

Master Flat Maker

Introduction

MasterFlatMaker is a program to assist in image calibration for astrophotography. It is a single-purpose program, taking a collection of flat frames and combining them into a single “master flat frame” for use in calibration. It can be used via a graphic user interface or via a command-line interface. Precompiled binaries are available for Mac and Windows platforms. It is written in Python 3.8 and should run anywhere that Python can run.

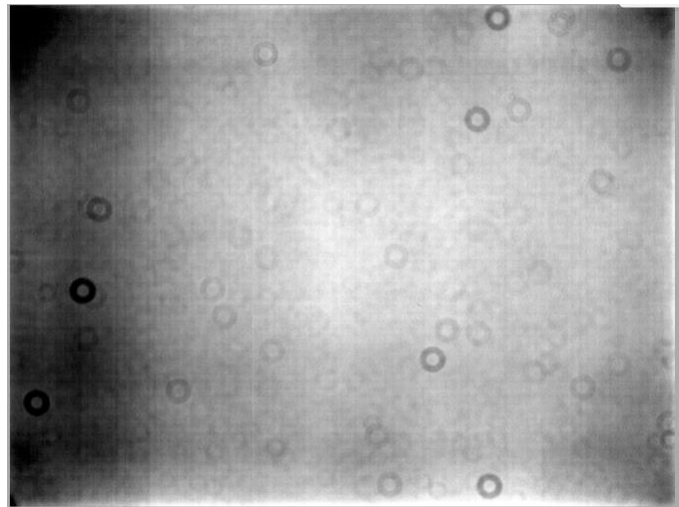
The program reads and produces only FITS files, not Jpegs or other image formats.

Foundation Concepts

Purpose of Flat Frames

Flat Frames are an important part of the post-processing of astrophotography images.

A flat frame is an exposure of a flat, evenly illuminated white or gray surface, with an exposure length that results in the pixels in the camera being exposed to about one-third to one-half of their capacity. The flat image should use the same filter and be at the same temperature as the images you plan to calibrate.



With the telescope pointed at a flat white surface, you would expect such a picture to be a uniform white or gray field, but it won't be. Instead it is a record of the imperfections in your optical system: irregularities in the focus of your lens, dust spots on your lens or mirror, dust spots on your camera or filter, etc.

In the example shown here, you can see

- Light fall-off at the corners of the field
- Uneven illumination
- “Dust Donuts” – specks of dust on the mirror or filter which, being out of focus, result in small circles

Image calibration software can use a flat frame to subtract such artifacts from your finished image.

Since dust spots can come and go, you would normally not save and reuse flat frames across multiple days but, rather, take fresh ones at the start or end of each imaging session.

Typically, you will take a large number of flat frames then combine them into a single “master” calibration frame. That combination of multiple flat frames into a single master flat is what MasterFlatMaker is for.

Precalibrating Input Flats

Unlike dark and bias frames, flat frames are actual images, so they need to be calibrated with dark frames just like ordinary images. You can provide a dark frame to MasterFlatMaker to subtract from the flats before they are combined, or have it search through a collection of dark frames to find the closest match. Since flat frames are often taken with very short exposure times, you can often use bias frames instead of dark frames for the calibration. Bias frames are, after all, just dark frames with an exposure length of zero. Ideally, however, you should use a dark frame with an exposure and temperature very similar to the flat frame being calibrated.

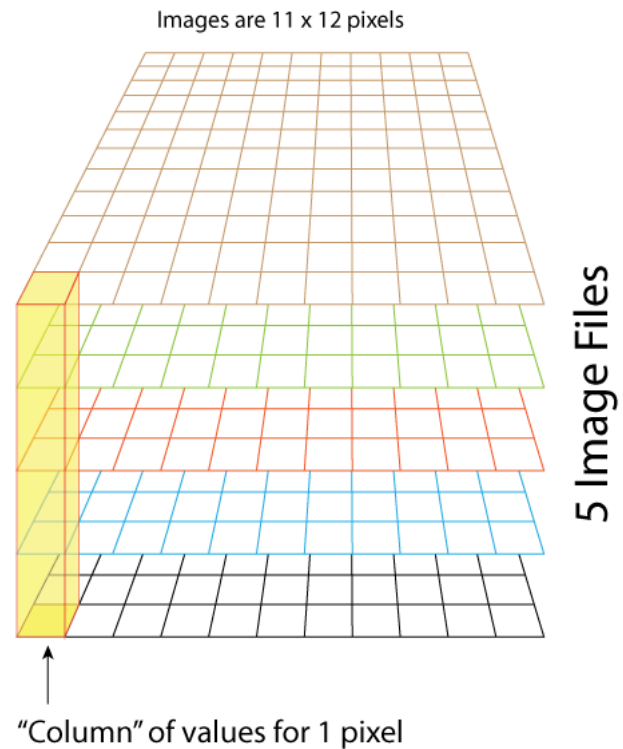
Combining Flat frames into a Master Flat

