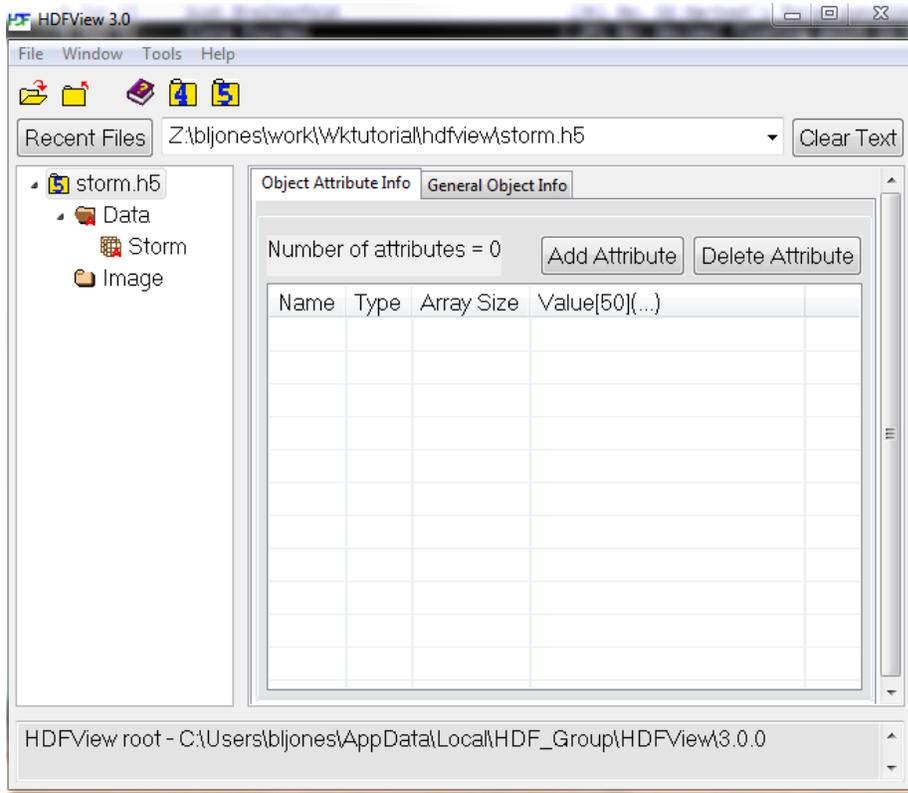
- Right click on the *Image* group, and select *New -> Dataset*.
- Enter the following information for the dataset. Leave the Datatype as is (INTEGER):

Dataset Name	Another Storm
Under Dataspace, Current size	57x57
Under Dataspace, Set Max Size	57x57
Storage Layout	Chunked
Chunk Size	20x20
Compression	gzip
Compression Level	9

