

HOLLIS AKINS

# THE DATA ANALYSIS AND PHOTOMETRY PIPELINE

GUILFORD COLLEGE CLINE OBSERVATORY



# *Introduction*

THE GUILFORD COLLEGE CLINE OBSERVATORY is an astronomical observatory used primarily by students for research and as a tool for learning physics. It is operated by the Physics department at Guilford College, a small liberal arts college in Greensboro, North Carolina.

## *Observatory Equipment*

THE FACILITY HOUSES several 8-inch Meade telescopes, a 10-inch telescope for spectroscopy, and a 16-inch telescope for optical CCD imaging. Our 10-inch telescope and spectroscopic imaging system were set up by Sean Kirwan '18 as a thesis project. Both the 10-inch and 16-inch telescopes are permanently mounted, and the 16-inch is remotely controllable. Much of the student work done at the observatory uses this 16-inch telescope, which at our level of light pollution allows us to study galaxies, nebulae, and stars down to roughly 13th magnitude with no problem.

## *Project Goals*

THERE ARE THREE primary methods for astronomers to look at the sky: direct imaging, spectroscopy, and photometry. Direct imaging and spectroscopy tools are available at our observatory for students to use and apply to their projects as needed; however, photometry has remained a slow and tedious process, requiring students to manually create apertures in CCD images.

The primary goal for this project was to develop a system in which we could perform more thorough and sophisticated photometric observations in research projects. Photometric automation would allow us to more easily study brightness variations in stars over time and the systematic limitations of our observatory.<sup>1</sup> Furthermore, these tools would allow future students to work with data gathered



Figure 1: Exterior view of the Guilford College Cline Observatory

<sup>1</sup> We can study the limitations of our observatory by learning what the fundamental noise limit in our photometric measurements is, and determining the limiting magnitude of our telescope.

over long periods, which would be tedious to gather by hand. This further helps make possible long-term photometry projects to be handed off from one student to the next. This possibility for one more level of abstraction in photometry allows for more thorough student research.

THIS DOCUMENT IS intended to be a summary of the work done over the summer of 2018, the preliminary results gathered, and the future work intended. This is not intended to be a technical discussion of the astrophysical implications of photometric observations, nor a technical report on the software development. Largely, the goal of this document is to describe my work, why it matters, and how it will function to help future students.

# Overview of the Software

ALL OF THE software detailed in this section lives on a Linux server and runs once a day on each image captured the previous night. The entirety of the data analysis pipeline was written in Python.

## *Programatic Image Calibration*

THE FIRST STEPS to processing new CCD images is to calibrate them using existing or new dark, bias, and flat-field frames. This is handled in the pipeline using Python's `astropy` module.

## *Making Master Calibration Frames*

Before we can calibrate any new CCD images, we must have master calibration frames to use—these are composites of single calibration frames, intended to remove any outliers and reduce noise levels. The pipeline scans the `days` directory for any new calibration frames. Any that it finds are calibrated relative to each other and combined via the median to remove outliers.<sup>2</sup> The resulting master calibration frames are then written to a new `.FITS` file.

We maintain a directory of darks, flats, and biases for each filter at binning `1x1`, and for narrow-band photometric filters at binning `2x2` and `3x3`.<sup>3</sup> The files in this directory can be overwritten by the pipeline, but the user must specify before the daily run that they want new calibration masters to overwrite the old ones. This is so we can maintain a set of well-exposed frames with good statistics and low noise, which won't be overwritten by a random frame taken carelessly.

## *Calibrating CCD Images*

We use python's `astropy` module for opening the data arrays in `.FITS` files and for writing arrays to new `.FITS` files. Each image that is captured is calibrated if and only if it is deemed worthwhile—a

<sup>2</sup> All bias frames are combined via the median, to form a bias master. The bias master is subtracted from each dark frame, which are subsequently combined via the median to form a dark master. Finally, the dark master is subtracted from each flat field, and the flat fields are combined via the median and normalized to 1.