You should take a large number of flat frames, so you get an average-smoothed picture of what your optical imperfections look like. However, most image calibration software wants a *single* flat frame to use for image calibration, so you must combine your collected flats into a single “master flat”. (If your image calibration software allows you to provide a collection of flat frames, then it is just combining them into a master internally. You can save time by doing the combination once, rather than having your imaging software repeat the process every time.)

There are several algorithms for doing this combination.

Imagine that you have taken a number of flat frames – the more the better – all of the same dimension and binning as the images you will be calibrating. Think of the multiple frames as layers stacked on top of one another: for any given pixel in your master image, there is a stack of input pixel values at that position - one for each of the layered frames. Let's call this set of values for a given pixel a "column" of values.

For example, suppose you are capturing images that are 11 pixels wide and 12 pixels high. There are $11 \times 12 = 132$ pixels in each image. Suppose we have taken 5 flat frames for calibration. Each of them is also 11×12 pixels in dimension. So, you could think of the collection of flat frames as a collection of $11 \times 12 = 132$ columns, each with 5 values in it (the values from 5 flat frames).



Several algorithms are available to combine the frames into a single master, each with advantages and disadvantages. Choosing the best algorithm depends on circumstances such as the number of frames to be combined and the amount and type of noise in the frames.

The two most basic combination methods are:

Mean

"Mean" combination combines all the frames using a simple average. Each pixel in the resulting image is the average of the pixels in that column.

This method gives the best signal-to-noise ratio (SNR). However, outlier pixels (pixels that are considerably brighter or darker than the others in their column) in any of the input frames will affect the result so things like stray cosmic ray hits will show through.

Median

This combines the frames by taking the *median* value from each column. The median is the middle value if the values are sorted into ascending order (or the average of the two middle values if there is an even number of values).

This tends to reject outlier noise such as cosmic ray hits since they are usually much brighter than the naturally occurring pixels in that region and won't end up being the middle values. However, it produces a lower SNR than mean combination.

So, this method is a better choice for frames that have sporadic noise but that, overall, are not very noisy.

Two more advanced algorithms get close to the SNR of Mean while reducing the impact of random noise, close to the performance of Median. They are:

Min-Max Clipped Mean

This method drops the minimum and maximum values (i.e. all the instances of the minimum value and all the instances of the maximum value) from each column and then calculates the Mean of the remaining values. The dropping function can be repeated more than once: for example, you could drop the minimum and maximum values, and the next-to-minimum and next-to-maximum values, and so on.

This works well with noise that is either very bright (cosmic ray hits) or very dark. However, since you are actually throwing away some data, it requires that you have a large number of input frames to work well. At least 10 frames, and preferably many more.

It also can have a problem when a given column only contains one or two values. Throwing away the minimum and maximum throws out all the values, and then the algorithm has to back up and try throwing away less. This can be quite slow.

Sigma-Clipped Mean

This method also throws away outlier values and then takes the average of the rest, but it uses statistical techniques to determine what to throw away. It works as follows:

For each column we calculate the *mean* and the *standard deviation* of the values in that column. (*Standard deviation* is a common statistical measure that gives an indication of the amount of variability in a set of data.)

Then we calculate the *z-score* of each value in the column. The *z-score* is the distance of the value from the mean in multiples of standard deviation. So, a *z-score* of 2 means that the value in question is 2 standard deviations away from the mean.

Finally, we discard any values with a *z-score* above a given threshold and calculate the mean of the remaining values. For example, a threshold of 2 means “drop any values in a column that are more than 2 Standard Deviations from the mean of that column”.

This method works very well with a large number of input frames, and is the recommended method if you have more than about 10 frames.