The dataset gets created and will show up under the *Image* folder in the tree view on the left. Its path is: "/Image/Another Storm"
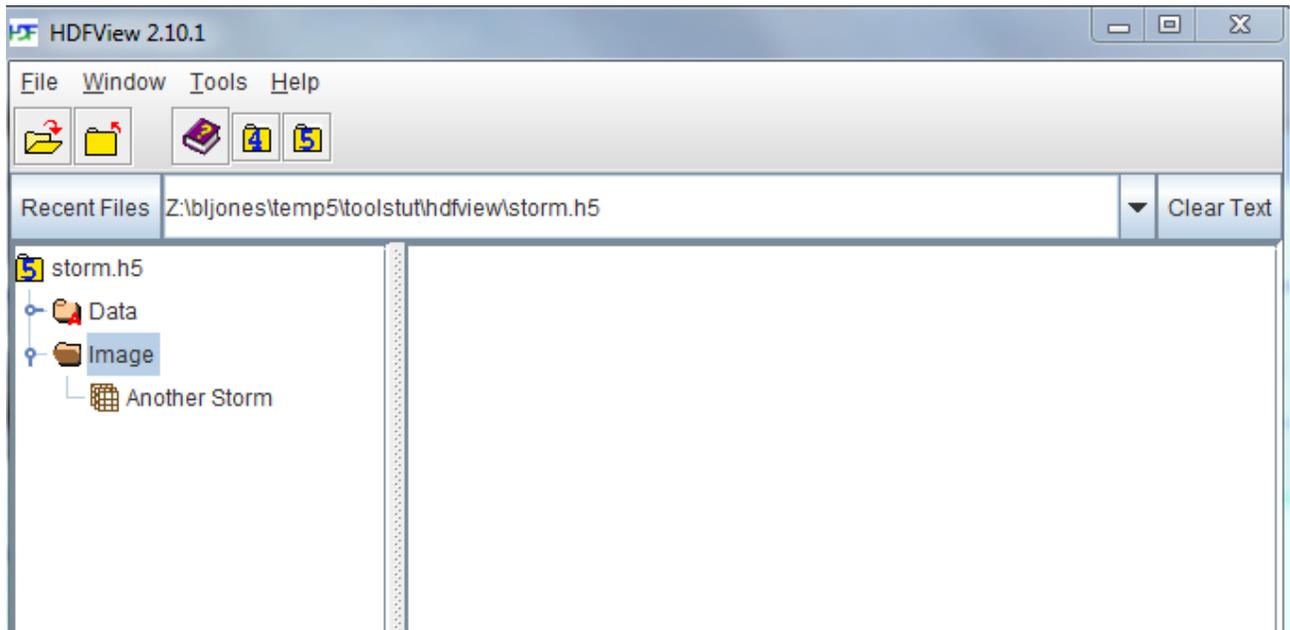
- Double left-mouse click on the *Another Storm* dataset.

A spreadsheet window will pop up on the right side of the screen, containing zeroes.
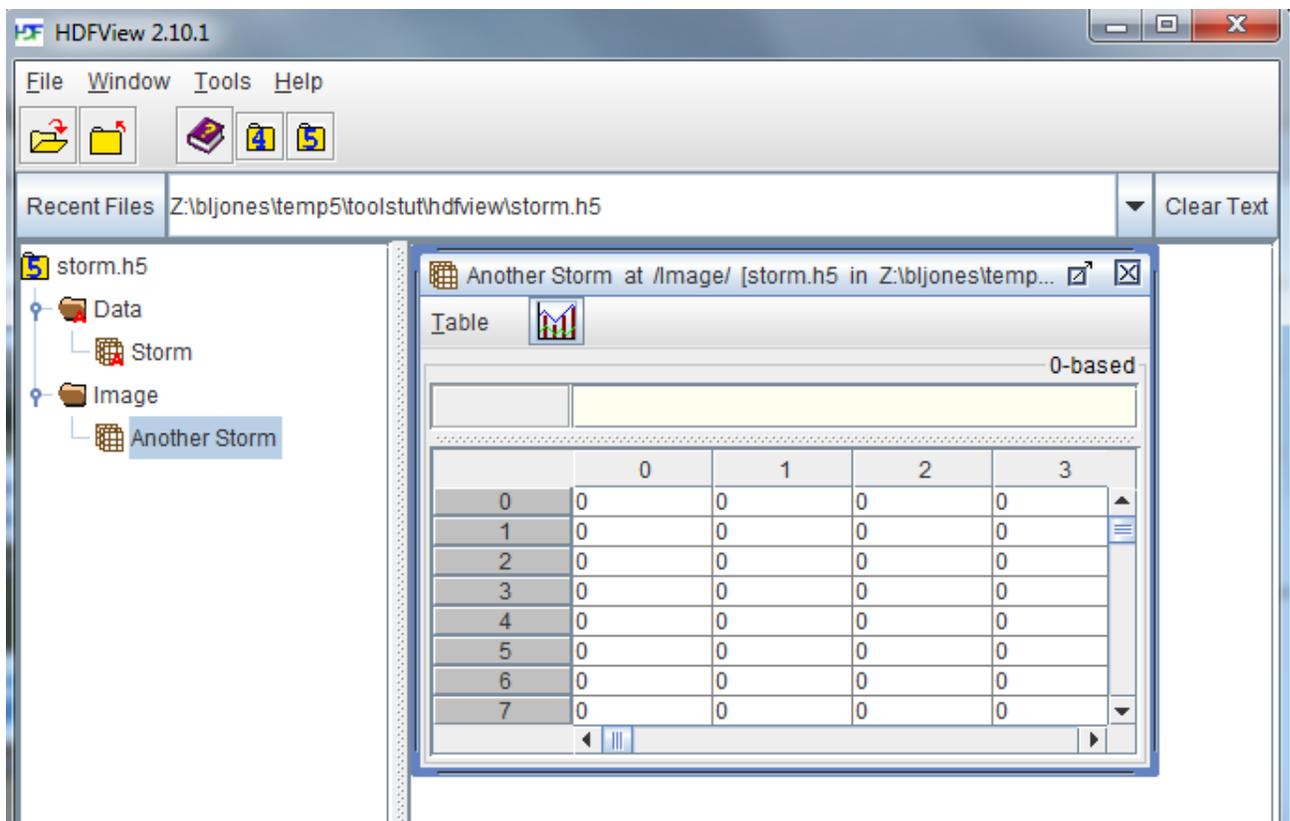
- Copy the data from the storm1.txt file into the dataset. (See the previous topic for copying storm1.txt into a dataset.)
- Close the table, and save the data.
- Right click on *Another Storm*, and select *Open As*.
- Select the *Image* button in the Dataset Selection window that pops up. Click the Ok button at the bottom of the window.