<sup>3</sup> Our filter wheel includes wide-band Red, Green, Blue, and Lum, as well as narrow-band V, B, R, and Halpha filters.

determination that is made based on several factors including image size and sensor temperature at capture. If we determine the image worthy of calibration, the program checks the FITS header to determine the appropriate filter, binning, and exposure times to use in the calculation. Images are calibrated according to the following formulas:

$$\text{bias\_corrected\_image} = \text{image} - \text{bias\_master} \quad (1)$$

$$\begin{aligned} \text{dark\_corrected\_image} &= \text{bias\_corrected\_image} \\ &\quad - (\text{exptime}/\text{dark\_exptime}) * \text{dark\_master} \end{aligned} \quad (2)$$

$$\text{final\_image} = \text{dark\_corrected\_image}/\text{flat\_master} \quad (3)$$

The final image is written to a new FITS file and stored in a separate directory.

### *Ensemble Photometry*

PHOTOMETRIC MEASUREMENT IS perhaps the most important step in the pipeline. We use a method called *ensemble photometry*, a form of differential aperture photometry that uses every star in the image to calibrate the instrumental magnitude of every other star. In other words, it uses every star in the image to calculate a *frame zero-point*, to which every instrumental magnitude is adjusted.

### *Plate Solution*

In order for the pipeline to perform aperture photometry in an automated fashion, it needs to know where in the sky each image was captured. The pipeline relies on the user to plate solve each image they want to be processed—this simplifies our program and gives the user control over which images go through the photometry process. For example, a user capturing images of a diffuse galaxy likely wouldn't have any need for photometric data—they also would be unlikely to plate solve their images manually.

We use the image link feature included in TheSkyX Pro to find an astrometric solution for our images. It writes a World Coordinate System (WCS) matrix into the FITS header for each image successfully plate solved, which we access using the `astropy.wcs` module. Using this module along with the python wrapper for the SExtractor program (`sep`), we gather a list of equatorial coordinates for each star detected in the image.

### *Aperture Photometry*

We use the SExtractor python wrapper `sep` to perform aperture photometry and annulus subtraction. The aperture radius is fixed at 30 pixels for binning 1x1. The radius is adjusted by a factor the  $x$  binning as the binning changes, so the radius for an image of binning 2x2 would be 15 pixels. We use a fairly large annulus ranging from 1.5 to 2.0 times the aperture radius, with an algorithm that removes any outlier pixels from the annulus before performing background subtraction. This ensures that any stars that show up in the annulus are not subtracted as background.

For each star in the image, we subtract out the background and sum the pixel values in the aperture to generate a flux value. This is then put on a logarithmic scale by the formula

$$\text{instmag} = -2.5 \log_{10} \text{flux} \quad (4)$$

We are then left with a list of instrumental magnitudes for each star in the image.

### *Catalog Comparison and Zero-Point Offset*

Now, with corresponding lists of equatorial coordinates and instrumental magnitudes, we programmatically look up each star in the catalog using the VizieR tool from the `astroquery` module. We search within a 3 arcsecond radius of the target coordinates for a catalog entry, and store the resulting filter-specific magnitude in a list of catalog magnitudes. The zero point is calculated based on these two lists by

$$\text{zero\_point} = \text{median}(\text{cmags} - \text{imags}) \quad (5)$$

where `cmags` and `imags` are arrays. The calibrated magnitudes are calculated by

$$\text{mags} = \text{instrumental\_mags} + \text{zero\_point} \quad (6)$$

The final data, including star ID, measured and catalog coordinates, measured and catalog magnitudes, and other variables are written to a `.csv` file.