In normally distributed data, setting the *z-score* threshold to 2.0 will reject about 5% of the data and keep about 95%. A lower *z-score* will reject more data (*z*=1.0 rejects about 32% and keeps about 68%), while a higher *z-score* will reject less data (*z*=3.0 keeps about 97%). Using this method requires a bit of experimentation. Start with a *z* threshold of 2.0 then reduce it with very noisy data or increase it with very clean data.

The *z* threshold is a floating-point number and should rarely be outside the range 1.0 to 3.0. A threshold of 2.0 will work well most of the time.

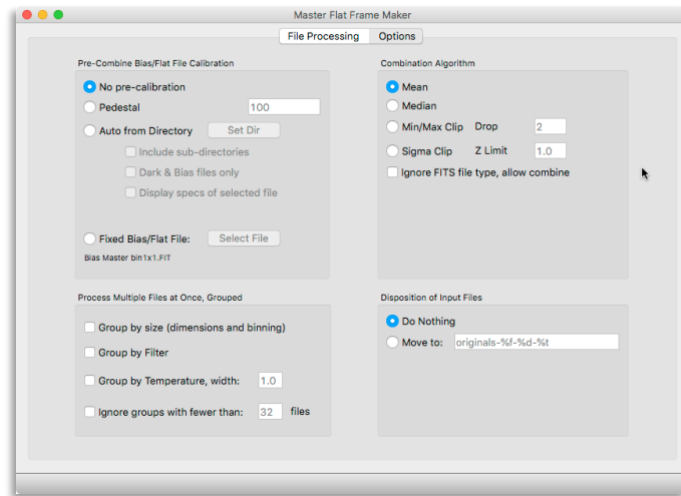
Two Modes of Application Operation

MasterFlatMaker can be used with an interactive graphic user interface (GUI), or from the command line with flags and filenames like a traditional UNIX command.

The command line option is intended to support scripting use of the program and combining it with other processes in your workflow. However, I recommend you start with the GUI to become familiar with the behavior of the program. The command line is less intuitive and does less error checking. The GUI is also the only way to modify the default settings in the program preferences.

Using the GUI

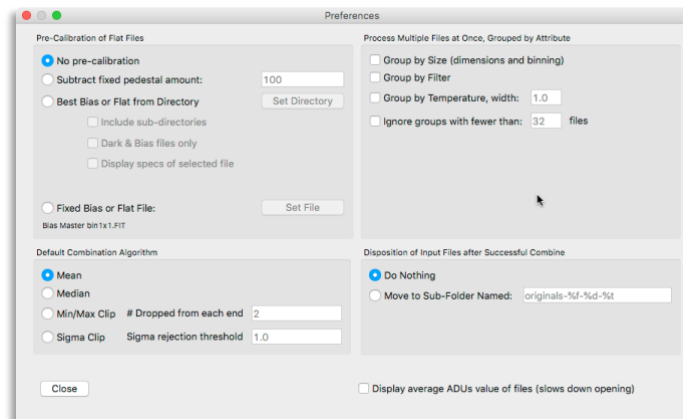
When you start the program as a Windows or Mac application, it will open in GUI mode, with the window shown here.



Preferences

Before we explore the main window, on your first use you should visit the Preferences window by selecting Preferences from the MasterFlatMaker (Mac) or File (Windows) menus.

The preferences window sets default values that are used when a new GUI session is started, and when the command line is used with a given setting not specified with a command line option.



The Preferences window is divided into four sections:

Pre-Calibration of Flat Files

In this section you specify if and how the flat files should be precalibrated, as discussed above. The last two options, "Automatic File from Directory" and "Fixed Dark or Bias File" both use a dark or bias frame. The difference is that "Automatic" will automatically pick the best-matching dark or bias file from a directory you specify, while "fixed file" has you specify the exact file to use.

Default Combination Algorithm

In this section you select the combination algorithm (Mean, Median, Min/Max Clip, or Sigma Clip) as described above. For the Min/Max and Sigma options you also specify the relevant numeric parameter, as described above.