You should see an image in the window that pops up on the right.
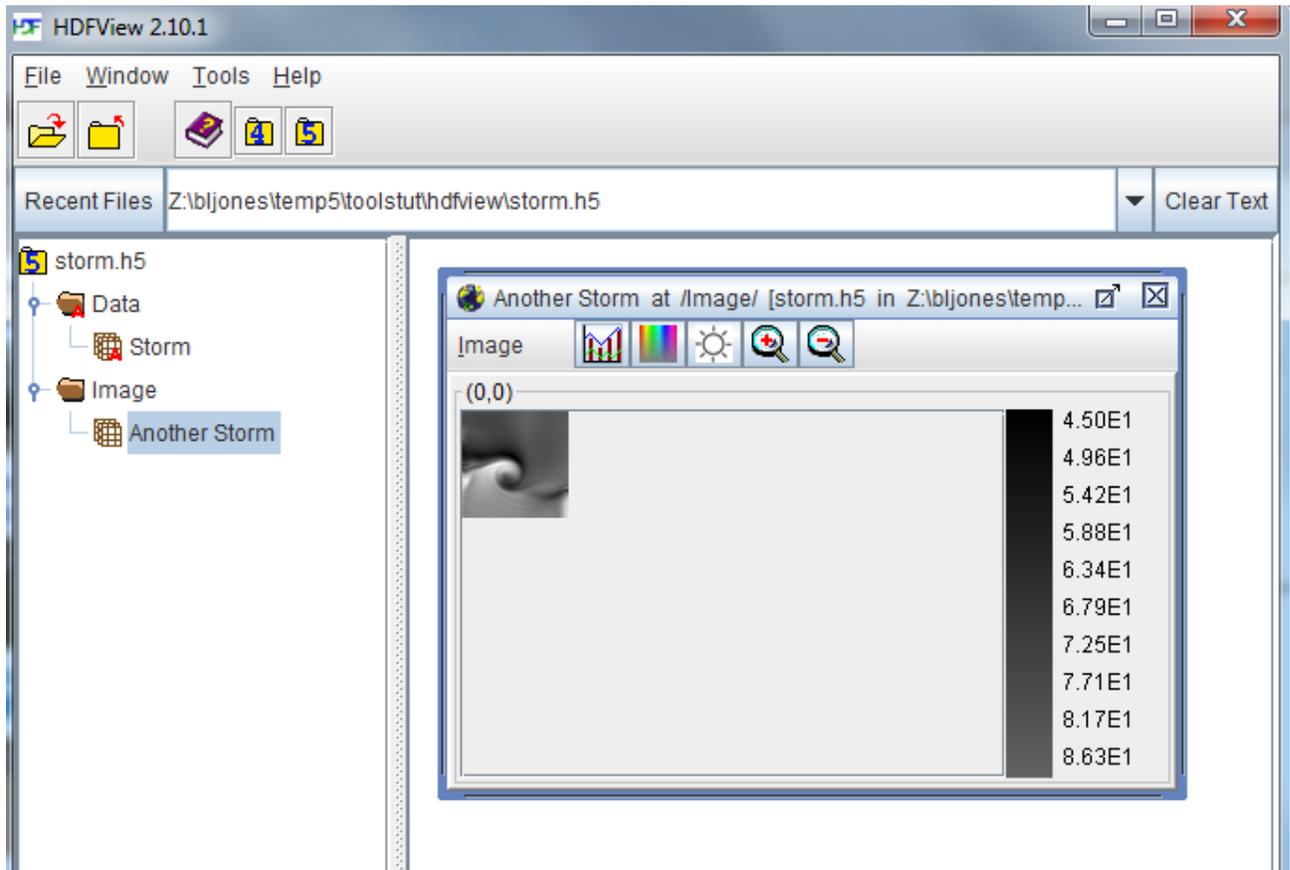
Creating an Image and a Subset