# Preliminary Results

## Variable Stars

Figure 2 and 3 show the results of the first photometric observations processed through the data pipeline: light curves for two variables stars. Figure 2 shows a light curve for V0457 Lac, a pair of stars in an eclipsing binary system. We can see in this curve a definite dip in the brightness of the system, from an average catalog magnitude of  $\approx 9.85$  to a minimum of  $\approx 10.05$ . Figure 3 shows a light curve for RR Lyrae, the eponym for the RR Lyrae class of variable stars. We see the star rise magnitude  $\approx 8.2$  to  $\approx 7.5$  before clouds obscured the image around 06:30 GMT.

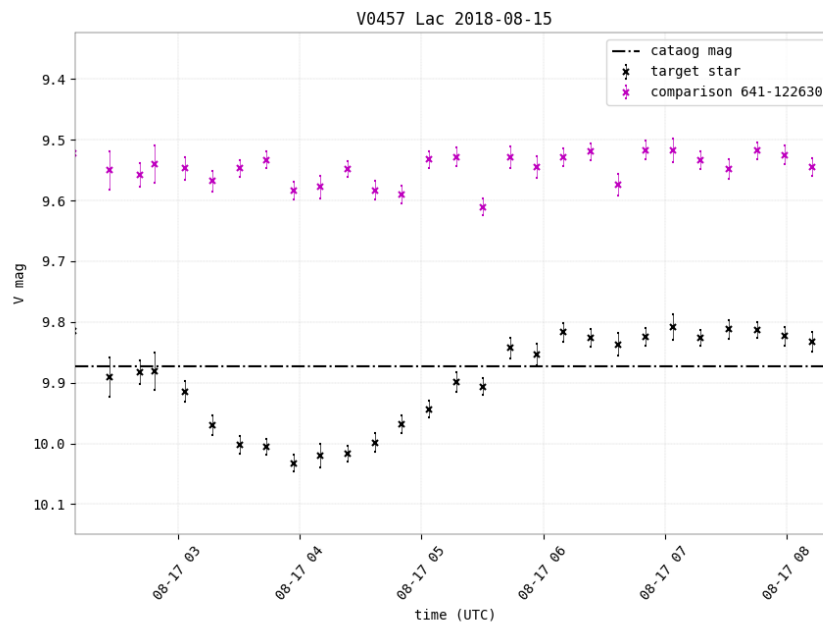


Figure 2: V047 Lac, an eclipsing binary

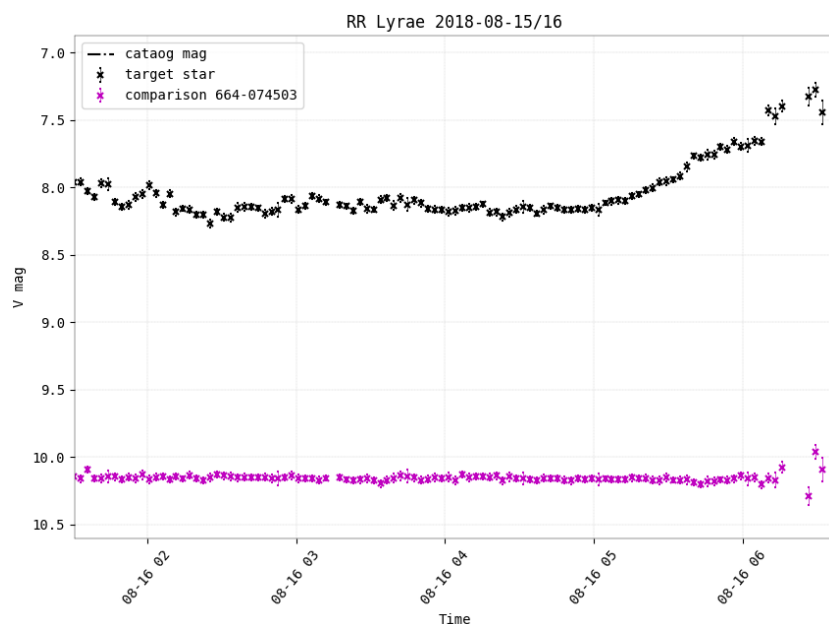


Figure 3: RR Lyrae captured on August 15, 2018

## Stellar Evolution

Figure 4 shows a plot of the apparent V band magnitude vs the color (B-V) for stars in the open cluster Messier 34. This type of plot is functionally similar to a Hertzsprung-Russel (HR) diagram, in which luminosity is plotted vs temperature. Luminosity and temperature are absolute quantities which we can only determine if the distance to the star is known. Since all of the stars in a specific cluster are approximately the same distance from us, their magnitudes relative to each other are not impacted by their relative distances.