Process Multiple Files at Once, Grouped by Attribute

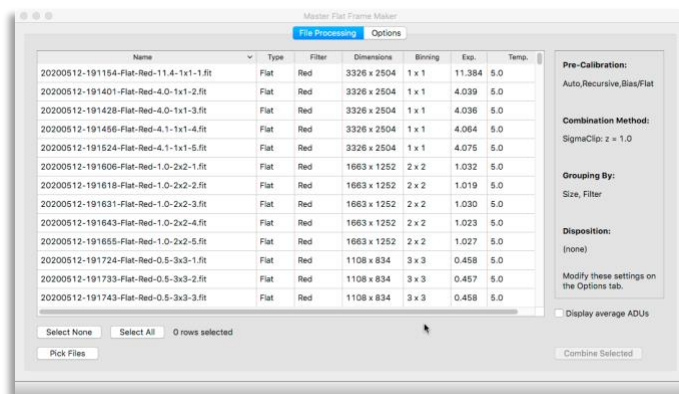
These settings allow the program to process a large number of flat frames into multiple outputs, grouped by size or other attributes. See the section “Processing Files in Groups”, below, for an explanation of these settings.

Disposition of Input Files

Optionally, MasterFlatMaker can move your input flat frames to a subfolder after they are processed. This can help keep files organized in certain workflows.

Main Window

The main window has a table, initially empty, that will list your Flat Frame files. There are a number of options that control the combination process. To save window space, these options are manipulated in the *Options* tab and summarized to the right of the file table on the *File Processing* tab.



Start by clicking “Pick Files” or by selecting Open from the File menu. In the dialog that opens, select all the flat frame files you will want processed. You will have a chance to further refine the list, so feel free to pick all the files in a folder if that is more convenient.

Note that the program must open each selected file to fetch the FITS metadata, so selecting a very large number of files may take a while. For example, on my computer, if I go into a folder containing 2,000 FITS files and do a “select all”, it takes about 45 seconds before the program responds with all the files listed. Be patient – it hasn’t died.

The selected files will be listed, along with some of their internal FITS metadata, in the file table. Click on column headings to sort the table by the various attributes. There is a checkbox to include each frame’s “average ADUs” figure in the table. This is the average exposure of the pixels in each image, which can help you verify that your flats are exposed properly. However, this option requires that each file be *read* to calculate the exposure level, so it will slow down the initial opening and display of the files.

Visit the Options tab to select your combination algorithm and other options, then go back to the File Processing tab and select one or more of your listed files to actually combine. Command-A or Control-A to quickly select all the files.

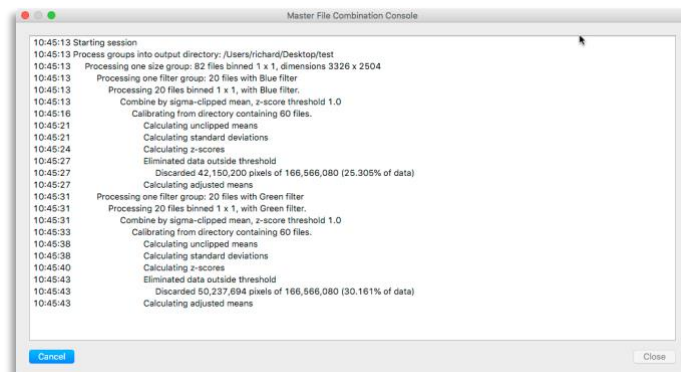
By default, the program will only allow you to select files whose FITS metadata says they are Flat files. (This is so you can just Command/Control-A to select all the files in a folder, then Command/Control-A to select all the files in this window, and you will end up with just the Flat frames, not any other stray files that happened to be in the folder.) However, some acquisition programs don't set that metadata correctly, so if you are certain the files you have picked are Flat files but MasterFlatMaker thinks they are not, you can click the "Ignore FITS file type" checkbox to bypass this check.

The "Combine Selected" button will be enabled once you have files selected and valid parameters chosen. If it is not enabled, something is not valid in your setup. Hovering over the disabled Combine button will give you a message explaining why it is disabled. The problem could be:

- No files selected (highlighted)
- Not enough files selected for the Min/Max method (there must be at least $2n + 1$ files for min/max parameter n).
- Selected files are not all flat frames and the "ignore" button is not checked.
- Selected files are not all the same size, unless you are processing by groups (see below).
- You have selected automatic calibration, but the directory of calibration files isn't specified or doesn't exist.

Once you have a valid plan, click the Combine button and you will be prompted for the name of your output file, which will then be created. Combining a large number of files from a high-pixel camera might take several seconds. (On my main computer, combining 32 8-megapixel files with the Sigma Clip method takes about 30 seconds.)