A previous topic demonstrated how to view any dataset as an image in HDFView. With HDFView you can also create an image to begin with, as is shown below.

- Right click on the *Data* group and select *New -> Image*.
- A window pops up on the right. Enter the following and then click Ok:

| Image name	Storm Image
Height	57
Width	57

- Close the dataset.
- Double left-mouse click on the *Data* group to see its contents.

You will see *Storm Image* listed with an "earth" icon to the left of it, indicating an image.
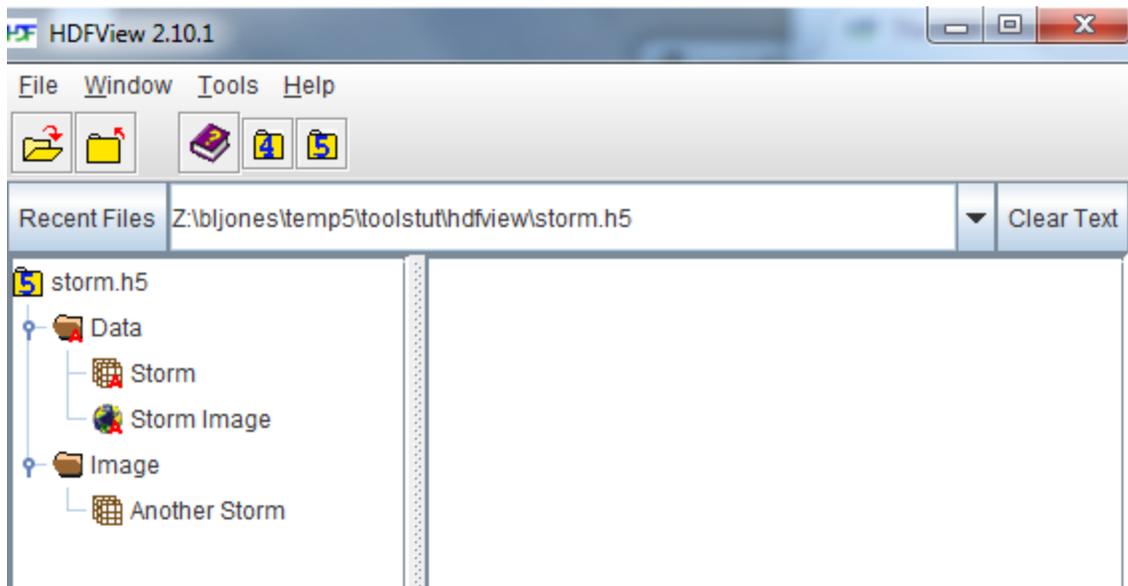
- Add data to the *Storm Image* dataset as was shown previously:

    - Right click on *Storm Image*, and select *Open As* to open the Dataset Selection window.
    - Click on the *Spreadsheet* button at the top left of the Dataset Selection window to view the image as a spreadsheet.
    - Copy the data from the storm1.txt file into the dataset.
    - Close the dataset and save the data.
- Left double click on *Storm Image* to see the image. Close the dataset.
- Right click on *Storm Image* and select *Show Properties* from the pop-up menu, to open the Properties window. Select the Attributes tab.

In the Properties window you will see an attribute called CLASS. This attribute indicates to HDFView that the dataset is an image. See the HDF5 Image and Palette Specification for further information.
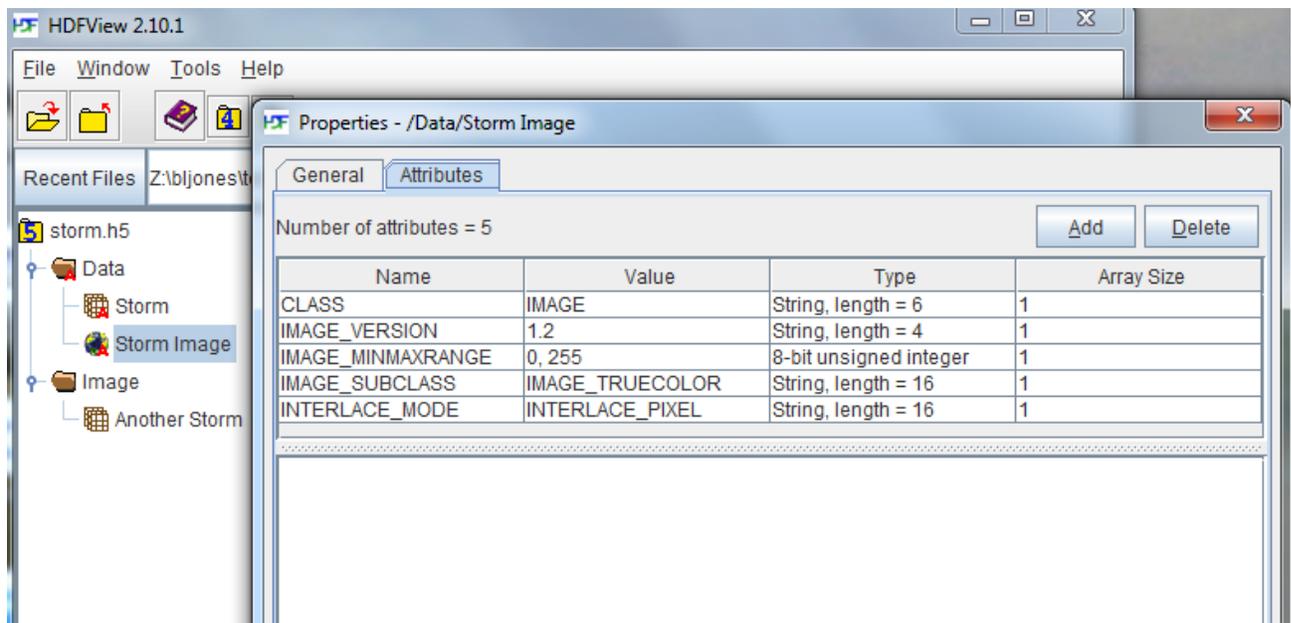

- Close the Properties window.
- Right click on *Storm Image* and select *Open As* to bring up the Data Selection window.
- Select a subset by clicking the left mouse on the image in the window and dragging the mouse. Notice that the Height and Width values change. Select to display it as an image. Click Ok.