We can see in this plot a fairly well-defined main sequence, highlighted in green. We also see several stars that fall outside of the main sequence, highlighted in red, which are classified as red giants. This plot was generated via a python script which searches for stored data points within 7 arcseconds of the specified target coordinates.<sup>4</sup> The pipeline requires only two CCD images of a cluster in order to generate the data required for these plots: one in V and one in B.

This ease of use allows for the possibility of extensive studies of stellar evolution in open and globular clusters, with the primarily limitation being resolving power and limiting magnitude.

Richmond, Michael. *Interpreting the HR diagram of stellar clusters*. <<http://spiff.rit.edu/classes/phys230/lectures/clusters/clusters.html>>

<sup>4</sup> This search radius can—and will need to—be adjusted according to the object in question.

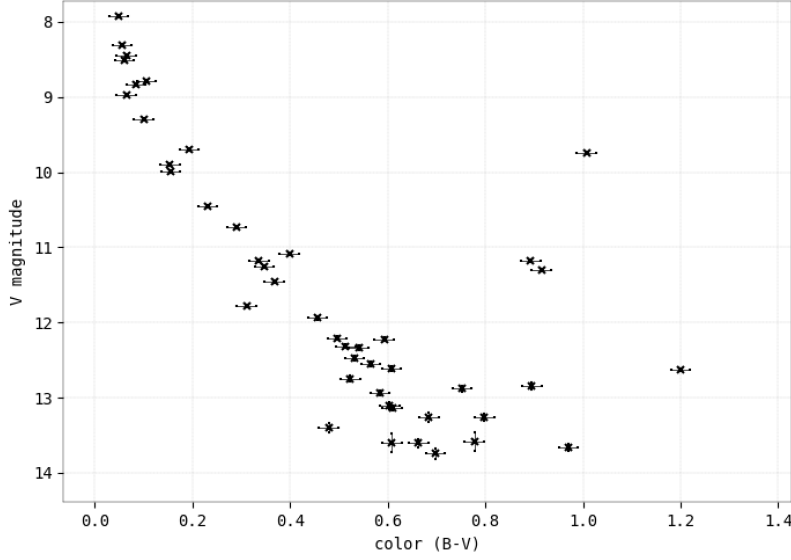


Figure 4: Plot of the apparent V mag vs color for stars in the open cluster Messier 34.

### *Future Work*

While the majority of the software development is complete, the capabilities of the data analysis pipeline have not yet been utilized to their potential. We hope to use the pipeline for several projects:

- calculations of variable star periods and the stellar physics associated with the type of star
- further studies of globular and open cluster evolution
- studies of occultations of asteroids or planets<sup>5</sup>
- a systematic study of the limiting magnitude of our facility

We hope to continue to develop the software to be as portable, error-proof, and useful as possible. A major step in this process would be a revamp of the data management structure, which at present is clunky and makes accessing the data difficult and discouraging.

### *Acknowledgements*

I would like to thank Don Smith for his mentorship in bringing this project to fruition, and for constantly pushing me to work my hardest. I would also like to thank Donald J. Cline for his contributions and support of our observatory, as well as the family of Rexford Adelberger for their contributions to departmental funding for student research.

<sup>5</sup> The initial completion of our pipeline software occurred just a week before the August 2018 Pluto occultation. It was our intent to observe this event and process photometric data through this software, but it was obscured by clouds only an hour before the event.