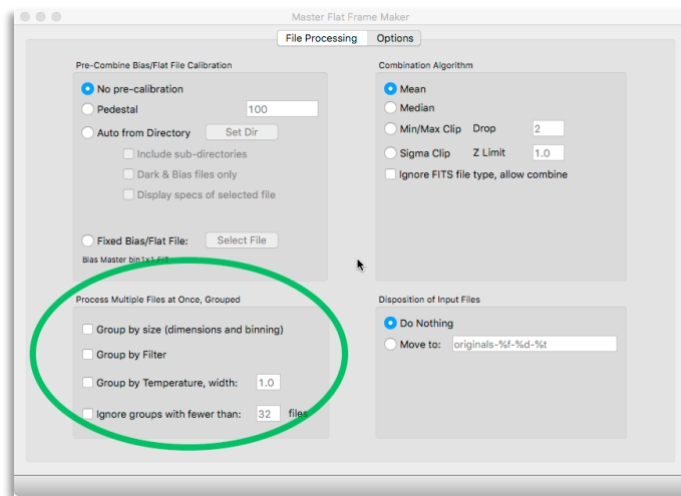
When using the GUI version, a console window will open during the combination operation, displaying progress results. This window also contains a Cancel button which will stop the combination task, in case you realize you forgot some important setting.



Processing Files in Groups

You may have a large set of Flat files in one directory but not want, or be able, to combine them all into a single master file. For example, they may be different binning levels, or using different filters.

On the Options tab, you can select “Group by size”, “Group by filter”, and “Group by temperature”. These are not mutually exclusive; you can select any combination of these settings. If selected, the program will process your files in groups and produce a separate master file for each group. Instead of prompting you for a single output file to produce, the program will prompt you for a *directory* to receive all the created master files.



Temperature grouping is different from size or filter grouping because, while image sizes and filter names are fixed values, it is quite possible for temperatures to vary by small amounts between frames. For example, suppose you have files with temperatures of:

0.1°
-9.9°
-15.07°
-0.1°
-10°
-15°
0°
-10.2°
-15.2°

You probably want that to be considered 3 temperature groups, not 9. Like this:

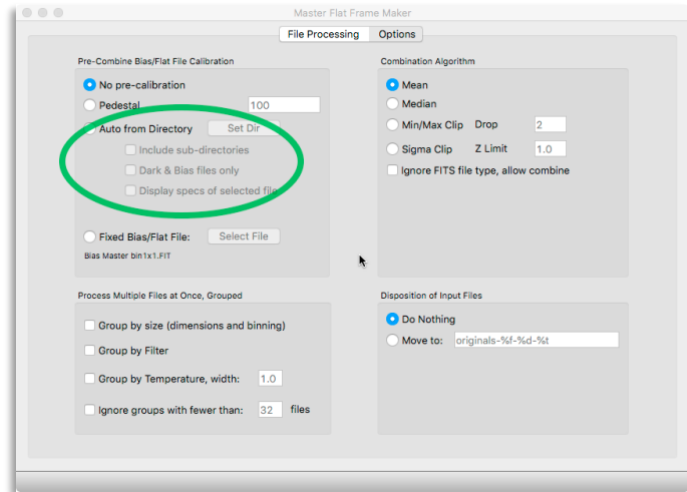
0.1°
-9.9°
-15.07°
-0.1°
-10°
-15°
0°
-10.2°
-15.2°

So, if you select temperature grouping, you must specify a width that the program will use to cluster the files into groups. A width of 1.0 (the default) will group files into groups where the temperature varies up to about 1 degree on either side of the group center. The width parameter can be between 0.1 and 50. The clustering uses statistical techniques, and the width parameter is not a “hard” limit — files outside the specified width may be included if the overall grouping result is better.

Automatic Selection of Precalibration Dark or Bias Files

If you are precalibrating your flat frames with a dark or bias frame, you might have a folder with several master dark and bias frames, taken with different binning levels and at different temperatures. The program can automatically select the best match from this library of dark and bias frames.