A window pops up with the selected portion of the image displayed in it.
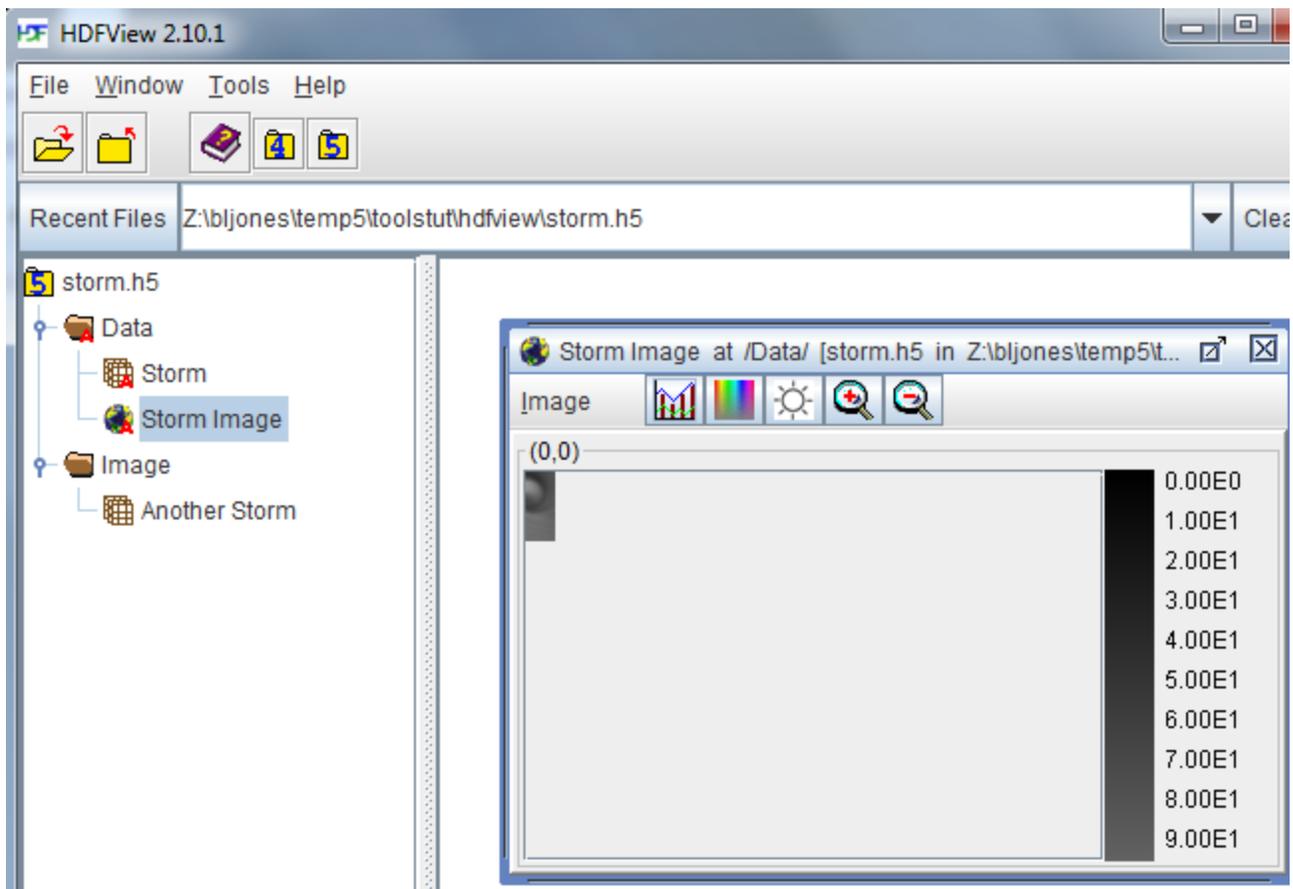
- Position the cursor in the middle of the image. Press Shift+Left Mouse button and hold, and then drag the mouse to select another subset.
- Select *Image->Write Selection to Image*. Enter *Subset* for the new image name. Click Ok. The *Subset* image will appear in the tree view on the left.
- Left double click on the image *Subset* to bring it up on the right.

The very small *Subset* image is displayed on the right side of the HDFView screen.
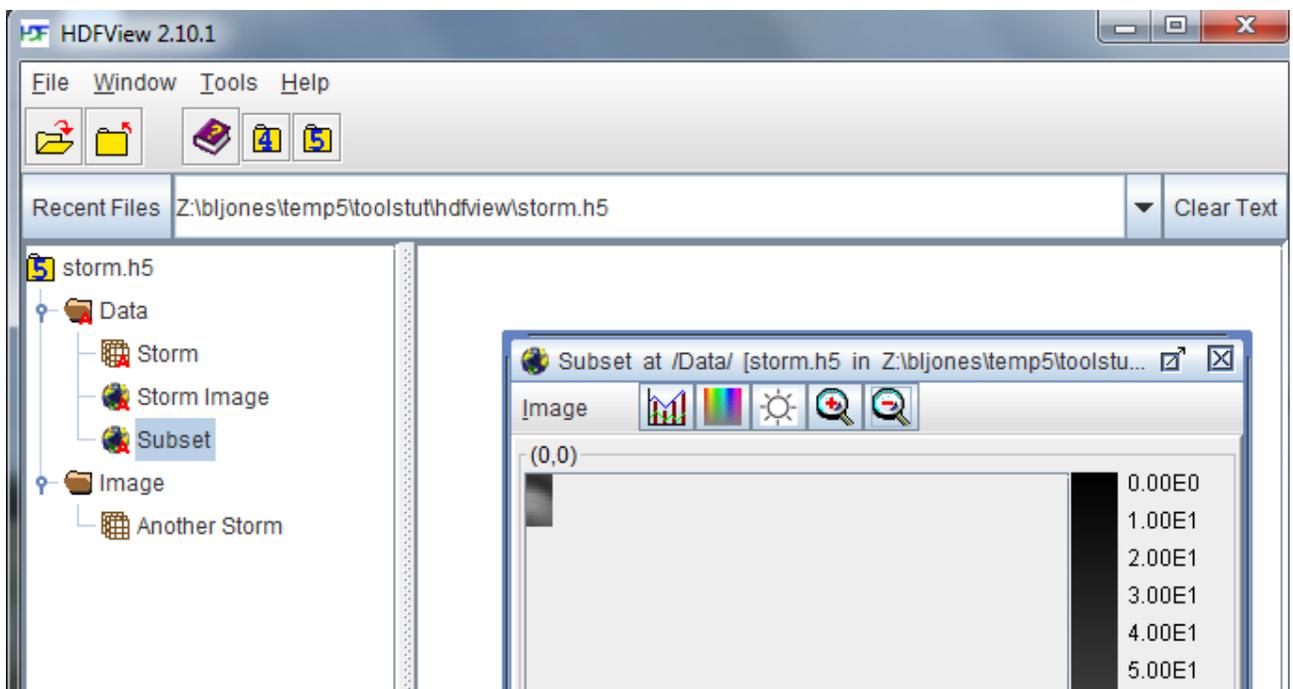

- Close the *Subset* image.

Creating a Table (Compound Dataset)

A dataset with a compound datatype contains data elements that consist of multiple fields. If the dataspace for the compound dataset is one-dimensional, then the dataset can be viewed as a table in HDFView, as is shown below.