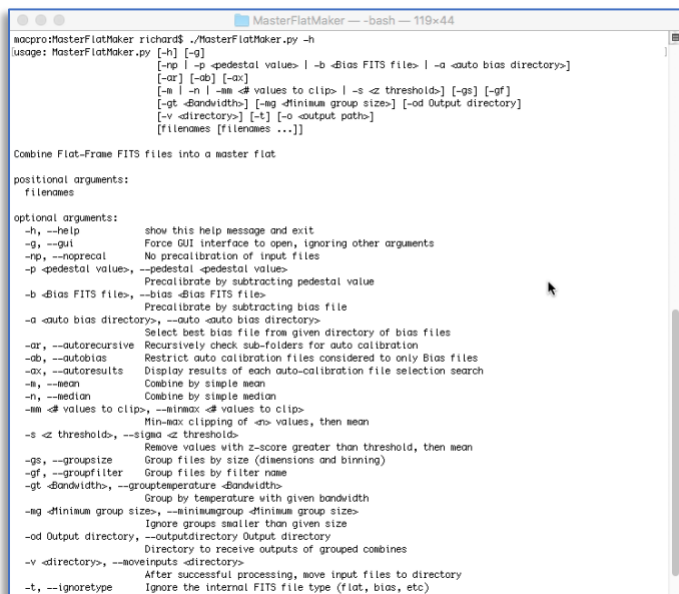
On the Options pane, select “Auto from Directory” as your precalibration method and click “Set Dir” to specify the directory containing all your dark and bias files. The program will then automatically select the best dark or bias file for each flat frame (the file with the correct binning and closest to the exposure time and temperature of your flat). If the folder contains files other than dark or bias files, you can click “Dark & Bias files only” to have the search consider only the dark and bias files.



Using the Command Line

To use command line mode, just run the program from your system's terminal or shell window (or, more commonly, from a shell script invoked from that window) and specify options and input files as command line arguments.

Run the program with the “-h” flag to get a brief summary of the available options. Every setting discussed in the GUI section above is available as a command-line option. If an important option is not specified on the command line, the value set in the Preferences will be used.



```
wooprot@MasterFlatMaker richard$ ./MasterFlatMaker.py -h
usage: MasterFlatMaker.py [-h] [-g]
                        [-rp] [-p <pedestal value>] [-b <Bias FITS file>] [-a <auto bias directory>]
                        [-ar] [-ab] [-ax]
                        [-m] [-n] [-m <# values to clip>] [-s <threshold>] [-gs] [-gf]
                        [-gt <bandwidth>] [-mg <minimum group size>] [-od <output directory>]
                        [-v <directory>] [-t] [-o <output path>]
                        [filenames [filenames ...]]

Combine Flat-Frame FITS files into a master flat

positional arguments:
  filenames

optional arguments:
  -h, --help            show this help message and exit
  -g, --gui             Force GUI interface to open, ignoring other arguments
  -rp, --noprca        No precalibration of input files
  -p <pedestal value>, --pedestal <pedestal value>
                        Precalibrate by subtracting pedestal value
  -b <Bias FITS file>, --bias <Bias FITS file>
                        Precalibrate by subtracting bias file
  -a <auto bias directory>, --auto <auto bias directory>
                        Select best bias file from given directory of bias files
  -ar, --autorecursive  Recursively check sub-folders for auto calibration
  -ab, --autobias       Restrict auto calibration files considered to only Bias files
  -ax, --autorecursive  Display results of each auto-calibration file selection search
  -m, --mean            Combine by simple mean
  -n, --median          Combine by simple median
  -mm <# values to clip>, --mm <# values to clip>
                        Min-max clipping of <#> values, then mean
  -s <threshold>, --sigma <threshold>
                        Remove values with z-score greater than threshold, then mean
  -gs, --groupsize      Group files by size (dimensions and binning)
  -gf, --groupfilter    Group files by filter name
  -gt <bandwidth>, --gt <bandwidth>
                        Group by temperature with given bandwidth
  -mg <minimum group size>, --minimumgroup <minimum group size>
                        Ignore groups smaller than given size
  -od <output directory>, --outputdirectory <output directory>
                        Directory to receive outputs of grouped combines
  -v <directory>, --moveinputs <directory>
                        After successful processing, move input files to directory
  -t, --ignoretype      Ignore the internal FITS file type (flat, bias, etc)
```

For common repeatable workflows, it is worth setting up a shell script to run the program via the command line. Since it doesn't have to pre-open all the files to populate the GUI window, it is faster.