- Right button click on the group *Data*. Select *New -> Compound DS*.
- A window pops up on the right side of the screen. Only fill in the following fields:

Dataset name	*Table*
Dataspace (Current size only)	4
Compound Datatype Properties: Number of Members	3
Compound Datatype Properties: *Name* / Datatype / Size	*Description* / string / 4 *Temperature* / float / 1 *Pressure* / double / 1

The window for creating a compound dataset is shown below.
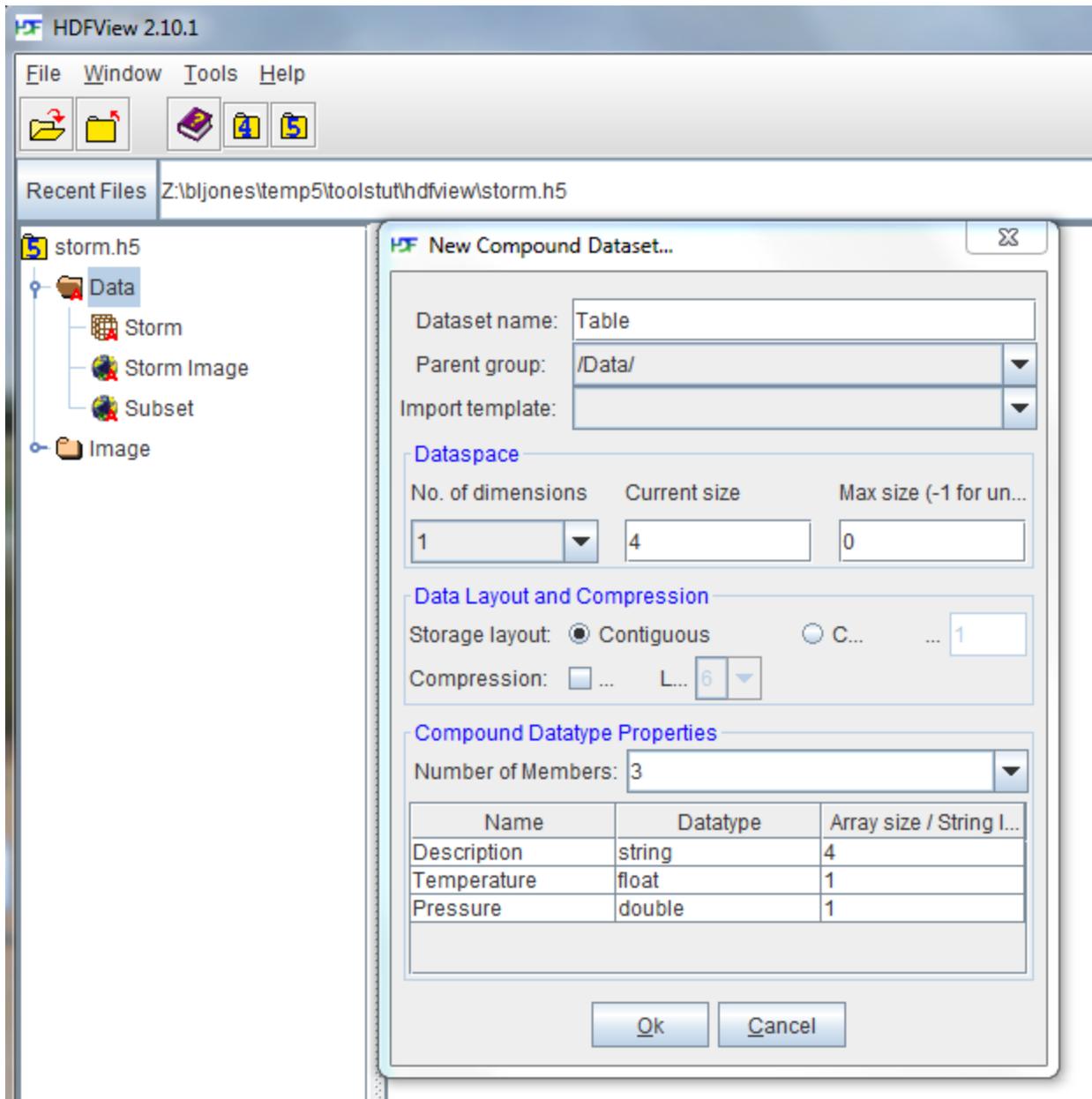
Click Ok at the bottom.
- Open the *Data* group (if it is not open) and double left click on the *Table* object.

The dataset will open in a window on the right. The fields in the dataset have no values written to them.

Close the dataset.